Journal of Statistics & Management Systems
ISSN 0972-0510 (Print), ISSN 2169-0014 (Online)
Vol. 27 (2024), No. 8, pp. 1499–1523
DOI : 10.47974/JSMS-981

A reduced-bias weighted least squares estimation of the extreme value index

By E. Ocran [§]
R. Minkah [†]
K. Doku-Amponsah *
Department of Statistics and Actuarial Science
University of Ghana
Box LG 115
Legon Accra
Ghana

Abstract

In this paper, we propose a reduced-bias estimator of the EVI for Pareto-type tails (heavy-tailed) distributions. This is derived using the weighted least squares method. It is shown that the estimator is asymptotically unbiased, asymptotically consistent and asymptotically normal under the second-order conditions on the underlying distribution of the data. The finite sample properties of the proposed estimator are studied through a simulation study. The results show that it is competitive to the existing estimators of the extreme value index in terms of bias and Mean Square Error. In addition, it yields estimates of $\gamma > 0$ that are less sensitive to the number of top-order statistics, and hence, it alleviate the problem of selecting an optimal tail fraction to some extent. The proposed estimator is further illustrated using practical datasets from pedochemical and insurance.

Subject Classification: *(2020) 62E20, 62G32, 62J05, 60F10.*

Keywords: *Extreme value theory, Extreme value index, Weighted least squares, Large deviations, Weak law of large numbers, Limit theorems.*

1. Introduction

Statistics of extremes deals with the estimations of the occurrences of rare events. Such events include high quantile, exceedance probability and return periods. The knowledge of the frequency and magnitude of these

[§] *E-mail:* **eocran@ug.edu.gh**
[†] *E-mail:* **rminkah@ug.edu.gh**
* *E-mail:* **kdoku-amponsah@ug.edu.gh (Corresponding Author)**

events help in planning to mitigate the effects of their occurrences. The primary parameter in estimating these events is the extreme value index (EVI), γ (or also known as the tail index). The extreme value index, either implicitly or explicitly, plays a vital role in estimating other extreme events [8, 15, 26]. Hence in statistics of extremes, the primary focus is the estimation of the EVI, γ. The EVI is crucial in extreme value analysis because all inference requires that the EVI is known. The EVI regulates the tail heaviness or flatness of the extreme value distribution, assuming the max-domain of attraction condition holds for the underlying distribution, *F*. We may have $\gamma \in \mathbb{R}$ and the heaviness of the tail function $\overline{F} := 1 - F$, increases with γ. However, in this paper, we shall consider the particular case where $\gamma > 0$. That is, we shall work in the domain of attraction for maxima of an extreme value distribution function (df),

$$H_\gamma(x) = \exp(-(1+\gamma x)^{-1/\gamma}), 1+\gamma x \geq 0. \tag{1}$$

This class of distributions has been found to be useful in diverse fields, including but not limited to insurance [31, 32, 38], finance [20, 26, 27], economics [16, 28, 34], telecommunication [18, 30] and Climatology [36, 40]. For $\gamma > 0$, the underlying distribution, *F* belongs to the Pareto-type distributions with the upper tail function expressed as

$$\overline{F} = x^{-1/\gamma} \ell_F(x);\ x > 1, \tag{2}$$

or has a tail quantile function defined by

$$U(x) = x^\gamma \ell_u(x) x > 1, \tag{3}$$

where ℓ_F and ℓ_u are slowly varying functions at infinity, that is,

$$\frac{\ell_{\mathcal{B}}(tx)}{\ell_{\mathcal{B}}(x)} \to 1, \text{ as } x \to \infty, \text{for all } t > 0,\ \mathcal{B} \in \{F, u\}. \tag{4}$$

Using the concept of the *de Bruyn conjugate*, it can be shown that ℓ_F and ℓ_u are linked [4].

Let $\chi = \{X_j\}_{j=1}^n$ denote a sequence of independent and identically distributed random variables with distribution function, *F*, belonging to the class of distributions defined by (2). Again, let $X_{j,n}$ denote the j^{th} order statistic associated with the sample χ. Define the weighted log-spacings

$$Z_j = j \log\left(\frac{X_{n-j+1,n}}{X_{n-j,n}}\right), 1 \leq j \leq k. \tag{5}$$

[17] and [1] demonstrated that the effect of bias can be accommodated by modelling the weighted log-spacing of the order statistics and also proposed the usage of a regression model on Z_j with $\left(\frac{j}{k+1}\right)^{-\rho}$ as the explanatory variable. It can be seen from [1] and [32] that the $Z_j's$ can be approximated with a regression model with exponential responses:

$$Z_j = \left(\gamma + b_{n,k}\left(\frac{j}{k+1}\right)^{-\rho}\right) f_j, 1 \le j \le k, \tag{6}$$

where $b_{n.k} = b((n+1)/(k+1)) \to 0$ as $k, n \to \infty$, $\rho < 0$ is a second-order parameter, and $f_j's$ follow exponential distribution with a unit mean. The authors stated that this representation permits the generation of reduced-bias estimates of γ. In the case of strict Pareto distributions, (6) reduces to

$$Z_j = \gamma f_j, j = 1, \ldots, k, \tag{7}$$

and the estimator of γ in this case is the usual Hill estimator [23].

Different estimation techniques have been proposed for estimating γ (see for example, [2, 6, 9, 11, 13, 23, 37]). The Hill estimator [23] till date plays a vital role in many applications and it is the most widely used estimator in practice. Using the Renyi's representation, the [23] estimator can be expressed as the mean of the weighted log-spacings,

$$\gamma_k^H = \frac{1}{k}\sum_{j=1}^{k} Z_j, \tag{8}$$

A desirable property of the Hill estimator is that, it has a minimal variance and asymptotically normal [12, 29]. A drawback which makes the usage of the Hill estimator in application difficult is its sensitivity to the number of top order statistics, *k*. This drawback is common to most semi-parametric estimators currently in existence. Any estimator which is sensitive to the value of *k* may be challenging to use in application since a little change in the value of *k* may result in large change in the value of the extreme value index. In addition it makes it challenging to select the optimal *k*.

Since the Hill estimator is obtained with $\ell_F(x) = 1$ in (2), in order to obtain a reduced-bias estimator, a functional form of ℓ_F is needed. Therefore [1] assume that for the Hall class of distributions [42] the slowly varying part ℓ_F satisfies a second-order condition :

Assumption (R_{ℓ_F}): There exists some auxiliary function $b > 0$ that is regularly varying with index $\rho < 0$ satisfying $b(x) \to 0$ as $x \to \infty$, such that for all $\lambda \ge 1$, as $x \to \infty$

$$\log \frac{\ell_F(\lambda x)}{\ell_F(x)} \approx b(x)k_\rho(\lambda) \tag{9}$$

with $k_p(\lambda) = \frac{\lambda^\rho - 1}{\rho}$. That is, under assumption R_ℓ, we assume the slowly varying part has a structure so we do not ignore it and hence reduces the bias accompanying the approximation of ℓ_F as a constant.

[3] further approximated (7) under assumption R_ℓ as

$$Z_j = \gamma + b_{n,k}C_j + \varepsilon_j \tag{10}$$

where $b_{n,k} = b(k/n) \to 0(k, n \to 0)$ is the slope, $C_j = (j/(k+1))^{-\rho}$ is the covariate and γ is the intercept of the regression line. The authors showed that the error term, ε_j is asymptotically normal with variance, γ^2, and mean, 0. The regression model in (10) has independent exponential responses [1]. Therefore, for fixed *j*, the Z_j's are approximately independent exponentially distributed with mean

$$\mu_j = \mathbb{E}(Z_j) = \gamma + b_{n,k}C_j. \tag{11}$$

Note that $1/(k+1)^{-\rho} \le C_j \le 1$ for every *j*, and hence, it follows that

$$\gamma + \frac{b_{n,k}}{(k+1)^{-\rho}} \le \mu_j \le \gamma + b_{n,k} \text{ for every } j \in \{1, 2, \ldots, k\}. \tag{12}$$

This paper seeks to propose a reduced-bias estimator that yields stable estimates over the number of top order statistics, *k*. It is worth mentioning that [9] has proposed the ridge regression as another reduced-bias estimator, which yields much more stable estimates compared to other existing estimators. However, our proposal is an alternative estimator that uses weighted least s on the $Z_j's$ in (10) to estimate γ and $b_{n,k}$ jointly. The parameter ρ is estimated externally using the minimum variance method introduced by [9].

The concept of estimating the extreme value index using a weighted least squares estimator is not new in the field of extreme value theory. Researchers like [6], [39], [24] and [25] have used the weighted least squares method to estimate the extreme value index under different conditions. First, [6] fitted a weighted least squares estimator to the tail of the Pareto quantile plot and estimated γ as the slope of the linear model. Second, [39] linearised (3) and then fitted weighted least squares to their error sum of squares. Third, [24] also employed the weighted least squares to estimate the extreme value index for heavy-tailed distributions whose slowly varying part is given as $\ell(x) = a(1 + bx^{-\beta})$, where $\beta > 0$ and $a, b \in \mathbb{R}$. Lastly, [25] estimated γ by applying the weighted least squares method to gamma

$$\log\left(\frac{X_{n-i,n}}{X_{n.k,n}}\right)=\gamma\log\frac{k+1}{i+1},$$

where $X_{n-i,n}=F_n^{\leftarrow}\left(1-\frac{i}{n}\right), i=0,1,\ldots,k.$

Our proposed methodology differs from the existing studies in terms of;

i. the definition of the weight function, and

ii. the form of the regression model.

Hence, to the best of our knowledge, our proposed methodology has not been considered in literature.

The rest of the paper is organised as follows. The reduced-bias weighted least square estimator and its asymptotic properties with their proofs are presented in section 2. In section 3, the proposed estimator's performance is compared to other existing estimators via a simulation study and a practical illustration. Lastly, in Section 4, we provide concluding statements of the study.

2. Materials and Methods

2.1 *The proposed estimator*

To jointly estimate γ and $b_{n,k}$ in (10) via our proposed weighted least squares method, we define the loss function for the weighted least squares estimator as

$$L_k(\gamma,b_{n,k};W)=\sum_{j=1}^{k}W_j(Z_j-\gamma-b_{n,k}C_j)^2, \tag{13}$$

for $j\in\{1,2,\ldots,k\}$. The weight function is defined as

$$W_j=1-\frac{j}{k+1}, 1\le j\le k, \tag{14}$$

where $W_j\in(0,1)$ and decreases linearly with respect to *j*. We minimise (13) with respect to γ and $b_{n,k}$ to obtain the following Weighted Least Squares (WLS) estimators:

$$\hat{b}_{n,k}=\frac{\Sigma_{j=1}^{k}\tilde{W}_j(C_j-\Sigma_{j=1}^{k}\tilde{W}_jC_j)Z_j}{\Sigma_{j=1}^{k}\tilde{W}_jC_j^2-(\Sigma_{j=1}^{k}\tilde{W}_jC_j)^2} \tag{15}$$

and

$$\hat{\gamma}^{+}_{wls} = \sum_{j=1}^{k} \tilde{W}_j Z_j - \hat{b}_{n,k} \sum_{j=1}^{k} \tilde{W}_j C_j, \tag{16}$$

where

$$\tilde{W}_j = \frac{W_j}{\sum_{j=1}^{k} W_j}, 1 \le j \le k \tag{17}$$

and $C_j = (j/(k+1))^{-\rho}$. Here, $\tilde{W}_j$ is normalised and sum up to 1, (i.e., $0 \le \tilde{W}_j \le 1$ and $\sum_{j=1}^{k} \tilde{W}_j = 1$) and ρ is a second-order parameter which would be estimated externally.

The parameter ρ plays an essential role when dealing with optimisation in extreme value analysis. The speed of convergence of the limiting distribution of the extremes and also the asymptotic normality of estimators of γ is regulated by (9). It can be observed that the rate of convergence of (9) is determined by ρ. The estimation of ρ is also vital in controlling the bias component of most EVI estimators (see [1, 17, 21, 22], among others). Therefore, it is important to have a good estimator for ρ for an efficient adaptive choice of the optimal k for EVI estimators. The most common estimator of ρ, which works well both in theory and in practice, was proposed by [19]. [9] have also proposed the minimum variance approach, which also performs well. In the simulation study and the practical illustration, we consider the minimum variance approach and the [19] estimator. However, we present the results for the minimum variance method since that yields much more stable EVI estimates for the practical datasets.

2.1.1 Properties of the Proposed estimator

In this section we demonstrate that, the proposed weighted least squares estimator, $\hat{\gamma}^{+}_{wls}$ is asymptotically unbiased, asymptotically consistent and asymptotically normal.

First, we compute the asymptotic mean square error (AMSE) of the estimator, $\hat{\gamma}^{+}_{wls}$. It is well known that the mean square error (MSE) can be decomposed into a variance term and a bias term. The bias is defined as $\mathbb{E}(\hat{\gamma}^{+}_{wls}) - \gamma$, the distance between the estimator's expected value and the parameter γ. An estimator is said to be unbiased if the bias is 0, in which case the MSE is just the variance of the estimator. The expected value of the estimator $\hat{\gamma}^{+}_{wls}$ is given by

$$
\begin{aligned}
\mathbb{E}(\hat{\gamma}^{+}_{wls}) &= \sum_{j=1}^{k} \tilde{W}_j \mathbb{E}(Z_j) - \hat{b}_{n,k} \sum_{j=1}^{k} \tilde{W}_j C_j \\
&= \sum_{j=1}^{k} \tilde{W}_j (\gamma + \hat{b}_{n,k} C_j) - \hat{b}_{n,k} \sum_{j=1}^{k} \tilde{W}_j C_j \\
&= \gamma \sum_{j=1}^{k} \tilde{W}_j \\
&= \gamma.
\end{aligned}
$$

Therefore, the asymptotic bias of $\hat{\gamma}^{+}_{wls}$, $Abias(\hat{\gamma}^{+}_{wls})$ is

$$
\begin{aligned}
Abias(\hat{\gamma}^{+}_{wls}) &= \mathbb{E}(\hat{\gamma}^{+}_{wls}) - \gamma \\
&= 0.
\end{aligned}
$$

Since $\hat{\gamma}^{+}_{wls}$ is asymptotically unbiased, the AMSE is the same as the asymptotic variance (Avar) of the estimator, which we proceed to find. Observe from (15) and (16) that we can also rewrite $\hat{\gamma}^{+}_{wls}$ as

$$
\hat{\gamma}^{+}_{wls} = \sum_{j=1}^{k} \tilde{W}_j \left(1 + \frac{S_1^2 - S_1 C_j}{S_2}\right) Z_j,
$$

where $S_1 = \sum_{j=1}^{k} \tilde{W}_j C_j$ and $S_2 = \sum_{j=1}^{k} \tilde{W}_j C_j^2 - (\sum_{j=1}^{k} \tilde{W}_j C_j)^2$. Now, let $\dot{S} := \sum_{j=1}^{k} \tilde{W}_j^2 (S_1 - C_j)$ and $\ddot{S} := \sum_{j=1}^{k} \tilde{W}_j^2 (S_1 - C_j)^2$ and observe that we have

$$
\begin{aligned}
Avar(\hat{\gamma}^{+}_{wls}) &= \sum_{j=1}^{k} \tilde{W}_j^2 \left(1 + \frac{S_1^2 - S_1 C_j}{S_2}\right)^2 Var(Z_j) \\
&= \gamma^2 \sum_{j=1}^{k} \tilde{W}_j^2 \left(1 + \frac{S_1^2 - S_1 C_j}{S_2}\right)^2 \\
&= \gamma^2 \left(\sum_{j=1}^{k} \tilde{W}_j^2 + \frac{2S_1}{S_2} \sum_{j=1}^{k} \tilde{W}_j^2 (S_1 - C_j) + \frac{S_1^2}{S_2^2} \sum_{j=1}^{k} \tilde{W}_j^2 (S_1 - C_j)^2 \right) \\
&= \gamma^2 \left(\sum_{j=1}^{k} \tilde{W}_j^2 + \frac{2S_1 \dot{S}}{S_2} + \frac{S_1^2 \ddot{S}}{S_2^2} \right) \\
&= \gamma^2 \left(\frac{4}{3k} + \frac{2S_1 \dot{S}}{S_2} + \frac{S_1^2 \ddot{S}}{S_2^2} \right) + o(1/k),
\end{aligned}
$$

where we have used $\lim_{k\to\infty}[k\sum_{j=1}^{k} \tilde{W}_j^2] = \lim_{k\to\infty}[\frac{4}{k}\sum_{j=1}^{k} W_j^2] = 4\int_0^1 (1-u)^2 du = 4/3$ in the last step. Hence, the AMSE of the weighted least squares estimator is given by

$$
0 < MSE(\hat{\gamma}^{+}_{wls}) = \gamma^2 \left(\frac{4}{3k} + \frac{4S_1 \dot{S}}{S_2} + \frac{S_1^2 \ddot{S}}{S_2^2} \right) + o(1/k). \tag{18}
$$

Note that, the following basic properties are required to study the asymptotic behaviour of the reduced-bias weighted least squares estimator.

Lemma 1: *Assume that* $\rho < 0$, *then, as* $k \to \infty$;

i. $S_1 = \sum_{j=1}^{k} \tilde{W}_j C_j \to 2/(1-\rho)(2-\rho)$.

ii. $S_2 = \sum_{j=1}^{k} \tilde{W}_j C_j^2 - (\sum_{j=1}^{k} \tilde{W}_j C_j)^2 \to \rho^2(5-\rho)/(1-2\rho)(1-\rho)^2(2-\rho)^2$.

iii. $\dot{S} = \sum_{j=1}^{k} \tilde{W}_j^2 (S_1 - C_j) \to 0$.

iv. $\ddot{S} = \sum_{j=1}^{k} \tilde{W}_j^2 (S_1 - C_j)^2 \to 0$.

Proof of Lemma 1. The following approximations are useful in the proof of Lemma 1: We recall the weight function and the normalised weight function as follows:

$$W_j = 1 - \tfrac{j}{k+1}, 1 \le j \le k$$

and

$$\tilde{W}_j = \frac{W_j}{\sum_{j=1}^{k} W_j}$$

Now the proof of Lemma 1 is as follows:

i. $S_1 = \sum_{j=1}^{k} \tilde{W}_j C_j$ can be rewritten as

$$S_1 = \frac{2}{k}\sum_{j=1}^{k} W_j C_j = \frac{2}{k}\sum_{j=1}^{k}\left(1-\frac{j}{k+1}\right)\left(\frac{j}{k+1}\right)^{-\rho}.$$

As $k \to \infty$, we have

$$\begin{aligned}\frac{1}{k}\sum_{j=1}^{k} W_j C_j &= \frac{1}{k}\sum_{j=1}^{k}\left(1-\frac{j}{k+1}\right)\left(\frac{j}{k+1}\right)^{-\rho}\\ &\approx \int_0^1 (u^{-\rho} - u^{1-\rho})du + o(1)\\ &= \frac{1}{(1-\rho)(2-\rho)} + o(1).\end{aligned}$$

Hence $S_1 \to 2/(1-\rho)(2-\rho)$ as $k \to \infty$.

ii. $S_2 = \sum_{j=1}^{k} \tilde{W}_j C_j^2 - (\sum_{j=1}^{k} \tilde{W}_j C_j)^2$ can also be expressed as

$$S_2 = \frac{2}{k}\sum_{j=1}^{k} W_j C_j^2 - S_1^2 \approx \frac{2}{k}\sum_{j=1}^{k}\left(1-\frac{j}{k+1}\right)\left(\frac{j}{k+1}\right)^{-2\rho} - \left(\frac{2}{(1-\rho)(2-\rho)}\right)^2.$$

As $k \to \infty$,

$$\frac{1}{k}\sum_{j=1}^{k} W_j C_j^2 = \frac{1}{k}\sum_{j=1}^{k}\left(1-\frac{j}{k+1}\right)\left(\frac{j}{k}\right)^{-\rho}$$
$$\approx \int_0^1 (u^{-2\rho}-u^{1-2\rho})du + o(1)$$
$$= \frac{1}{2(1-\rho)(1-2\rho)} + o(1).$$

This implies that, as $k \to \infty$;

$$S_2 = \frac{2}{2(1-\rho)(1-2\rho)} - \frac{4}{(1-\rho)^2(2-\rho)^2} + o(1)$$
$$= \frac{\rho^2(5-\rho)}{(1-2\rho)(1-\rho)^2(2-\rho)^2} + o(1).$$

That is, as $k \to \infty$, $S_2 \to \rho^2(5-\rho)/(1-2\rho)(1-\rho)^2(2-\rho)^2$.

iii. The expression $\dot{S} = \sum_{j=1}^{k} \tilde{W}_j^2 (S_1 - C_j)$ can be rewritten as

$$\dot{S} = \frac{4}{k^2}\sum_{j=1}^{k} W_j^2 (S_1 - C_j).$$

Note from $0 \le W_j \le 1$ and $0 \le C_j \le 1$ that we have

$$-\sum_{j=1}^{k} W_j^2 C_j \le \sum_{j=1}^{k} W_j^2 (S_1 - C_j) \le 2\sum_{j=1}^{k} W_j^2$$

$$-\frac{4}{k^2}\sum_{j=1}^{k} W_j^2 C_j \le \frac{4}{k^2}\sum_{j=1}^{k} W_j^2 (S_1 - C_j) \le \frac{8}{3k}.$$

Therefore, we have,

$$-\frac{4}{k} \le \frac{4}{k^2}\sum_{j=1}^{k} W_j^2 (S_1 - C_j) \le \frac{8}{3k},$$

which gives

$$0 \le \lim_{k\to\infty} \frac{4}{k^2}\sum_{j=1}^{k} W_j^2 (S_1 - C_j) \le 0.$$

Hence $\dot{S} \to 0$ as $k \to \infty$.

- We can also express $\ddot{S} = \sum_{j=1}^{k} \tilde{W}_j^2 (S_1 - C_j)^2$ as

$$\ddot{S} = \frac{4}{k^2}\sum_{j=1}^{k} W_j^2 (S_1 - C_j)^2.$$

We observe that $C_j^2 \le (S_1 - C_j)^2 \le 4$ and hence, we have

$$\frac{4}{k^2}\sum_{j=1}^{k} W_j^2 C_j^2 \le \frac{4}{k^2}\sum_{j=1}^{k} W_j^2 (S_1 - C_j)^2 \le \frac{16}{k^2}\sum_{j=1}^{k} W_j^2 \le \frac{16}{3k}.$$

Now, using

$$0 \le \frac{4}{k^2}\sum_{j=1}^{k} W_j^2 C_j^2 \le \frac{4}{k}, \tag{19}$$

we obtain $0 \le \frac{4}{k^2}\sum_{j=1}^{k} W_j^2 (S_1 - C_j)^2 \le \frac{16}{3k}$ and $0 \le \lim_{k\to\infty} \frac{4}{k^2}\sum_{j=1}^{k} W_j^2 (S_1 - C_j)^2 \le 0.$ Thus we have, $\ddot{S} \to 0$ as $k \to \infty$ which completes the proof of Lemma 1.

A desirable property of an estimator is consistency. If more observations are included, we hope to obtain a lot of information about the unknown parameter. From Lemma 1, the AMSE of $\hat{\gamma}_{wls}^{+}$ approaches 0, as $k \to \infty$. This implies that the estimator $\hat{\gamma}_{wls}^{+}$ is "MSE-Consistent".

We conclude this section by stating and proving the sampling distribution of the weighted least squares estimator, $\hat{\gamma}_{wls}^{+}$.

Theorem 2: *Suppose (3) and (9) holds. Assume also that ρ is estimated by a consistent estimator. Let $k \to \infty, k/n \to 0$, as $n \to \infty$, then for $\sqrt{k} b_{n,k} \to_p 0$*

$$\sqrt{3k}(\hat{\gamma}_{wls}^{+} - \gamma)/2\gamma \to_d N(0,1). \tag{20}$$

The following five Lemmas will be required to prove Theorem (2). We write

$$\mathcal{M}_k(C_j) := \frac{\tilde{W}_j (C_j - \sum_{j=1}^{k} \tilde{W}_j C_j)}{\sum_{j=1}^{k} \tilde{W}_j C_j^2 - (\sum_{j=1}^{k} \tilde{W}_j C_j)^2}. \tag{21}$$

Lemma 3: *Let $C_j = \left(\frac{j}{k+1}\right)^{-\rho}, \tilde{W}_j = \frac{W_j}{\sum_{j=1}^{k} W_j}, j \in \{1,2,\ldots,k\}$ and $\rho < 0$. Then as $k \to \infty$,*

$$\mathcal{M}(C_j) \to o(1/k^{1+\omega}),$$

where $0 < \omega < 0.1$.

Proof of Lemma 3. We observe that

$$-1/|k^{1+\omega} + 5\rho| \le \mathcal{M}_k(C_j) \le C_j^2 / S_2 |k^{1+\omega} + 5\rho|.$$

Therefore, we have $\mathcal{M}_k(C_j) \to o(1/k^{1+\omega}), 0 < \omega < 0.1$ as $k \to \infty$, which completes the proof of Lemma 3.

Lemma 4: *Let* $Z_1, Z_2, Z_3, \ldots$ *be independent random variables,* $a_i \in \mathbb{R}, i \geq 1$ *and for all* $p > 0$ $\mathbb{E}|Z_i^p| < \infty$. *We write*

$$Y = \sum_{j=1}^{n} a_i Z_i,$$

then the moment generating function (m.g.f) of Y *is given as*

$$M_Y(t) = \prod_{i=1}^{n} M_{Z_i}(a_i t),$$

where M_{Z_i} *is the m.g.f of* Z_i.

Proof of Lemma 4. Note that $Z_1, Z_2, Z_3, \ldots$ are independent random variables and the moment generating function of Y is given as

$$\begin{aligned} M_Y(t) &= \mathbb{E}(e^{Yt}) \\ &= \mathbb{E}(e^{t(a_1 Z_1 + a_2 Z_2 + \ldots + a_n Z_n)}) \\ &= \mathbb{E}(e^{a_1 Z_1 t} e^{a_2 Z_2 t} \ldots e^{a_n Z_n t}) \\ &= \prod_{i=1}^{n} M_{Z_i}(a_i t). \end{aligned}$$

Lemma 5: *Assume* $Z_1, Z_2, Z_3, \ldots$ *are independent random variables on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *and for all* $p > 0$, $\mathbb{E}|Z_i^p| < \infty$, *then for any* $\varepsilon > 0$

$$\lim_{k \to \infty} \mathbb{P}\left(\sqrt{k} b_{n,k} \geq \varepsilon\right) = 0. \tag{22}$$

Proof of Lemma 5. The proof of this lemma would be obtained by large deviations techniques applied to weighted sum of independent random variables $\mathcal{M}_k(C_j)Z_j$. We observe from (21) and (15) that $b_{n,k}$ is expressible as $\hat{b}_{n,k} = \sum_{j=1}^{k} \mathcal{M}_k(C_j)Z_j$. Recall that the Z_j's are independent exponentials with finite mean $\gamma + b_{n,k}/(k+1)^{-\rho} \leq \mu_j \leq \gamma + b_{n,k}$. Therefore, the moment generating function of Z_k is given as

$$M_{Z_k}(t) = \frac{1}{1 - \mu_k t}$$

Using Lemma 4, the moment generating function of $\hat{b}_{n,k}$ is given by

$$M_{\hat{b}_{n,k}}(t) = \prod_{j=1}^{k} M_{Z_j}(\mathcal{M}_k(C_j)t)$$

$$\therefore \qquad M_{\hat{b}_{n,k}}(kt) = \prod_{j=1}^{k} \frac{1}{1 - \mu_j k \mathcal{M}_k(C_j)t}.$$

Note from Lemma 3 that, $\mathcal{M}_k(C_j) \to o(1/k^{1+\omega})$ as $k \to \infty$; hence we have $\mu_j k \mathcal{M}_k(C_j) \to o(1/k^{1+\omega})$ $k \to \infty$, $k/n \to 0$, as $n \to \infty$. Therefore, we can find some $\delta > 0$ such that

$$-\delta \le \mu_j k \mathcal{M}_k(C_j) \le \delta.$$

This implies,

$$(1-\delta+\delta^2-\delta^3+\ldots)^k \le M_{\hat{b}_{n,k}}(kt) \le (1+\delta+\delta^2+\delta^3+\ldots)^k$$

$$\log(1-\delta+\delta^2-\delta^3+\ldots) \le \lim_{k\to\infty} \frac{1}{k} \log M_{\hat{b}_{n,k}}(kt) \le \log(1+\delta+\delta^2+\delta^3+\ldots)$$

Now taking the limit $\delta \downarrow 0$ and applying the Sandwich Theorem, we have

$$\lim_{k\to\infty} \frac{1}{k} \log M_{\hat{b}_{n,k}}(kt) = 0. \tag{23}$$

Equation (23) means that the limiting logarithmic moment generating function converges to 0 with speed k. Hence by the Gärther-Ellis theorem [14, Theorem 2.3.6, p. 44], $\hat{b}_{n,k}$ satisfies a large deviation principle (LDP) in the space of non-negative real numbers with speed k and rate function $I(x)$ given by

$$\begin{aligned} I(x) &= \sup_{\lambda\in\mathbb{R}} \{\lambda x - 0\} \\ &= \sup_{\lambda\in\mathbb{R}} \{\lambda x\} \\ &= \begin{cases} 0 \text{ if } x = 0 \\ \infty \text{ if } x \ne 0 \end{cases} \end{aligned}$$

where $x \in [0,\infty)$ and $\lambda > 0$. If $\hat{b}_{n,k}$ obeys LDP in the space $[0,\infty)$ with speed k and a rate function $I(x)$, then for every $\varepsilon > 0$, we have

$$\limsup_{k\to\infty} \frac{1}{k} \log \mathbb{P}(\sqrt{k}\hat{b}_{n,k} > \varepsilon) = \limsup_{k\to\infty} \frac{1}{k} \log \mathbb{P}(\hat{b}_{n,k} > \tau_k) \le -\inf_{x\in(0,\infty)} I(x), \tag{24}$$

where $\tau_k = \frac{\varepsilon}{\sqrt{k}}$. As the non- typical behaviour of the rate function is when $x \ne 0$, we have

$$\lim_{k\to\infty} \mathbb{P}(\sqrt{k}\,\hat{b}_{n,k} > \varepsilon) \le 0.$$

Thus, $\sqrt{k}\hat{b}_{n,k} \to_p 0$, as $k \to \infty$ which ends the Proof of Lemma 5.

Lemma 6: *Suppose that* $Z_1, Z_2, \ldots$ *are independent random variables such that* $\mathbb{E}(Z_k) = \mu_k$ *and* $Var(Z_k) = \sigma_k^2 < \infty$, *then there exists* $\delta > 0$ *such that*

$$\lim_{k\to\infty}\frac{1}{S_n^{2+\delta}}\sum_{j=1}^{k}\mathbb{E}(|Z_j-\mu_j|^{2+\delta})=0. \tag{25}$$

Proof of Lemma 6. The $Z_1, Z_2, Z_3, \ldots$ are independent but not identical.

$$\begin{aligned}
\mathbb{E}(|Z_k-\mu_k|^{2+\delta}) &= \int_0^\infty |Z_k-\mu_k|^{2+\delta} f(z_k)dz_k \\
&= -\int_0^{\mu_k}(Z_k-\mu_k)^{2+\delta} f(z_k)dz_k + \int_{\mu_k}^{\infty}(Z_k-\mu_k)^{2+\delta} f(z_k)dz_k \\
&= \frac{e^{-1}}{\mu_k}\left\{\frac{(-\mu_k)^{3+\delta}(7+2\delta)}{(3+\delta)(4+\delta)}+\mu_k^{3+\delta}(2+\delta)!\right\}\{1+o(1)\} \\
&= \mu_k^{2+\delta}\eta(\delta)\{1+o(1)\},
\end{aligned}$$

where $\eta(\delta)=\frac{1}{e}\left\{\frac{(-1)^{3+\delta}(7+2\delta)}{(3+\delta)(4+\delta)}+(2+\delta)!\right\}<\infty$ for $0<\delta<\infty$.

Now, the next part of the proof requires, we sum the above expectation over *k*;

$$\begin{aligned}
&\sum_{j=1}^{k}\mathbb{E}(|Z_j-\mu_j|^{2+\delta}) \\
&= \eta(\delta)\sum_{j=1}^{k}\mu_j^{2+\delta}\{1+o(1)\} \\
&= \eta(\delta)\sum_{j=1}^{k}(\gamma+b_{n,k}C_j)^{2+\delta}\{1+o(1)\} \\
&\le \eta(\delta)\sum_{j=1}^{k}(\gamma+b_{n,k})^{2+\delta}\{1+o(1)\}, \text{since } \gamma+\frac{b_{n,k}}{(k+1)^{-\rho}}<\gamma+b_{n,k}C_j<\gamma+b_{n,k} \\
&= k\eta(\delta)(\gamma+b_{n,k})^{2+\delta}\{1+o(1)\}
\end{aligned}$$

Now define

$$T_j = Z_j-\mu_j, 1\le j\le k,$$

$$S_n^2 = Var\left(\sum_{j=1}^{k}T_j\right)=\sum_{j=1}^{k}Var(Z_j-\mu_j)=\sum_{j=1}^{k}Var(Z_j)=\sum_{j=1}^{k}Var(\varepsilon_j)=k\gamma^2.$$

Hence,

$$\begin{aligned}
\lim_{k\to\infty}\frac{1}{S_n^{2+\delta}}\sum_{j=1}^{k}\mathbb{E}(|Z_j-\mu_j|^{2+\delta}) &= \lim_{k\to\infty}\frac{k\eta(\delta)(\gamma+b_{n,k})^{2+\delta}\{1+o(1)\}}{(k\gamma^2)^{1+\delta/2}} \\
&= 0, \text{ since } b_{n,k}\to 0, \text{ as } k\to\infty.
\end{aligned}$$

We write $\lim_{k\to\infty}\mathbb{E}(Y_k):=\mu$, $\lim_{k\to\infty}Var(Y_k):=\sigma^2$ and observe from Lemma (5) and (6) the Proof of Theorem 2 as follows: Lemma (5) establishes that

$\sqrt{k}\hat{b}_{n,k}$ converges in probability to 0 as k becomes large. Lemma (6) shows that Lyapunov's condition for the central limit theorem holds [7, p. 359-362], that is, $Y_k := \sqrt{3k}(\hat{\gamma}^{+}_{wls} - \gamma) / 2\gamma \to_d N(\mu, \sigma^2)$, as $k \to \infty$. Therefore, all we need to complete the proof of Theorem 2 is to estimate the parameters μ and σ^2 under the condition $\sqrt{k}\hat{b}_{n,k} \to_p 0$.

Recall that Y_k converges in distribution to normally distributed random variable with mean μ and variance σ^2. We compute the parameters μ and σ^2 as follows:

$$\begin{aligned}
\mu &= \lim_{k\to\infty} \mathbb{E}(Y_k) = \lim_{k\to\infty}[\frac{\sqrt{3k}}{2\gamma}(\mathbb{E}(\hat{\gamma}^{+}_{wls}) - \gamma)] \\
&= \lim_{k\to\infty}[\frac{\sqrt{3k}}{2\gamma}(\sum_{j=1}^{k}\tilde{W}_j\mathbb{E}(Z_j) - \gamma)] \text{ (since } \sqrt{k}\hat{b}_{n,k} \to 0 \text{ as } k \to \infty) \\
&= \lim_{k\to\infty}[\frac{\sqrt{3k}}{2\gamma}\sum_{j=1}^{k}\gamma\tilde{W}_j + \frac{\sqrt{3k}}{2\gamma}\hat{b}_{n,k}\sum_{j=1}^{k}\tilde{W}_jC_j - \frac{\sqrt{3k}}{2\gamma}\gamma] = 0
\end{aligned}$$

Further, we have

$$\begin{aligned}
\sigma^2 &= \lim_{k\to\infty} Var(Y_k) = \lim_{k\to\infty}[\frac{3k}{4\gamma^2}\sum_{j=1}^{k}\tilde{W}_j^2 Var(Z_j)] \\
&= \lim_{k\to\infty}[\frac{3k}{4}\sum_{j=1}^{k}\tilde{W}_j^2] = \lim_{k\to\infty}[1 + o(1/k)] = 1.
\end{aligned}$$

Hence $\sqrt{k}\hat{b}_{n,k} \to_p 0$ and $Y_k \to_d N(0,1)$, as $k \to \infty$, as required.

3. Results and Discussion

We present a simulation study that compares the performance of the proposed extreme value index estimator with some existing estimators in the literature.

Table 1

Heavy-tailed distributions from the Pareto-type distribution

Distribution	$1 - F(x)$	$\ell_F(x)$	γ
Burr	$(1 + x^{\tau})^{-\lambda}$	$(1 + x^{-\tau})^{-\lambda}$	$1/\lambda\tau$
Fréchet	$1 - exp(-x^{-\alpha})$	$1 - \frac{x^{-\alpha}}{2} + O(x^{-\alpha})$	$1/\alpha$

3.1 *Simulation study*

In this simulation study, the reduced-bias weighted least square estimator is compared to the Hill [23], the least square [3], the Bias-corrected Hill [11] and the ridge regression [9] estimators. The estimators are illustrated for two distributions from the Pareto-type distribution as shown in Table 1, taking 1000 repetitions of samples of sizes 50 and 200. The bias and MSE are plotted as a function of k to investigate the behaviour of the estimators' sample paths. For each distribution, we consider three different values of $\gamma \in (0,1]$ to investigate the performance of the proposed estimator for varying values of γ. In particular, we consider $\gamma = 0.10, 0.50$ and 1.00, since these values are usually used in simulation studies (see for example [5, 10, 32, 33]). The following distributions are used:

- Burr (η, τ, λ)
  - $\eta = 1, \tau = \sqrt{10}$ and $\lambda = \sqrt{10}$, so that $\gamma = 0.10$.
  - $\eta = 1, \tau = \sqrt{2}$ and $\lambda = \sqrt{2}$, so that $\gamma = 0.50$.
  - $\eta = 1, \tau = 2$ and $\lambda = 1/2$, so that $\gamma = 1.00$.

- Fréchet (α)
  - $\alpha = 10, 2$ and 1, so that $\gamma = 0.10, 0.50$ and 1.00, respectively.

Table 2 presents the notations of the extreme value index estimators used in the simulation study.

Table 2

Notations of the Estimators

Estimators	Notation
Hill	HILL
Bias-corrected Hill	BCHILL
Least square	LS
Ridge regression	RR
reduced-bias Weighted least square	WLS

The simulation results for the Burr distribution are presented in Figures 1 - 3, and that of the Fréchet distribution are shown in Figures 4 - 6. The sample paths of the proposed estimator, WLS, are close to that of LS in most cases, and this means that the two estimators are competitively close to each other.

In the case of the Burr distribution, WLS generally outperforms the other estimators in terms of MSE and bias for small EVI $(\gamma = 0.1)$, and the

WLS estimator is mostly the second-best estimator in terms of bias and MSE for $\gamma = 0.50$. However, in the case of $\gamma = 1.00$, the sample paths of WLS and BCHILL are generally close for small to medium values of k. The WLS estimator mostly has the lowest bias and MSE among the three reduced-bias estimators (i.e., WLS, LS and RR) for the Burr distribution.

Furthermore, in the case of the Fréchet distribution, the WLS is mostly the best performing estimator in terms of bias. The sample paths of WLS and LS are mainly close to each other. However, WLS generally appears to outperform LS in terms of bias. For a large sample, i.e., $n = 200$, the WLS estimator outperforms the other estimators in terms of MSE for large values of k.

The MSE plots of the WLS estimator are stable over the middle region of k and thus makes the determination of the optimal k, defined by $k_0 = \arg\min_k [MSE\{\hat{\gamma}^+_{wls}(k)\}]$, easier.

Given the role of ρ, the WLS was also compared to the other estimators, using the [19] estimator for ρ, and although it was not universally the best in terms of bias and MSE, it performed well compared to the other estimator. However, for brevity we omit these results.

In summary, the reduced-bias weighted least squares estimator generally yields lower bias and MSE, which are stable over a long range of k values. Thus, it can be considered an appropriate estimator of the extreme value index for samples generated from the Pareto-type distribution.

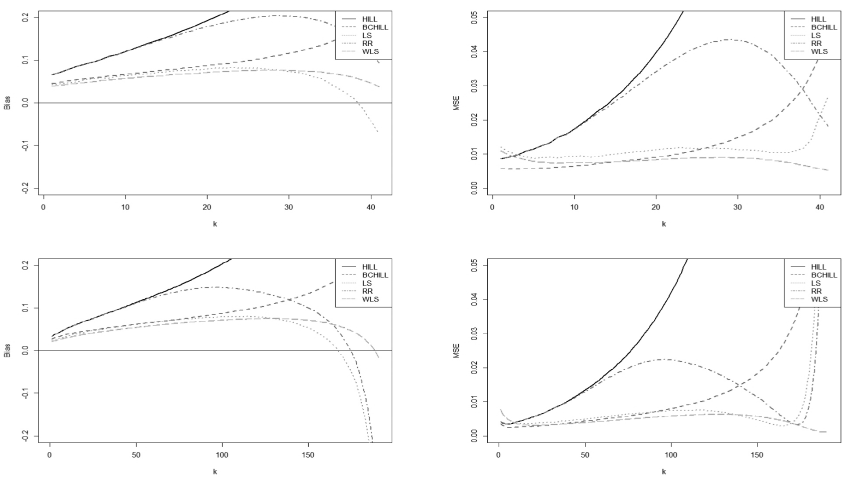

Figure 1

Results for Burr distribution with γ = 0.1: Bias(left panel); MSE(right panel). First row: n = 50; second row: n = 200

3.2 *Practical Illustration*

To illustrate the practical application of the proposed estimator, we consider the estimation of the extreme value index of the underlying distribution of the Secura Belgian reinsurance claim size data and the calcium content (measured in mg/100g of dry soil) of soil from a particular city (NIS code 61072) in the Condroz region of Belgium. These data sets have been studied extensively in the extreme value theory literature such as [4]. In this application, we lookout for estimators whose graph is near horizontal and smooth.

The Secura Belgian Reinsurance data contains 371 automobile claims in euros which occurred from 1989 to 2001. [4] have demonstrated that the Secura Belgian reinsurance data has a heavy tail. The plots of the extreme value index estimates as a function k is shown in Figure 7.

Also, the Condroz data contains 1505 observations, and seven of these values deviate from the rest of the observations. [4] removed the top seven Calcium measurements before modelling the data. However, we fitted the biased-reduced weighted least square to the complete data to study the actual tail behaviour of the data. [4] have shown that the Ca measurements in the Condroz data belongs to the Pareto domain of attraction. The extreme value index estimates as a function of k of the Ca measurements of the Condroz data is illustrated in Figure 8.

In the case of the Secura Belgian data, the HILL estimator is very sensitive to the choice of k, and this makes its interpretation challenging. The BCHILL and LS estimators exhibit some moderate stability in k. The ridge regression estimates are stable for k values between 110 and 220, while the reduced-bias weighted least square yields stable estimates for $k \geq 120$. The WLS shows stability across a larger region of k, making it easier to specify the value of the estimated extreme value index.

Regarding the Condroz data, the plots of HILL and BCHILL estimators are difficult to interpret for large values of k, and this is because the estimators are very sensitive to the choice of k. However, they are less sensitive for k values between 270 and 500. The extreme value index estimates for the RR, LS and WLS estimators are stable and approximately equal to 0.26 for k values in the interval [283, 1097], [505, 1107] and [710, 1230] respectively.

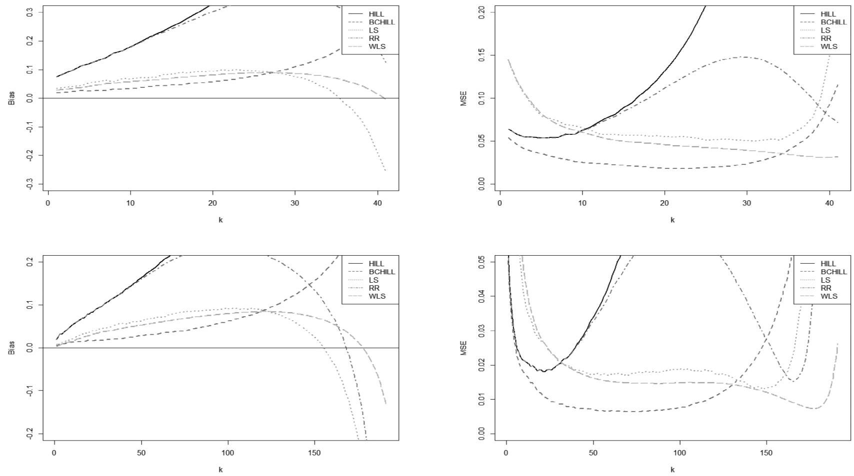

Figure 2

Results for Burr distribution with $\gamma = 0.5$: Bias (left panel); MSE (right panel). First row: $n = 50$; second row: $n = 200$.

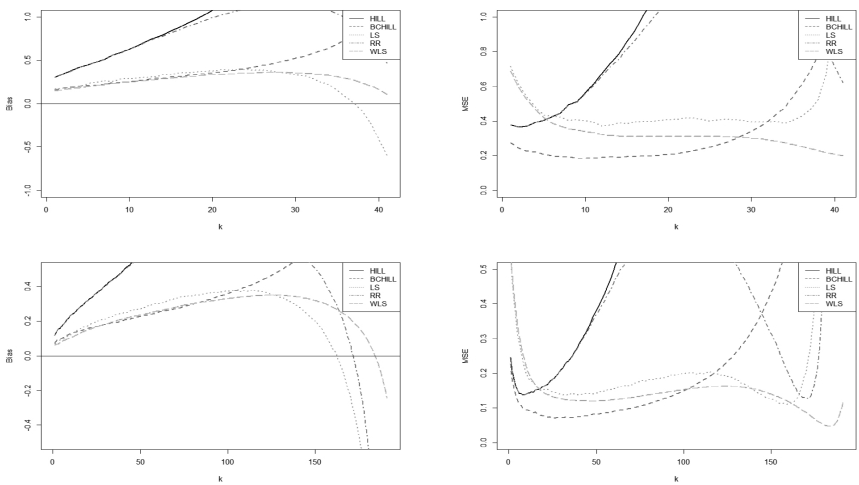

Figure 3

Results for Burr distribution with $\gamma = 1.0$: Bias (left panel); MSE (right panel). First row: $n = 50$; second row: $n = 200$.

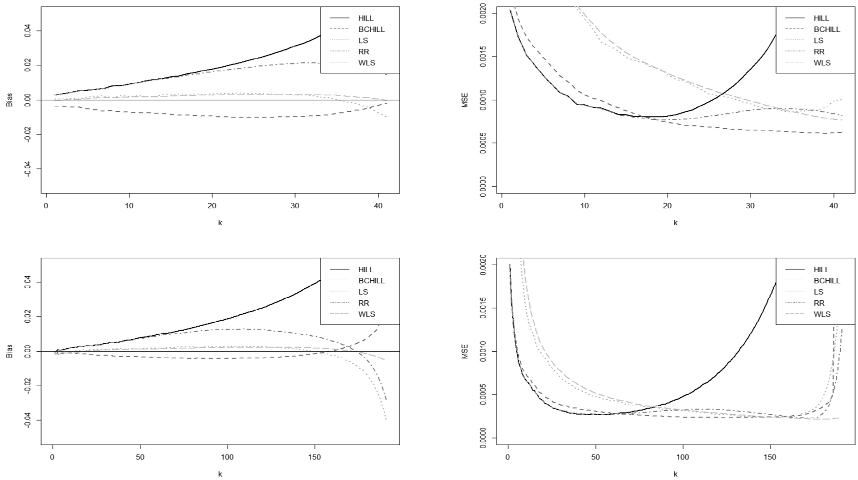

Figure 4

Results for Fréchet distribution with $\gamma = 0.1$: Bias (left panel); MSE(rightmost panel). First row: $n = 50$; second row: $n = 200$

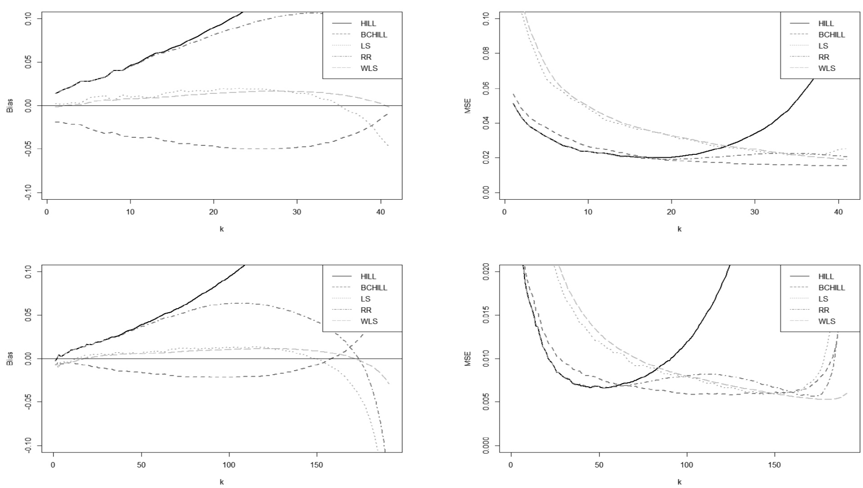

Figure 5

Results for Fréchet distribution with $\gamma = 0.5$: Bias (left panel); MSE (right panel). First row: $n = 50$; second row: $n = 200$.

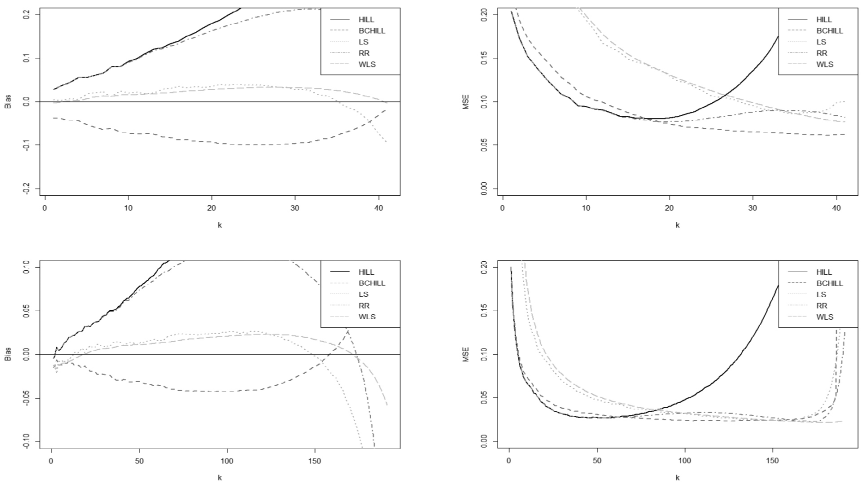

Figure 6

Results for Fréchet distribution with γ = 1.0 :: Bias (left panel); MSE (right panel). First row: n = 50; second row: n = 200.

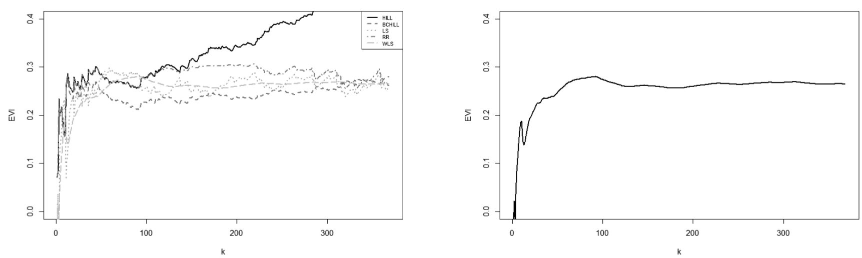

Figure 7

Claim sizes of Secura Belgian Reinsurance Data: $\hat{\gamma}^{+}_{wls}$ on the right

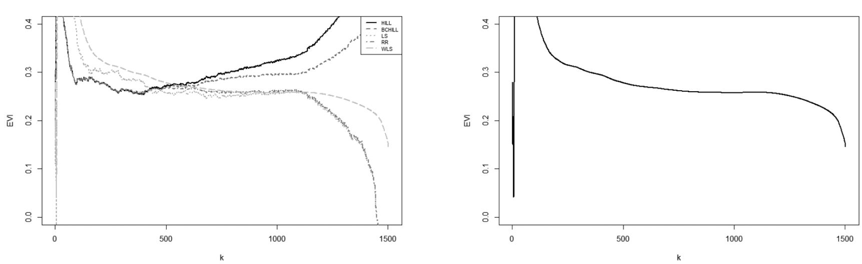

Figure 8

Calcium content of soil. Condroz Data: $\hat{\gamma}^{+}_{wls}$ on the right

The ρ estimator by [19] was also considered for the application but the result is omitted. The reduced-bias estimators produced stable estimates over some k values but not as horizontal as when the minimum variance [9] approach is used, especially in the case of the Secura Belgian claim data whose sample size is relatively small.

4. Conclusion

This study proposed an alternative reduced-bias estimator for the extreme value index, $\gamma(\gamma > 0)$, using the exponential regression model introduced by [3]. Specifically, we proposed a reduced-bias weighted least squares estimator of the extreme value index using the exponential regression model for the weighted log-spacings. We showed that the proposed estimator is asymptotically unbiased, asymptotically consistent and asymptotically normal. In addition, it performed favourably well in terms of bias and MSE relative to the other EVI estimators considered in the study. Furthermore, it produces stable estimates that are less sensitive to the tail sample fraction, k. Hence it can easily be used in the selection of the appropriate value of k which is an ongoing research area in extreme value analysis. In application, we suggest practitioners plots $\hat{\gamma}_{wls}^{+}$ as a function of k to aid in the selection of an optimal k for the proposed estimator.

Unlike the least squares and the ridge regression estimators in the literature, the asymptotic variance of the proposed estimator is $\sigma^2 = 4/3\gamma^2$ and does not depend on ρ. Also, the asymptotic variance of the proposed estimator is less than that of the least squares estimator for $-6 \leq \rho < 0$. Moreover, we envisage that if we choose a random weight, we can attain an asymptotic variance that is as smaller as that of the Hill or the bias-corrected Hill estimators. This is a topic that deserves further studies and will be considered in a subsequent work.

Data Availability

The data sets used in this study to support the findings are from [4] and are available at https://lstat.kuleuven.be/Wiley/.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgement

Note, a draft of the article has previously appeared online via arxiv.org as a preprint. See, [35]. Also, Ocran, E. would like to thank the University of Ghana Building a New Generation of Academics in Africa (BANGA-Africa) Project (funded by Carnegie Corporation of New York) for providing financial support for this Ph.D research work.

References

[1] J. Beirlant, G. Dierckx, Y. Goegebeur, and G. Matthys, ”Extreme value index estimation and an exponential regression model,” *Extremes,* vol. 2, no. 2, pp. 177–200 (1999).

[2] J. Beirlant, G. Dierckx, and A. Guillou, ”Estimation of the extreme-value index and generalized quantile plots,” *Bernoulli,* vol. 11, no. 6, pp. 949–970 (2005).

[3] J. Beirlant, G. Dierckx, A. Guillou, and C. Starica, ”On exponential representations of log-spacings of extreme order statistics,” *Extremes,* vol. 5, pp. 157–180 (2002).

[4] J. Beirlant, Y. Goegebeur, J. Teugels, and J. Segers, Statistics of Extremes: Theory and Applications, vol. 558, John Wiley & Sons (2004).

[5] J. Beirlant, G. Maribe, and A. Verster, ”Using shrinkage estimators to reduce bias and MSE estimation of heavy tails,” *Statistical Journal,* vol. 17, no. 1, pp. 91–108 (2019).

[6] J. Beirlant, P. Vynckier, and J. L. Teugels, ”Extreme value index estimation, Pareto quantile plots regression diagnostics,” *Journal of the American Statistical Association*, vol. 91, no. 436, pp. 1659–1667 (1996).

[7] P. Billingsley, Probability and Measure, 3rd ed., vol. 26, Wiley Series in Probability and Mathematical Statistics, New York (1983).

[8] M. Bladt, H. Albrecher, and J. Beirlant, ”Trimmed extreme value estimators for censored heavy-tailed data,” *Electronic Journal of Statistics*, vol. 15, no. 1, pp. 3112–3136 (2021).

[9] S. Buitendag, J. Beirlant, and T. de Wet, ”Ridge regression estimators for the extreme value index,” *Extremes*, vol. 22, no. 2, pp. 271–292 (2018).

[10] I. Cabral, F. Caeiro, and M. I. Gomes, ”On the comparison of several classical estimators of the extreme value index,” *Communications in Statistics - Theory and Methods*, pp. 1–18 (2020).

[11] F. Caeiro, M. I. Gomes, and D. Pestana, "Direct reduction of bias of the classical Hill estimator," *Statistical*, vol. 3, no. 2, pp. 113–136 (2005).

[12] L. De Haan and S. Resnick, "On asymptotic normality of the Hill estimator," *Stochastic Models,* vol. 14, no. 4, pp. 849–866 (1998).

[13] A. L. M. Dekkers, J. H. J. Einmahl, and L. de Haan, "A moment estimator for the index of an extreme-value distribution," *The Annals of Statistics,* vol. 17, no. 4, pp. 1833–1855 (1989).

[14] A. Dembo and O. Zeitouni, Large Deviations Techniques and Applications, vol. 38, Springer Science & Business Media (2009).

[15] S. Elvidge and M. J. Angling, "Using extreme value theory for determining the probability of Carrington-like solar flares," *Space Weather*, vol. 16, no. 4, pp. 417–421 (2018).

[16] P. Embrechts, "Extremes in economics and the economics of extremes," in Extreme Values in Finance, Telecommunications, and the Environment, pp. 188–202, Chapman and Hall/CRC (2003).

[17] A. Feuerverger and P. Hall, "Estimating a tail exponent by modelling departure from a Pareto distribution," *The Annals of Statistics,* vol. 27, no. 2, pp. 760–781 (1999).

[18] B. Finkenstadt and H. Rootz´en, eds., Extreme Values in Finance, Telecommunications, and the Environment, Chapman & Hall/CRC (2004).

[19] M. Fraga Alves, L. de Haan, and T. Lin, "Estimation of the parameter controlling the speed of convergence in extreme value theory," *Mathematical Methods of Statistics*, vol. 12, no. 2, pp. 155–176 (2003).

[20] K. Gkillas and P. Katsiampa, "An application of extreme value theory to cryptocurrencies," *Economics Letters,* vol. 164, pp. 109–111 (2018).

[21] M. I. Gomes, L. De Haan, and L. H. Rodrigues, "Extreme value index estimation for heavy-tailed models: Accommodation of bias in weighted log-excesses," *Journal of the Royal Statistical Society: Series B* (Statistical Methodology), vol. 70, no. 1, pp. 31–52 (2008).

[22] M. I. Gomes, U. D. Lisboa, M. Joa, and U. Te, "Asymptotically unbiased estimators of the extreme value index based on external estimation of the second order parameter," *Extremes*, pp. 5–31 (2002).

[23] M. B. Hill, "A simple general approach to inference about the tail of a distribution," *Annals of Statistics*, vol. 3, no. 5, pp. 1163–1174 (1975).

[24] R. Huisman, K. G. Kool, C. J. M. Koedijk, and F. Palm, "Tail-index estimates in small samples," *Journal of Business & Economic Statistics*, vol. 19, no. 2, pp. 208–216 (2001).

[25] J. H¨usler, D. Li, and S. M¨uller, "Weighted least squares estimation of the extreme value index," *Statistics & Probability Letters*, vol. 76, no. 9, pp. 920–930 (2006).

[26] M. M. Kithinji, P. N. Mwita, and A. O. Kube, "Adjusted extreme conditional quantile autoregression with application to risk measurement," *Journal of Probability and Statistics*, vol. 2021 (2021).

[27] F. Longin, ed., Extreme Events in Finance: A Handbook of Extreme Value Theory and Its Applications, John Wiley & Sons (2016).

[28] V. Marimoutou, B. Raggad, and A. Trabelsi, "Extreme value theory and value at risk: Application to oil market," *Energy Economics,* vol. 31, no. 4, pp. 519–530 (2009).

[29] D. M. Mason, "Laws of large numbers for sums of extreme values," *The Annals of Probability*, pp. 754–764 (1982).

[30] N. Mehrnia and S. Coleri, "Wireless channel modeling based on extreme value theory for ultra-reliable communications," *IEEE Transactions on Wireless Communications,* vol. 21, no. 2, pp. 1064–1076 (2021).

[31] R. Minkah, "Extreme value index estimation of the generalized Pareto distribution using a pivot from a transformed Pareto distribution," *Science and Development Journal*, vol. 4, no. 1, pp. 1–19 (2020).

[32] R. Minkah, T. de Wet, and A. Ghosh, "Robust estimation of Pareto-type extreme value index through an exponential regression model," *Communications in Statistics - Theory and Methods*, pp. 1–19 (2021).

[33] R. Minkah, T. Doku-Amponsah, and D. Tertius, "On extreme value index estimation under random censoring," *African Journal of Applied Statistics*, vol. 5, no. 2, pp. 419–445 (2018).

[34] S. B. Muela, C. L. Mart´ın, and R. A. Sanz, "An application of extreme value theory in estimating liquidity risk," *European Research on Management and Business Economics*, vol. 23, no. 3, pp. 157–164 (2017).

[35] F. C. Onwuegbuche, S. B. Affognon, E. P. Enoc, M. O. Akinade, et al., "Application of extreme value theory in predicting climate change induced extreme rainfall in Kenya," I*nternational Journal of Statistics and Probability*, vol. 8, no. 4, pp. 85–94 (2019).

[36] J. Pickands, "Statistical inference using extreme order statistics," *The Annals of Statistics*, vol. 3, no. 2, pp. 119–131 (1975).

[37] C. Rohrbeck, E. F. Eastoe, A. Frigessi, and J. A. Tawn, ”Extreme value modelling of water-related insurance claims,” *The Annals of Applied Statistics*, vol. 12, no. 1, pp. 246–282 (2018).

[38] L. Viharos, ”Weighted least-squares estimators of tail indices,” Probability and Mathematical Statistics-Wroclaw University, vol. 19, no. 2, pp. 249–265 (1999).

[39] C. Yozgatlgil and M. T¨urke¸s, ”Extreme value analysis and forecasting of maximum precipitation amounts in the western Black Sea subregion of Turkey,” *International Journal of Climatology*, vol. 38, no. 15, pp. 5447–5458 (2018).

[40] P. Hall, ”On some simple estimates of an exponent of regular variation,” *Journal of the Royal Statistical Society: Series B*, vol. 44, no. 1, pp. 37–42 (1982).

Received December, 2021

Revised July, 2022