A regression method estimation of the extreme value index using decreasing dependent random weights

By

Emmanuel Adjei (edjei@aims.edu.gh)

May 2022

AN ESSAY PRESENTED TO AIMS-GHANA IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF MASTER OF SCIENCE IN MATHEMATICS

DECLARATION

This work was carried out at AIMS-Ghana in partial fulfilment of the requirements for a Master of Science Degree.

I hereby declare that except where due acknowledgement is made, this work has never been presented wholly or in part for the award of a degree at AIMS-Ghana or any other University.

Student: Emmanuel Adjei

Supervisor: Prof. Kwabena Doku-Amponsah

ACKNOWLEDGEMENTS

My supervisor, Prof. Kwabena Doku-Amponsah, generously provided knowledge and expertise. He deserves special thanks for his insightful and helpful feedback during the conception and production of this work. His generosity in giving his time was greatly appreciated. I would want to express my gratitude to Dr. Eric Ocran for his assistance throughout the simulation. He has demonstrated what a good scientist (and person) should be by his actions. Additionally, this endeavour would not have been possible without the support from the African Institute for Mathematical Sciences, Ghana for their financial support throughout my studies. I would also like to thank all the tutors for their support throughout my studies.

DEDICATION

The essay is dedicated to my family and numerous friends. I owe a special debt of gratitude to my church family for their unwavering support throughout the process. I will always appreciate all they have done, especially Charlotte Anogye for her support throughout high school and Harrison Obeng Debrah for his continuous guidance and support throughout high school and tertiary education.

Bernice Brobbey and Eric Donkor, my dearest friend, who has been there for me throughout my master's degree, deserves special recognition for this work.

Abstract

In this essay, a regression method for the extreme value index of a Pareto-type distribution is presented using the weighted least squares approach. The estimator is proven to be consistent, asymptotically normal, and unbiased considering the second-order assumption on the data distribution. Simulations are carried out to analyse the estimators' finite sample behaviour. The outcome of the simulation reveals that the estimator competes very well with available estimators of the Extreme Value Index in literature with regards to bias and Mean Square Error (MSE). The estimate were found not to be so sensitive to the number of top-order statistics and can therefore be employed for choosing the best tail fraction. A practical illustration was carried out with the proposed estimator using insurance and pedochemical data sets.

Contents

1. Introduction

The crucial conclusion reached by Fisher and Tippett in 1928 on the various limit laws of the sample maximum appears to have given rise to the notion that extreme value theory (EVT) is something unique and distinct from classical central limit theory [5]. There are numerous statistical tools available for deriving information about certain measures in a statistical distribution. Extreme statistics is concerned with estimating the occurrence of uncommon events such as high quantile, exceedance likelihood, and return periods by considering a limit as a failure threshold beyond which some form of failure or end-of-life event will occur. Knowing the frequency and size of these events aids in the formulation of strategies to reduce their consequences [2, 28]. The EVT addresses these extreme events by categorizing continuous distributions based on the behaviour of the tail region or their extreme realizations. This extreme value index (EVI), γ, is the fundamental parameter in estimating these events because the EVI must be known for all inference because the EVI is crucial in extreme value analysis. The extreme value index controls the tail heaviness of the extreme value distribution, assuming that the underlying distribution's max-domain of attraction criterion holds. The heaviness of the tail function $\bar{P} = 1 - P$ gets larger with respect to γ, where $\gamma \in (-\infty, \infty)$. Our concentration will be for a specific scenario where $\gamma > 0$ in this essay. This means that we will operate in the domain of attraction for a peak of an extreme value distribution function,

$$G_\gamma(y) = e^{-(1+\gamma y)^{-1/\gamma}}, \ y \geq -1/\gamma.$$

This type of distribution has been proven to be effective in a variety of sectors, including re-insurance [27, 25, 34], finance [16, 20, 21], risk management [23, 22, 1], telecommunication [24, 15] and climatology [30]. For $\gamma > 0$, the fundamental distribution, the upper tail function of P is a distribution of the Pareto type expressed as

$$1 - P(y) = y^{-1/\gamma}\varphi_F(y) \text{ as } y \to \infty, \tag{1.0.1}$$

where φ on $(0, \infty)$ is a positive, Lebesgue-measurable function slowly varying at ∞ and it is expressed as

$$\lim_{t\to\infty} \frac{\varphi(yt)}{\varphi(t)} = 1, \ \ t > 1. \tag{1.0.2}$$

For Pareto-type distributions, γ is strictly positive [28].

Suppose $Y_1, Y_2, ..., Y_m$ denotes random variables selected from a distribution that are independent and identically distributed (i.i.d), and a member of the maximum domain of attraction of the Pareto family of distributions, then for some auxiliary sequences of constants $\{d_m > 0; m \geq 1\}$ and $\{c_m; m \geq 1\}$ [19]

$$\lim_{m\to\infty} P\left(\frac{\max\{Y_1, Y_2, ..., Y_m\} - c_m}{d_m} \le y\right) = e^{-(1+\gamma y)^{-1/\gamma}},\ y \ge -1/\gamma, \tag{1.0.3}$$

where $c_m = n\mu_m$, $d_m = \sqrt{n\sigma_m}$ and $\gamma \ge 0$. Because every conclusion in extreme value analysis rely on the tail index, the estimation of γ continues to attract a lot of attention in statistics of extremes. We aim for estimators with the least variance and bias in practice. Estimating the tail index can be done using a parametric or semi-parametric method in order to reduce bias [10, 8, 37, 35]. In this study, however, we use a semi-parametric method to generate reduced-bias estimators.

Tail index estimators in the semi-parametric framework are reliant on the k largest observations, based on these conditions of k:

(i) **Condition 1:** $k(m) \to \infty$ as $m \to \infty$.

(ii) **Condition 2:** $k(m) \to O(m)$ as $m \to \infty$.

The HILL estimator is the universally employed semi-parametric tail index estimator [18]. Using the maximum likelihood estimator, Hill [18] estimates $\gamma > 0$ and approximates the top-order statistics with a Pareto-type distribution. Among the semi-parametric estimators, the Hill estimator has the lowest asymptotic variance, and it is particularly sensitive to the choice of k [5]. Because of this flaw, the estimator is difficult to use in reality, particularly when determining the tail fraction. Hill [18] established the tail estimator as

$$\hat{\gamma}_k^H = \frac{1}{k}\sum_{j=1}^{k}\left(\ln\left(Y_{m-j+1,m}\right) - \ln\left(Y_{m-k,m}\right)\right). \tag{1.0.4}$$

By Renyi's representation, equation (1.0.4) can be expressed as

$$\hat{\gamma}_k^H = \frac{1}{k}\sum_{j=1}^{k} Z_j,$$

where

$$Z_j = j\ln\left(\frac{Y_{m-j+1,m}}{Y_{m-j,m}}\right) \tag{1.0.5}$$

and $Y_{j,m}$ represent the j^{th} order-statistics in relation to $Y_{j,1} \le Y_{j,2} \le Y_{j,3} \le \cdots \le Y_{j,m}$.

Because of its widespread use, the Hill estimator has been subjected to a large number of generalizations: Consider the works for example, [12, 13, 3, 6, 9, 31, 32, 33]. The bias-corrected Hill estimator is another estimate that is likewise a modification of the Hill estimator [11].The bias-corrected Hill estimator yields stable tail index estimates that are slightly sensitive in choosing k compared to the other estimators in literature.

We present alternative tail index estimators that, under certain conditions, empirically give substantially more stable tail index estimate and achieve the Hill estimator's least asymptotic variance. In actuality, reliable tail index estimators are in high demand since they partially solve the problem of selecting the best k.

1.1 Aim of Study

A reduced-bias estimator of the extreme value index for Pareto-type distributions with heavy-tailed via the random weight and approximate deterministic weight function is proposed. The weight will be chosen as a function of the empirical cumulative distribution of the underlying distribution of the Condroz and Secura Belgian reinsurance data set. The performance of the proposed estimators was compared with those of other semi-parametric tail index estimators that are available in the literature.

1.2 Scope of Study

Under the semi-parametric context, we present the weighted least squares estimator to estimate the tail index from Pareto-type distributions of a given data in the Condroz and Secura Belgian reinsurance data set. The performance is measured by plotting the MSE and bias against the top-order statistics which provides more information as possible about a distribution's tail.

1.3 Literature Review

Existing literature on reduced-bias least squares estimation of extreme value index are reviewed in this section.

Vynckier and Teugels [36] showed that the tail index estimators can be used to calculate the slope at the right upper tail of a Pareto quantile plot using a weighted least squares technique. The amount of extreme values taken into account when using extreme value statistics to estimate the Pareto tail index is crucial. When estimating Pareto indices in a slowly decreasing moving element of a Pareto type model, bias is one of the most difficult to calculate. When the bias problem is not as severe, well-known estimators like Hill [18] and the moment estimator, Dekkers et al. [14] are required for adaptive selection of the sample fraction to utilize in such estimation techniques. Beirlant et al. [3] showed that maximum likelihood estimators based on a regression model for upper order statistics can be used to design solutions for the presented issues. The bias-variance trade-off for a given data collection can also be inferred using this technique.

Gomesa and Martins [17] showed that the removal of a bias component of $a_{m,k}$ considerably improves the estimation of the tail index γ, both in terms of sample pathways and in terms of the stability region around the goal value, and from the point of view of the MSE structure, which now has a "bath-tube"' form as a function of k, making the choice of an optimal sample fraction less relevant and suggested that the behaviour of the ρ-estimator considered has indeed a high impact on the distributional properties of the final estimator of γ, and must be carefully chosen.

Ocran et al. [28] gave another reduced-bias estimator for the extreme value index, $\gamma > 0$. Their proposed estimator was found to be unbiased, consistent, and normal asymptotically. Furthermore, in terms of bias and mean square error, it performs favourably well compared with all the reduced-bias estimators studied. It also gives stable estimates that are slightly sensitive to the tail sample fraction (k), making it simple to utilize in determining the best value for k.

Using alternative reduced-bias estimators, Ocran et al. [29] presented regularised weighted least squares and weighted least squares. They concluded that the proposed estimators' tail index estimations were generally stable and smooth throughout a larger k path than the Hill estimator, least square estimator and the bias-corrected Hill estimator.

Buitendag et al. [10] applied the ridge regression to a distribution of Pareto-type with positive EVI and found that in terms of accuracy, the RMSE and bias lie between the least-squares estimators and the expectation. They concluded that the EVI estimates of the ridge regression are horizontal and more stable than the HILL estimator and the least square estimator.

1.4 Organisation of Study

The structure of this essay is divided into four (4) chapters. Chapter 1 contains a general introduction, the study's goal, the scope of the investigation, a literature review that examines the work of previous researchers on the issue, their techniques, and the study's organization. The suggested estimator's asymptotic features are discussed in Chapter 2 of the paper. The simulation design, discussion of simulation findings, and conclusion of these results are all included in Chapter 3 of the study. The study's fourth chapter contains practical examples.

2. Materials and Methods

This section introduces the reduced-bias weighted least square estimator, as well as its asymptotic features. We propose a regression model given as

$$Y_j = \gamma + a_{m,k}\delta_j + \varepsilon_j, \quad j \in \{1, 2, \cdots, k\} \tag{2.0.1}$$

where $a_{m,k} = a\left(\dfrac{k}{m}\right)$ is the slope of Y_j, $\delta_j = \left(\dfrac{j}{k+1}\right)^{-\rho}$ is the co-variate (an independent variable that can influence the outcome of a given statistical trial, but which is not of direct interest), γ is the intercept in (2.0.1), $\varepsilon_i \sim N\left(0, \gamma^2\right)$ and ρ is estimated externally due to it second-order nature.

The proposed estimator employs weighted least squares on the $Y_j's$ in (2.0.1), to deduce γ and $a_{m,k}$, the coefficient of the model from the loss function of the weighted least square is given by

$$L\left(\gamma, a_{m,k}; \ell\right) = \sum_{j=1}^{k} \ell_j \left(Y_j - \gamma - a_{m,k}\delta_j\right)^2 \tag{2.0.2}$$

for $1 \leq j \leq k$. We define the weight function as

$$\ell_j = \left(1 - \left[F(Y_j)\right]^{\theta(k)} \frac{j}{k+1}\right), 1 \leq j \leq k, \tag{2.0.3}$$

where $\theta(k)$ is chosen such that $\lim_{k\to\infty} \frac{\theta(k)}{k} < \infty$ and $F(Y_j)$ is the empirical cumulative distribution of Zj. We write

$$\Delta := \lim_{k\to\infty} \frac{\theta(k)}{k} < \infty \tag{2.0.4}$$

Thus, ℓ_j is decreasing linearly with respect to j, $\ell_j \in (0, 1)$.

We observe that $\theta(k) = \Delta k + 0(1)$, ℓ_j is random through $F(Y_j)$, and when the index $\theta(k) = 0$, ℓ_j is deterministic. To be specific, when $\theta(k) = 0$, the weight function $\ell_j = 1 - \dfrac{j}{k+1}$.

We minimise (2.0.1) with regard to γ and $a_{m,k}$ to derive the Weighted Least Squares (WLS) coefficients:

Differentiating the loss function partially in respect to γ

$$\frac{\partial L_j}{\partial \gamma} = -2\sum_{j=1}^{k} \ell_j \left(Y_j - \gamma - \hat{a}_{m,k}\delta_j\right)$$

$$0 = \sum_{j=1}^{k} \ell_j Y_j - \gamma \sum_{j=1}^{k} \ell_j - \hat{a}_{m,k} \sum_{j=1}^{k} \ell_j \delta_j \tag{2.0.5}$$

$$\gamma \sum_{j=1}^{k} \ell_j = \sum_{j=1}^{k} \ell_j Y_j - \hat{a}_{m,k} \sum_{j=1}^{k} \ell_j \delta_j \tag{2.0.6}$$

$$\gamma = \sum_{j=1}^{k} \overline{\ell_j} Y_j - \hat{a}_{m,k} \sum_{j=1}^{k} \overline{\ell_j} \delta_j . \tag{2.0.7}$$

$$\text{since } \overline{\ell_j} = \frac{\ell_j}{\sum_{j=1}^{k} \ell_j} \tag{2.0.8}$$

$$\tag{2.0.9}$$

Differentiating partially with respect to the $a_{m,k}$

$$\frac{\partial L_j}{\partial a_{m,k}} = -2\delta_j \sum_{j=1}^{k} \ell_j \left(Y_j - \gamma - a_{m,k}\delta_j\right) \tag{2.0.10}$$

$$0 = \sum_{j=1}^{k} \delta_j \ell_j Y_j - \gamma \sum_{j=1}^{k} \ell_j \delta_j - \hat{a}_{m,k} \sum_{j=1}^{k} \ell_j \delta_j^2 \tag{2.0.11}$$

$$\hat{a}_{m,k} \sum_{j=1}^{k} \ell_j \delta_j^2 = \sum_{j=1}^{k} \delta_j \ell_i Y_j - \gamma \sum_{j=1}^{k} \ell_j \delta_j, \text{ dividing through by } \sum_{j=1}^{k} \ell_j \tag{2.0.12}$$

$$\hat{a}_{m,k} \sum_{j=1}^{k} \left(\frac{\ell_j}{\sum_{j=1}^{k} \ell_j} \right) \delta_j^2 = \sum_{j=1}^{k} Y_j \left(\frac{\ell_j}{\sum_{j=1}^{k} \ell_j} \right) \delta_j - \gamma \sum_{j=1}^{k} \left(\frac{\ell_j}{\sum_{j=1}^{k} \ell_j} \right) \delta_j \tag{2.0.13}$$

$$\hat{a}_{m,k} \sum_{j=1}^{k} \overline{\ell_j} \delta_j^2 = \sum_{j=1}^{k} \overline{\ell_j} Y_j \delta_j - \gamma \sum_{j=1}^{k} \overline{\ell_j} \delta_j \tag{2.0.14}$$

From (2.0.7), $\gamma = \sum_{j=1}^{k=} \overline{\ell_j} Y_j - \hat{a}_{m,k} \sum_{j=1}^{k} \overline{\ell_j} \delta_j$. Putting Equation (2.0.7) into Equation (2.0.14), we have

$$\hat{a}_{m,k}\sum_{j=1}^{k}\overline{\ell_j}\delta_j^2 = \sum_{j=1}^{k}\delta_j\overline{\ell_j}Y_j - \left(\sum_{j=1}^{k}\overline{\ell_j}Y_j - \hat{a}_{m,k}\sum_{j=1}^{k}\overline{\ell_j}\delta_j\right)\sum_{j=1}^{k}\overline{\ell_j}\delta_j \quad (2.0.15)$$

$$\hat{a}_{m,k}\sum_{j=1}^{k}\overline{\ell_j}\delta_j^2 = \sum_{j=1}^{k}\delta_j\overline{\ell_j}Y_j - \sum_{j=1}^{k}\overline{\ell_j}Y_j\sum_{j=1}^{k}\overline{\ell_j}\delta_j + \hat{a}_{m,k}\left(\sum_{j=1}^{k}\overline{\ell_j}\delta_j\right)^2 \quad (2.0.16)$$

$$\hat{a}_{m,k}\sum_{j=1}^{k}\overline{\ell_j}\delta_j^2 - \hat{a}_{m,k}\left(\sum_{j=1}^{k}\overline{\ell_j}\delta_j\right)^2 = \sum_{j=1}^{k}\delta_j\overline{\ell_j}Y_j - \sum_{j=1}^{k}\overline{\ell_j}Y_j\sum_{j=1}^{k}\overline{\ell_j}\delta_j \quad (2.0.17)$$

$$\hat{a}_{m,k}\left[\sum_{j=1}^{k}\overline{\ell_j}\delta_j^2 - \left(\sum_{j=1}^{k}\overline{\ell_j}\delta_j\right)^2\right] = \sum_{j=1}^{k}\overline{\ell_j}\left(\delta_j - \sum_{j=1}^{k}\overline{\ell_j}\delta_j\right)Y_j \quad (2.0.18)$$

$$\hat{a}_{m,k} = \frac{\sum_{j=1}^{k}\overline{\ell_j}\left(\delta_j - \sum_{j=1}^{k}\overline{\ell_j}\delta_j\right)Y_j}{\sum_{j=1}^{k}\overline{\ell_j}\delta_j^2 - \left(\sum_{j=1}^{k}\overline{\ell_j}\delta_j\right)^2}. \quad (2.0.19)$$

where $0 \le \ell_j \le 1$ and $\sum_{j=1}^{k}\overline{\ell}_j = 1$, by normalising ℓ_j.

We write

$$\mathfrak{A}_k(F,\delta_j) := \frac{\sum_{j=1}^{k}\overline{\ell_j}\left(\delta_j - \sum_{j=1}^{k}\overline{\ell_j}\delta_j\right)}{\sum_{j=1}^{k}\overline{\ell_j}\delta_j^2 - \left(\sum_{j=1}^{k}\overline{\ell_j}\delta_j\right)^2}. \quad (2.0.20)$$

2.1 The Proposed Estimators' Asymptotic Properties

We shall show in this section that the weighted least square estimator is asymptotically normal, unbiased and consistent. To show that the estimator is unbiased, we have to show that the bias$(\gamma) = 0$, where bias is expressed as $\mathbb{E}[\gamma] - \gamma$.

2.1.1 Lemma. [29] Assume that ρ is estimated by a consistent estimator $\hat{\rho}$ and Equation (2.0.4) holds, then as $k \to \infty$ and $\frac{k}{m} \to 0$;

i. $S_1(\rho, F) = \sum_{j=1}^{k}\overline{\ell_j}\delta_j \xrightarrow{a.s} \frac{1}{1-\rho}$.

ii. $S_2(\rho, F) = \sum_{j=1}^{k} \overline{\ell_j} \delta_j^2 - \left(\sum_{j=1}^{k} \overline{\ell_j} \delta_j \right)^2 \xrightarrow{a.s} \dfrac{\rho^2}{(1-2\rho)(1-\rho)^2}.$

See Appendix for the proof of lemma 2.1.1.

2.1.2 Lemma (Unbiasedness). Suppose ρ is estimated by a consistent $\hat{\rho}$ satisfying

$$\mathbb{E}_{\hat{\rho}} \left[\frac{1}{1-\hat{\rho}} \right] < \infty.$$

Then,$\hat{\gamma}$ is unbiased.i.e. $\lim_{k\to\infty} \mathrm{bias}(\gamma) = \lim_{k\to\infty} \mathbb{E}[\gamma] - \gamma = 0.$

Proof. Note $-\mathcal{O}(1) \le [F(Y_j)]^{k\Delta} \le \mathcal{O}(1)$, for all j, almost surely and hence we have

$$\left(1 - \mathcal{O}(1)\frac{j}{k+1}\right) \le \ell_j \le \left(1 + \mathcal{O}(1)\frac{j}{k+1}\right).$$

This gives

$$\frac{\pi(k)}{k}\Big(\sum_{j=1}^{k} Y_j - \hat{a}_{m,k} \sum_{j=1}^{k} \delta_j(\hat{\rho}) \Big) - \mathcal{O}(1)\frac{\pi(k)}{k} \sum_{j=1}^{k} |Y_j - \delta_j| \le \hat{\gamma} \tag{2.1.1}$$

$$\le \frac{\pi(k)}{k}\Big(\sum_{j=1}^{k} Y_j - \hat{a}_{m,k} \sum_{j=1}^{k} \delta_j(\hat{\rho}) \Big) + \mathcal{O}(1)\frac{\pi(k)}{k} \sum_{j=1}^{k} |Y_j - \delta_j| \tag{2.1.2}$$

Now

$$\begin{aligned}
\mathbb{E}\left(|Y_j - \delta_j| \big| \hat{\rho} = \rho\right) &= \int_0^\infty |Y_j - \delta_j| f(Y_j)\, dY_j \\
&= \int_{\delta_j}^\infty (Y_j - \delta_j)\, f(Y_j)\, dY_j - \int_0^{\delta_j} (Y_j - \delta_j)\, f(Y_j)\, dY_j \\
&= \int_{\delta_j}^\infty (Y_j - \delta_j)\, \frac{1}{\delta_j} e^{\left(-\frac{1}{\delta_j} Y_j\right)} dY_j - \int_0^{\delta_j} (Y_j - \delta_j)\, \frac{1}{\delta_j} e^{\left(-\frac{1}{\delta_j} Y_j\right)} dY_j \\
&= \int_0^\infty \frac{\mu}{\delta_j} e^{\left(-\frac{(\mu+\delta_j)}{\delta j}\right)} d\mu - \int_{\delta_j}^0 \frac{\mu}{\delta_j} e^{\left(-\frac{(\mu+\delta_j)}{\delta_j}\right)} d\mu
\end{aligned}$$

$$= e^{-1}\left(\int_0^\infty \frac{\mu}{\delta_j} e^{-\dfrac{1}{\delta_j}\mu} d\mu - \int_{-\delta_j}^0 \frac{\mu}{\delta_j} e^{-\dfrac{\mu}{\delta j}} d\mu\right)$$

$$= e^{-1}\left(\int_0^\infty \frac{\mu}{\delta_j} e^{-\dfrac{\mu}{\delta_j}} d\mu - \int_{-\delta_j}^0 \left(\frac{\mu}{\delta_j} - \frac{\mu^2}{\delta_j^2}\right) d\mu\right)\{1+\mathcal{O}(1)\}$$

$$= e^{-1}\left(\frac{1}{\delta_j}\frac{1!}{\left(\dfrac{1}{\delta_j}\right)^2} + \frac{\delta_j}{2} + \frac{\delta_j}{3}\right)\{1+\mathcal{O}(1)\}$$

$$= \frac{11\delta_j}{6e}\{1+\mathcal{O}(1)\},$$

which gives $\mathbb{E}\left(|Y_j-\delta_j|\right) = \mathbb{E}\left\{\mathbb{E}\left(|Y_j-\delta_j|\middle|\hat{\rho}\right)\right\} = \dfrac{11}{6e}\{1+\mathcal{O}(1)\}\,\mathbb{E}_{\hat{\rho}}\,\delta_j(\hat{\rho})$. Using (2.1.2) we have

$$\frac{\pi(k)}{k}\Big(\sum_{j=1}^k \mathbb{E}(Y_j) - \hat{\mathbb{E}}_{\hat{\rho}}\Big[a_{m,k}(\hat{\rho})\sum_{j=1}^k \delta_j(\hat{\rho})\Big]\Big) - \mathcal{O}(1)\frac{\pi(k)}{k}\sum_{j=1}^k \frac{11}{6e}\{1+\mathcal{O}(1)\}\,\mathbb{E}_{\hat{\rho}}\,\delta_j(\hat{\rho})$$
$$\le \mathbb{E}(\hat{\gamma})$$
$$\le \frac{\pi(k)}{k}\Big(\sum_{j=1}^k \mathbb{E}(Y_j) - \hat{\mathbb{E}}_{\hat{\rho}}\Big[a_{m,k}(\hat{\rho})\sum_{j=1}^k \delta_j(\hat{\rho})\Big]\Big) + \mathcal{O}(1)\frac{\pi(k)}{k}\sum_{j=1}^k \frac{11}{6e}\{1+\mathcal{O}(1)\}\,\mathbb{E}_{\hat{\rho}}\,\delta_j(\hat{\rho}) \tag{2.1.3}$$

This gives

$$\pi(k)\gamma - \mathcal{O}(1)\pi(k)\frac{11}{6e}\{1+\mathcal{O}(1)\}\,\mathbb{E}_{\hat{\rho}}\left[\frac{1}{1-\hat{\rho}}\right] \tag{2.1.4}$$

$$\le \mathbb{E}(\gamma) \le \pi(k)\gamma - \mathcal{O}(1)\pi(k)\frac{11}{6e}\{1+\mathcal{O}(1)\}\,\mathbb{E}_{\hat{\rho}}\left[\frac{1}{1-\hat{\rho}}\right] \tag{2.1.5}$$

Taking limit of both sides we obtain

$$\lim_{k\to\infty} \mathrm{bias}(\gamma) = \lim_{k\to\infty} \mathbb{E}[\gamma] - \gamma = \gamma - \gamma = 0. \tag{2.1.6}$$

□

2.1.3 Lemma (Asymptotic Consistency). Suppose ρ is estimated by consistent $\hat{\rho}$ satisfying $\mathbb{E}_{\hat{\rho}}\left[\dfrac{1}{1-\hat{\rho}}\right] < \infty$. Then, as $k\to\infty$ we have

(i) $\sqrt{k}\hat{a}_{m,k}(\hat{\rho}) \to 0_p$

(ii) $\mathrm{AMSE}(\gamma) \to 0.$

Proof. (i) The proof of $\sqrt{k}\hat{a}_{m,k}(\hat{\rho}) \to 0_p$ requires the use of Large Deviations (LD). Thus we shall show via LD that, $\forall\ \epsilon > 0$, then, $\lim_{k\to\infty} \mathbb{P}\left(\sqrt{k}\hat{a}_{m,k} \geq \epsilon\right) = 0.$

From Equation (2.0.19) and Equation (2.0.20), $\hat{a}_{m,k}(F, \hat{\rho}) = \sum_{j=1}^{k} \mathfrak{A}_k(F, \delta_j)Y_j$. Given $\{\hat{\rho} = \rho\}$, Y_j is exponentially distributed with mean

$$\mu_j(\rho) = \mathbb{E}[Y_j|\hat{\rho} = \rho] = \gamma + \hat{a}_{m,k}(\hat{\rho}, \theta)\delta_j(\hat{\rho}).$$

Hence, the moment-generating function of $\sqrt{k}a_{m,k}$ is

$$N_{\hat{a}_{m,k}(\rho)}(kt) = \left\{\prod_{j=1}^{k}\left(\frac{1}{1 - k\mu_j(\rho)\mathfrak{A}_k(F, \delta_j)t}\right)\right\}. \tag{2.1.7}$$

Notice from [28, Lemma 2.3.2] that

$$-\frac{k^{\rho}}{|k^{1+\tau+\rho} + m^{\rho}|} \leq \mathfrak{A}_k(F, \delta_j) \leq \frac{k^{\rho}\delta_j}{S_2(\rho, F)\,|k^{1+\tau+\rho} + m^{\rho}|},$$

and therefore, we have $\mathfrak{A}_k(F, \delta_j) \to \mathcal{O}\left(\frac{1}{k^{1+\tau}}\right)$, $0 < \tau \leq 0.1.$ as $k \to \infty$. Choose $1 > \phi > 0$ such that $-\phi \leq k\mu_j\mathfrak{A}_k(F, \delta_j) \leq \phi$. then we have,

$$\prod_{j=1}^{k}\left(\frac{1}{1+\phi}\right) \leq N_{a_{m,k}}(kt) \leq \prod_{j=1}^{k}\left(\frac{1}{1-\phi}\right) \tag{2.1.8}$$

$$\left(\frac{1}{1+\phi}\right)^{k} \leq N_{a_{m,k}}(kt) \leq \left(\frac{1}{1-\phi}\right)^{k} \tag{2.1.9}$$

$$\left(1 - \phi + \phi^2 - \phi^3 + \ldots\right)^k \leq N_{a_{m,k}}(kt) \leq \left(1 + \phi + \phi^2 + \phi^3 + \ldots\right)^k \tag{2.1.10}$$

$$k\log\left(1 - \delta + \phi^2 - \phi^3 + \ldots\right) \leq \log N_{a_{m,k}}(kt) \leq k\log\left(1 + \phi + \phi^2 + \phi^3 + \ldots\right) \tag{2.1.11}$$

$$\log\left(1 - \phi + \phi^2 - \phi^3 + \ldots\right) \leq \lim_{k\to\infty}\frac{1}{k}\log N_{a_{m,k}}(kt) \leq \log\left(1 + \phi + \phi^2 + \phi^3 + \ldots\right) \tag{2.1.12}$$

Using the Squeeze Theorem and setting $\phi \longrightarrow 0$, we get

$$\lim_{k\to\infty}\frac{1}{k}\log N_{\hat{a}_{m,k}}(kt) = 0.$$

As a result, we obtain 0 through the convergence of the logarithmic moment-generating function with speed k, and $a_{m,k}$ follows the Large Deviation Principle (LDP), according to the Gärtner

Ellis Theorem and a $I(y)$, rate function expressed as

$$I(y) = \sup_{\psi\in\mathbb{R}}\{\psi y - 0\} = \sup_{\psi\in\mathbb{R}}\{\psi y\} \tag{2.1.13}$$

$$= \begin{cases} 0 \text{ if } y = 0 \\ \infty \text{ if } y \neq 0 \end{cases} \tag{2.1.14}$$

where $y \in [0,\infty)$ and $\psi > 0$. If $a_{m,k}$ satisfies the LDP with $I(y)$, then $\exists\ \epsilon > 0$ such that

$$\mathbb{P}\left(\sqrt{k}\hat{a}_{m,k} > \epsilon\right) = \mathbb{P}\left(\hat{a}_{m,k} > \frac{\epsilon}{\sqrt{k}}\Big|\hat{\rho} = \rho\right)\mathbb{P}\left[\hat{\rho} = \rho\right] \tag{2.1.15}$$

$$= \mathbb{P}\left(\hat{a}_{m,k} > b_k\Big|\hat{\rho} = \rho\right)\mathbb{P}\left[\hat{\rho} = \rho\right] \leq e^{-kI(y)+o(k)}, \tag{2.1.16}$$

where $b_k = \epsilon/\sqrt{k}$. When $y \neq 0$, we achieve the typical behaviour of $I(y)$. Consequently,

$$\lim_{k\to\infty}\mathbb{P}\left(\sqrt{k}a_{m,k} > \epsilon\right) \leq 0.$$

Thus, $\sqrt{k}a_{m,k} \to_p 0$ as $k \to \infty$ and the proof is now complete.

(ii) $\mathrm{AMSE}(\gamma) \to 0$. We observe from Lemma 2.1.3 (i) that $\gamma = \sum_{j=1}^{k}\bar{\ell}_j Y_j + \mathcal{O}_p(1)$ and therefore we have

$$\mathrm{Var}(\gamma) = \frac{\pi^2(k)}{k^2}\sum_{j=1}^{k}\mathrm{Var}\left(\ell_j Y_j\right).$$

Now, we note $\left(1 - \mathcal{O}(1)\dfrac{j}{k+1}\right)^2 \mathrm{Var}\left(Y_j\right) \leq \mathrm{Var}\left(\ell_j Y_j\right) \leq \left(1 + \mathcal{O}(1)\dfrac{j}{k+1}\right)^2 \mathrm{Var}\left(Y_j\right)$

Therefore we have

$$\mathrm{Var}(\gamma) = \frac{\pi^2(k)}{k^2}\left[\sum_{j=1}^{k}\mathrm{Var}\left(Y_j\right) + \mathcal{O}(1)\right] \tag{2.1.17}$$

$$= \frac{\pi^2(k)}{k^2}\left[\sum_{j=1}^{k}\gamma^2 + \mathcal{O}(1)\right] \tag{2.1.18}$$

$$= \gamma^2\frac{\pi^2(k)}{k} + \frac{\pi^2(k)}{k}\mathcal{O}(1) \tag{2.1.19}$$

We estimate the asymptotic mean square as follows:

$$\lim_{k\to\infty}\mathrm{AMSE}(\gamma) \leq \lim_{k\to\infty}\left[\gamma^2\frac{\pi^2(k)}{k} + \frac{\pi^2(k)}{k}\mathcal{O}(1)\right] + \lim_{k\to\infty}\left(\mathbb{E}(\hat{\gamma}) - \gamma\right)^2 = \left(\lim_{k\to\infty}\mathbb{E}(\hat{\gamma}) - \gamma\right)^2 = 0$$

□

2.1.4 Theorem (Asymptotic Normality). *[29] Suppose Equation* (1.0.2) *and Equation* (2.1.2) *are satisfied. Assume also that* ρ *is estimated by a consistent estimator* $\hat{\rho}$ *that satisfies* $\mathbb{E}_{\hat{\rho}}\left[\frac{1}{1-\hat{\rho}}\right] < \infty$. *Then if Conditions* 1 *and* 2 *hold, and* $\sqrt{k}a_{m,k} \to 0$, *then we have*

$$\sqrt{k}\,(\hat{\gamma} - \gamma) \to_d N(0, \gamma^2).$$

Proof. First we can use Lemma 2.1.3 (i) and $\left(1 - \mathcal{O}(1)\frac{j}{k+1}\right) \leq \ell_j \leq \left(1 + \mathcal{O}(1)\frac{j}{k+1}\right)$ to reduce the estimator to

$$\sqrt{k}\hat{\gamma} \stackrel{d}{=} \frac{\sqrt{k}\pi(k)}{k}\sum_{j=1}^{k}\ell_j Y_j = \frac{\sqrt{k}\pi(k)}{k}\sum_{j=1}^{k} Y_j + \sqrt{k}\mathcal{O}(1)_p(1/k) = \sqrt{k}\pi(k)\gamma^H + \sqrt{k}\mathcal{O}(1)_p(1/k)$$

Therefore, it follows that

$$\begin{aligned}
\sqrt{k}(\hat{\gamma} - \gamma) &\stackrel{d}{=} \sqrt{k}\left(\pi(k)\hat{\gamma}^H - \gamma\right) - \pi(k)\sqrt{k}\gamma + \pi(k)\sqrt{k}\gamma + \sqrt{k}\mathcal{O}(1)_p(1/k) \\
\sqrt{k}(\hat{\gamma} - \gamma) &\stackrel{d}{=} \sqrt{k}\pi(k)\hat{\gamma}^H - \sqrt{k}\gamma - \pi(k)\sqrt{k}\gamma + \pi(k)\sqrt{k}\gamma + \sqrt{k}\mathcal{O}(1)_p(1/k) \\
\sqrt{k}(\hat{\gamma} - \gamma) &\stackrel{d}{=} \sqrt{k}\pi(k)\left(\hat{\gamma}^H - \gamma\right) + \sqrt{k}\gamma(\pi(k) - 1) + \sqrt{k}\mathcal{O}(1)_p(1/k) \\
\sqrt{k}(\hat{\gamma} - \gamma) &\stackrel{d}{=} \sqrt{k}\pi(k)\left(\hat{\gamma}^H - \gamma\right) + \sqrt{k}\Big(\gamma(\pi(k) - 1) + \mathcal{O}(1)_p(1/k)\Big) \\
\sqrt{k}(\hat{\gamma} - \gamma) &\stackrel{d}{=} \sqrt{k}\pi(k)\left(\hat{\gamma}^H - \gamma\right) + \sqrt{k}\Big(\gamma\mathcal{O}(1)(1/k) + \mathcal{O}(1)_p(1/k)\Big) \\
\sqrt{k}(\hat{\gamma} - \gamma) &\stackrel{d}{=} \sqrt{k}\pi(k)\left(\hat{\gamma}^H - \gamma\right) + \mathcal{O}(1)(1/\sqrt{k})
\end{aligned}$$

□

Using the asymptotic normality properties of Hill, we have that

$$\sqrt{k}(\hat{\gamma} - \gamma) \stackrel{d}{=} \pi(k)\sqrt{k}(\hat{\gamma}^H - \gamma) + \mathcal{O}(1)(1/\sqrt{k}) \xrightarrow{d} N(0, \gamma^2),$$

and thus, the completion of the proof of 2.1.4.

3. Simulations

Simulations were carried out to investigate the finite sample properties and sample path behaviour of the proposed EVI estimator. 1000 replications of samples of size $m = 50$ and 200 are generated for Fréchet and Burr distribution as shown in Table 3.1. These two distributions are employed because they belong to the Hall class of distributions and also they are commonly used in simulation studies see [27, 10, 26]. With regard to the following EVI, we consider $\gamma = 0.1, 0.5,$ and 1.0 respectively due to its constant use in simulation studies in [10, 27, 26]. The performance of our proposed estimators, the Hill [18], the Beirlant et al. [4], the Caeiro et al. [11], These estimators are investigated by plotting the bias and the MSE as a function of k see [28], to enable us scrutinize the behaviour of the estimates as k change. We consider $\alpha = 10, 2$ and 1.0 for the Fréchet distributions, and the mixes for the Burr XII distribution with

i $\eta = \sqrt{10}$, $\tau = \sqrt{10}$,

ii $\eta = \sqrt{2}$, $\tau = \sqrt{2}$ and

iii $\eta = 2$, $\tau = 1/2$.

The notations of the estimators of the EVI employed in the simulations are shown in Table 3.2

3.1 Simulation Design

In this simulation, we consider two classes for the proposed EVI estimates depending on the behaviour of $\theta(k)$. In the case of $\ell_j = \left(1 - (F(Y_j))^{\theta(k)} \frac{j}{k+1}\right)$, the $F(Y_j)$ will be replaced with it's empirical counter part. These classes are used when the weight function is considered to be approximately deterministic and when it is considered to be random.

Distribution	f(y)	Mean	Variance	γ
Burr type XII	$\frac{\tau\eta y^{\tau-1}}{(1+y^\tau)^{1+\eta}}$	$\eta\beta\left(\eta - \frac{1}{\tau}, 1 + \frac{1}{\tau}\right)$	$\eta\beta\left(\eta - \frac{2}{\tau}, 1 + \frac{2}{\tau}\right) - \eta^2\beta^2\left(\eta - \frac{1}{\tau}, 1 + \frac{1}{\tau}\right)^2$	$\frac{1}{\tau\eta}$
Fréchet	$\frac{\alpha}{Y^{\alpha+1}} e^{-y^{-\alpha}}$	$\Gamma\left(1 - \frac{1}{\alpha}\right)$	$\Gamma\left(1 - \frac{2}{\alpha}\right) - \left(\Gamma\left(1 - \frac{1}{\alpha}\right)\right)^2$	$\frac{1}{\alpha}$

Table 3.1: Distributions from the Pareto-type distribution with heavy-tailed

where Γ and β are gamma and Beta functions.

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

Table 3.2: Notations of the Estimators

The following algorithm is used in the simulation process:

i. Simulate $m(50 \text{ and } 200)$ random observations, say $Y_1, Y_2, \ldots, Y_m$ from the Fréchet and Burr type XII distribution.

ii. Transform the observation into the weighted log-spacing i.e $Z_j = j \ln\left(\dfrac{Y_{m-j+1,m}}{Y_{m-j,m}}\right)$.

iii. Compute the Covariate values (δ_j) for $j = 1, 2, \ldots, k$.

iv. Compute ρ (second-order parameter), using the minimum variance approach introduced by Buitendag et al. [10].

v. Compute the weight, $\ell_j = \left(1 - \left[F(Y_j)\right]^{\theta(k)} \dfrac{j}{k+1}\right)$. The $F(Y_j)$ in the weight function is estimated using the empirical cumulative distribution function of the weighted log-spacing, considering $\theta(k) = \dfrac{1}{k}$ and k.

vi. Compute the values of $a_{m,k}$ at each given values of k for the samples in step (ii).

vii. Compute the EVI at each values of k for the sample in step (ii).

viii. Repeat step (i) to step (vii), R replications1 $(R = 1000)$ to obtain $\hat{\gamma}_{1,j}, \hat{\gamma}_{2,j}, \hat{\gamma}_{3,j}, \ldots, \hat{\gamma}_{j,k}$ at each values of k.

ix. At each k, compute the bias$(\hat{\gamma}) = \dfrac{1}{R}\displaystyle\sum_{j=1}^{R}(\hat{\gamma}_{j,k} - \gamma)$ and MSE$(\hat{\gamma}) = \dfrac{1}{R}\displaystyle\sum_{j=1}^{R}(\hat{\gamma}_{j,k} - \gamma)^2$.

x. Plot the bias$(\hat{\gamma})$ and MSE$(\hat{\gamma})$ against k separately.

3.2 Discussion of Simulation Results

Simulations are used to compare the suggested estimator's performance to that of other existing estimators in the literature.

Figures 3.1 – 3.6 illustrate the results for the Burr type XII and Fréchet distribution with various tail indexes for approximately deterministic weight. Also, Figures 3.7 – 3.12 show a simulation display for the Burr type XII and Fréchet distribution with different tail indexes for random weight functions.

3.2.1 EVI Estimators with Approximated Deterministic Weight

Figures 3.1 – 3.3 illustrate the simulation display for the Burr distribution with various tail indexes that are approximately deterministic. In this first class, we pick $\theta(k)$ in a way that $\Delta \to 0$, as $k \to \infty$. In particular, we choose $\theta(k) = \dfrac{1}{k}$ using the Burr type XII distribution for $m(50$ and $200)$ with $\gamma = 0.1, 0.5.1.0$ respectively.

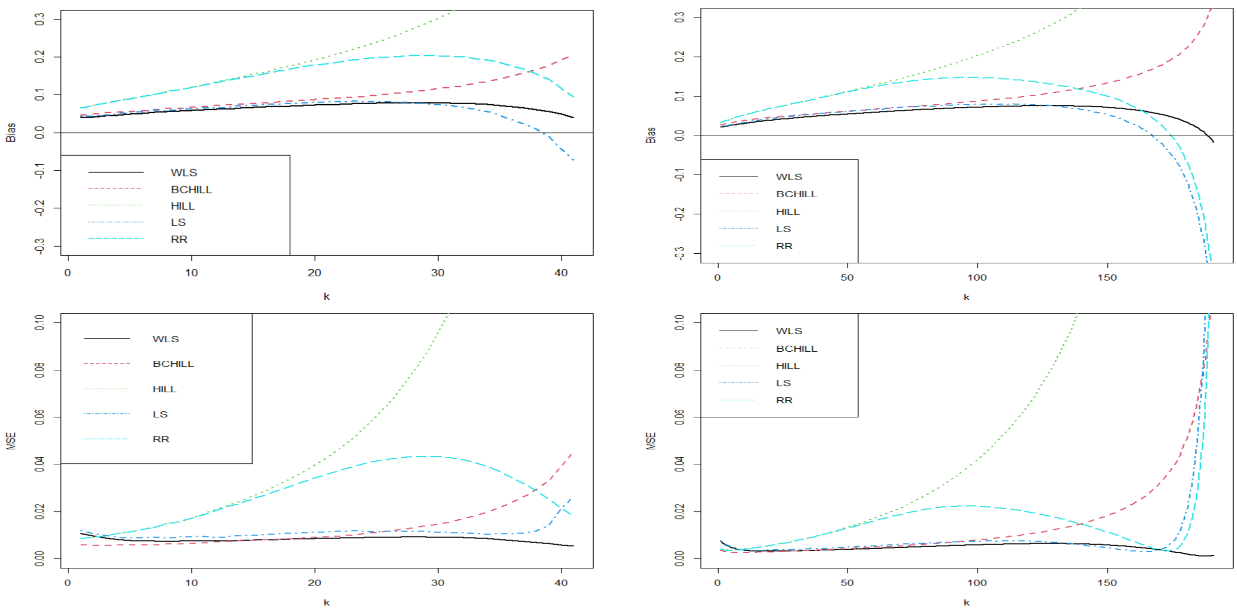

Figure 3.1: $\gamma = 0.1$

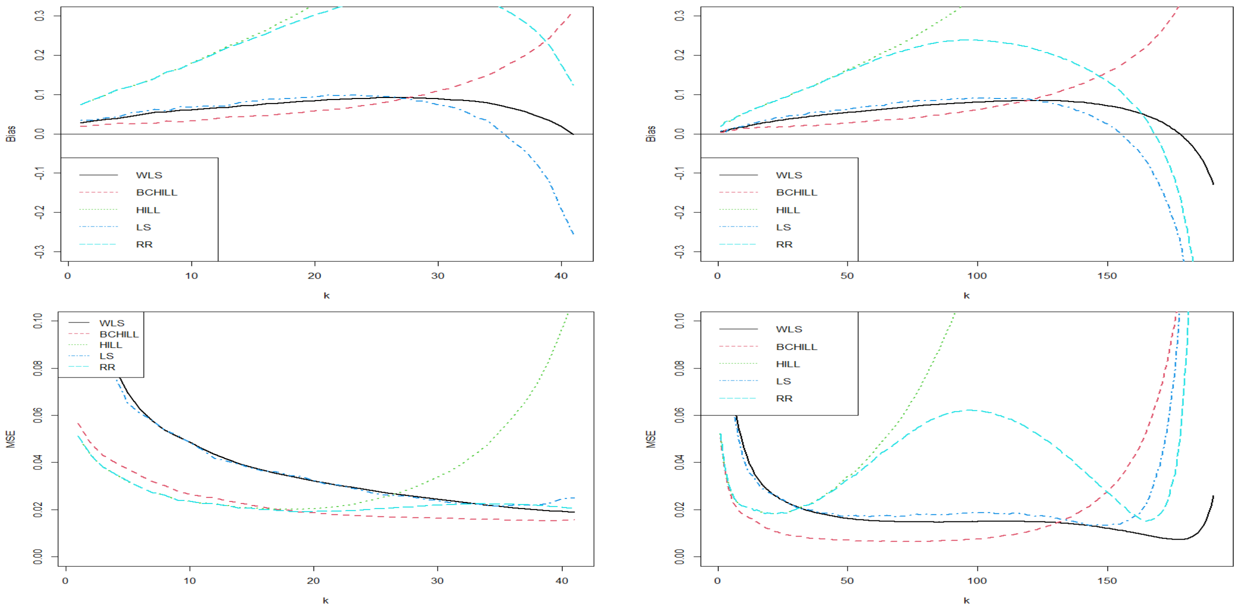

Figure 3.2: $\gamma = 0.5$

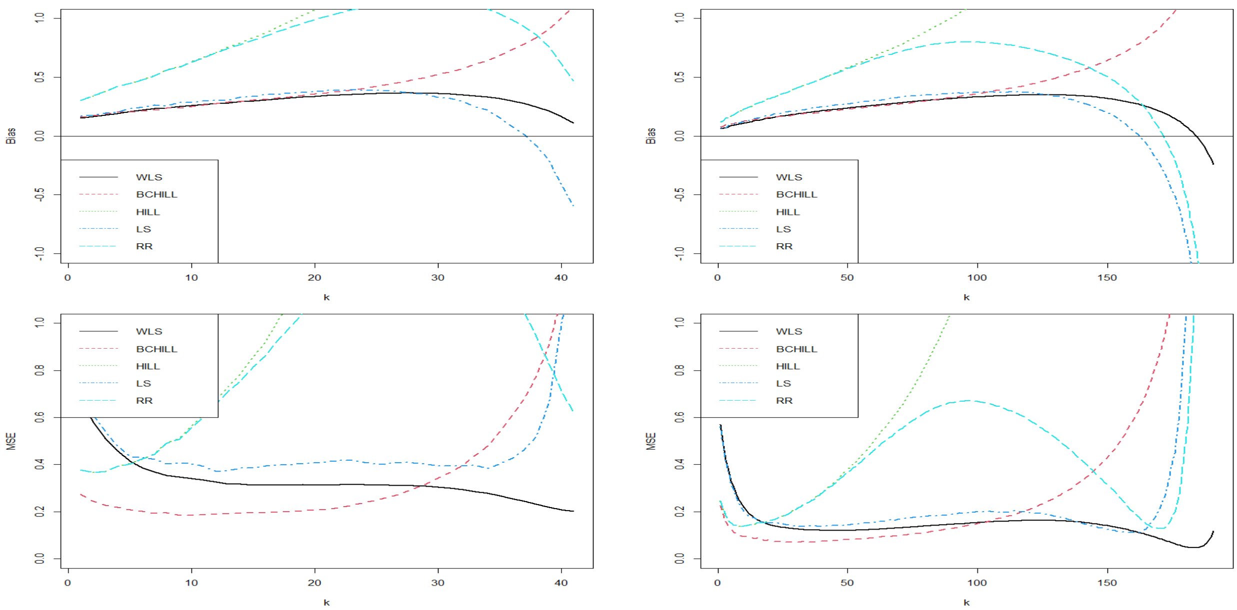

Figure 3.3: $\gamma = 1.0$

Figure 3.1 $-$ 3.3 presents the EVI estimators with approximated deterministic weight using Burr type XII distribution. It can be observed that, WLS performs favourably well in terms of MSE and bias. The WLS and LS mostly have the same sample path for small to central values of k in terms of bias and in terms of MSE when $m = 200$. Except for the WLS, the HILL, BCHILL, LS, and RR estimators are very sensitive to large values of k. The WLS is less sensitive to k as compared to the other estimators.

Figures 3.4 – 3.6 illustrate the simulation display for the Fréchet distribution with various tail indexes that are approximately deterministic. In this case, we pick $\theta(k)$ in a way that $\Delta \to 0$, as $k \to \infty$, in particular we choose $\theta(k) = \dfrac{1}{k}$ using the Fréchet distribution for $m(50$ and $200)$ with $\gamma = 0.1, 0.5.1.0$ respectively.

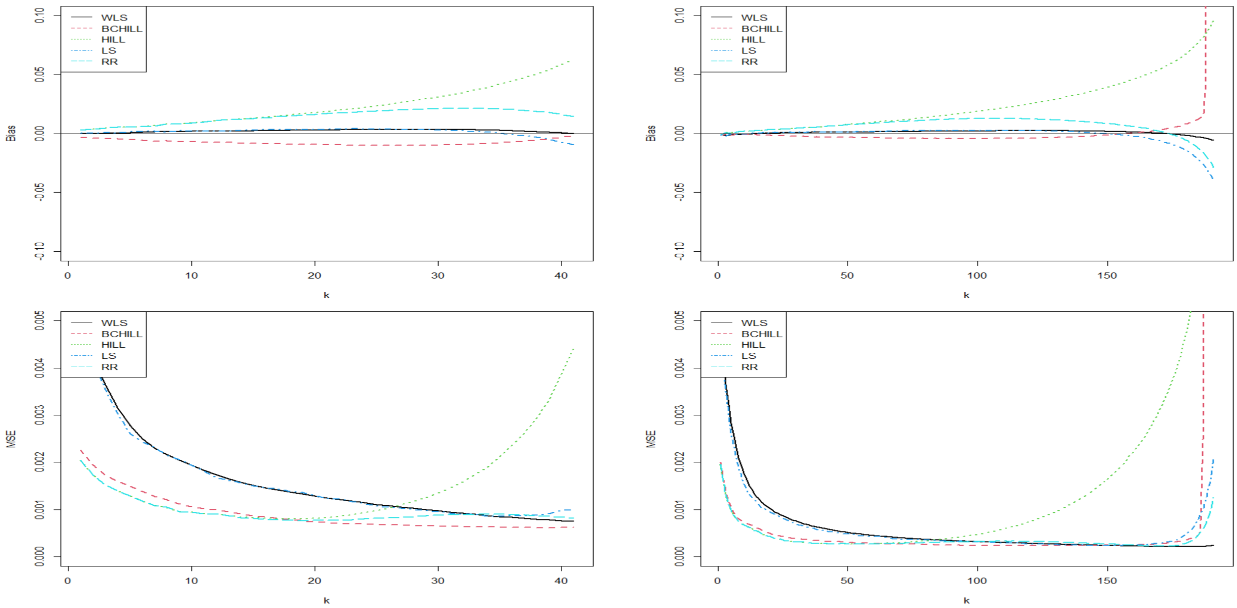

Figure 3.4: $\gamma = 0.1$

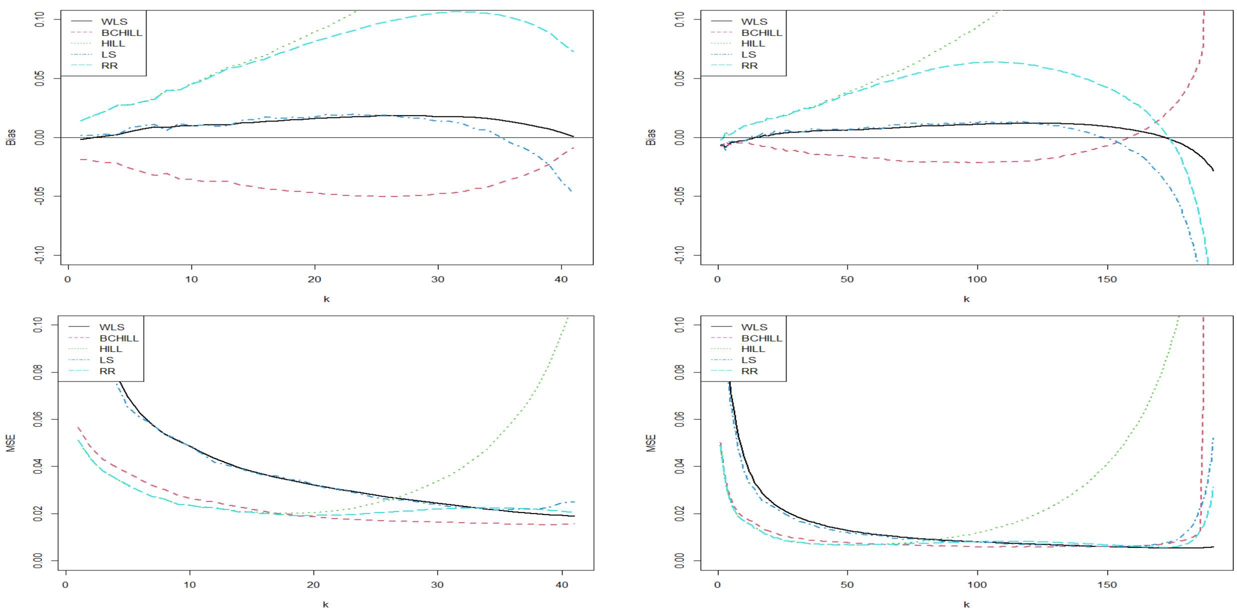

Figure 3.5: $\gamma = 0.5$

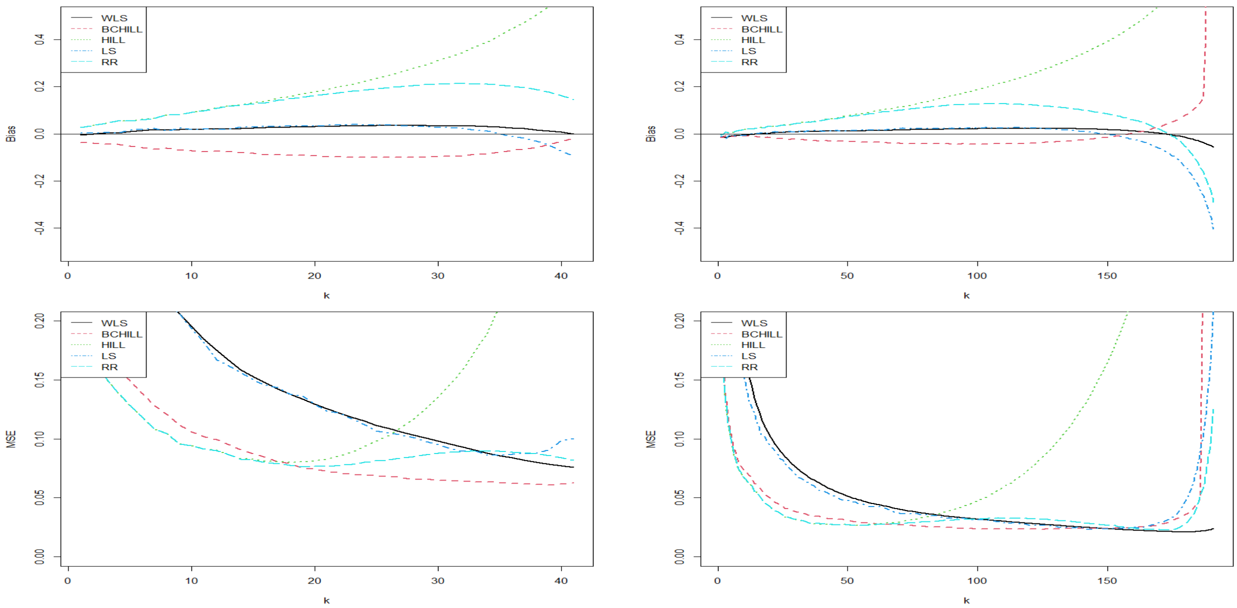

Figure 3.6: $\gamma = 1.0$

Figure 3.4 – 3.6 presents the EVI estimators with approximated deterministic weight using Fréchet distribution. It can be observed that, WLS is performs favourably well in terms of MSE and bias.The bias is almost horizontal. The WLS, BCHILL and LS mostly have the same sample path for small to central values of k in terms of bias and in terms of MSE when $m = 200$.

3.2.2 EVI Estimators with Random Weight

Figures 3.7 – 3.9 illustrate the simulation display for the Burr distribution with various tail indexes with random weight. For this second class, we consider $\theta(k)$ such that Δ approaches 1, as k approaches infinity, particularly, we select $\theta(k) = k$ using the Burr type XII distribution for $m(50$ and $200)$ with $\gamma = 0.1, 0.5.1.0$ respectively.

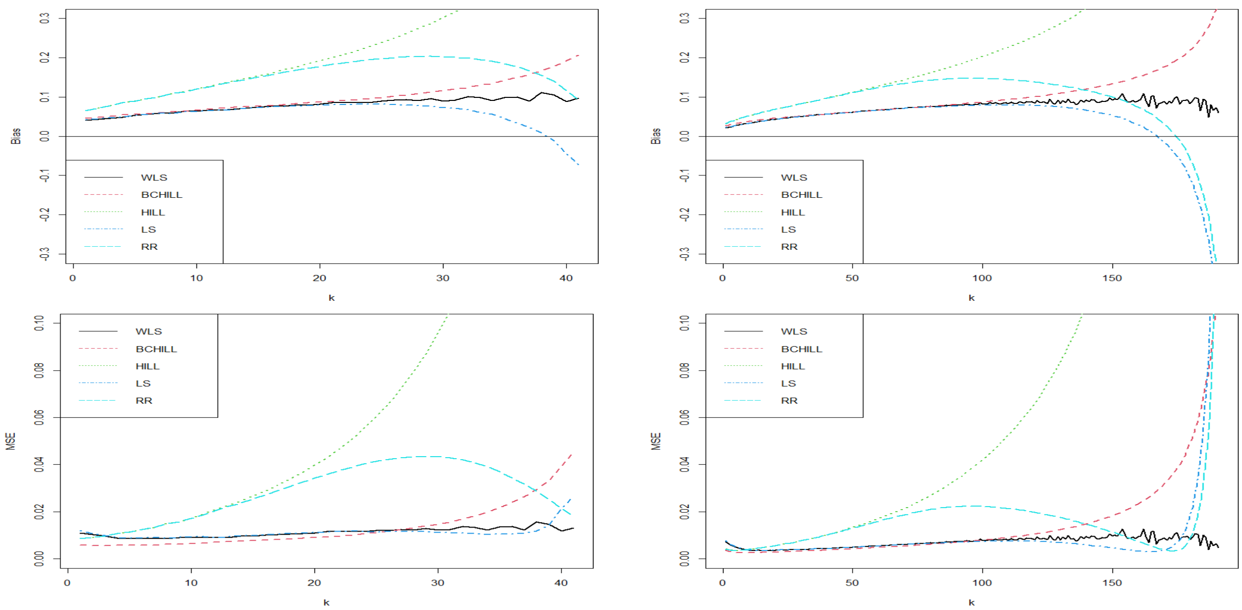

Figure 3.7: $\gamma = 0.1$

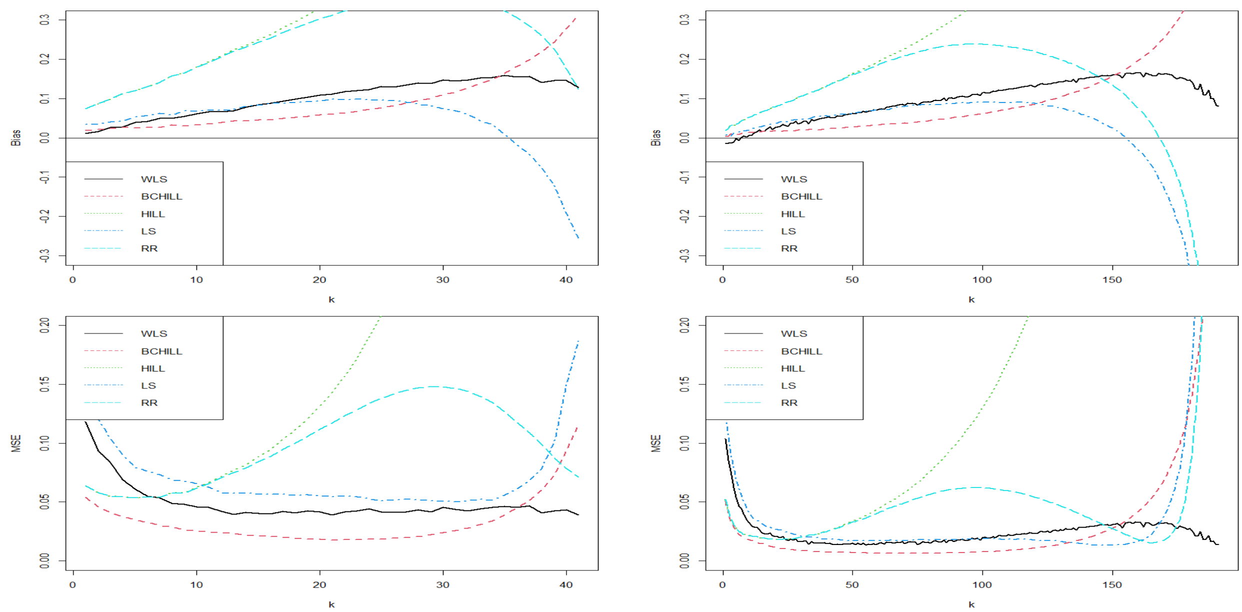

Figure 3.8: $\gamma = 0.5$

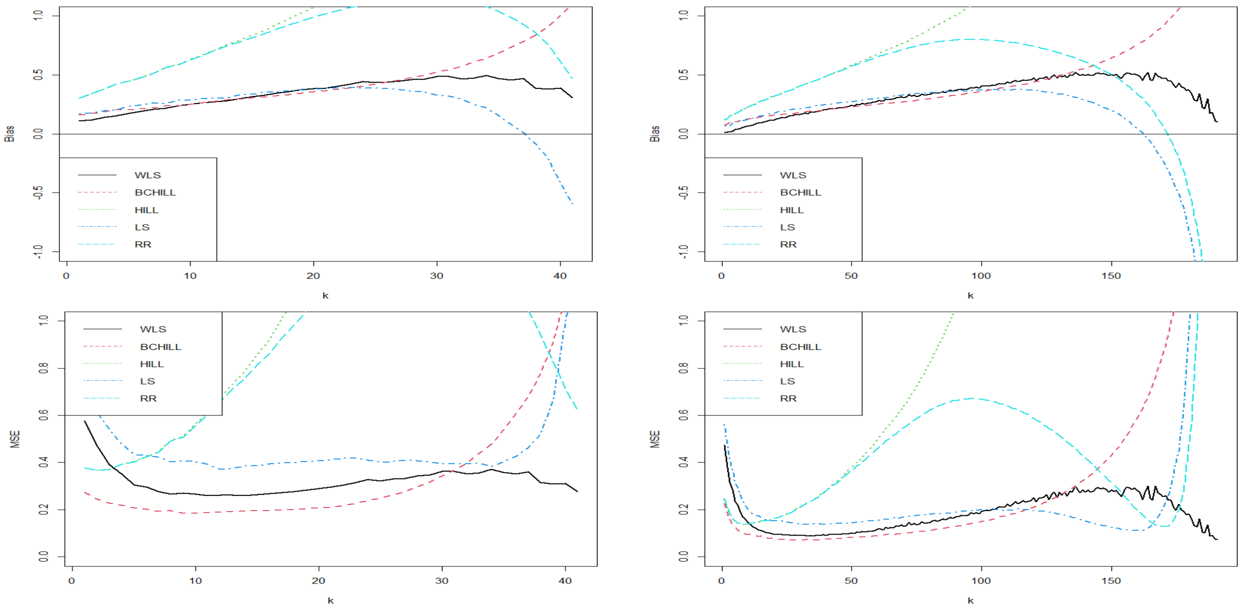

Figure 3.9: $\gamma = 1.0$

From Figure 3.7 – 3.9, It can be observed that WLS performs favourably well with the BCHILL in terms of MSE and bias for $(\gamma = 0.1, 1.0)$ and can be seen as the second best estimator to the BCHILL.

Figures 3.10 − 3.12 illustrate the simulation display for the Fréchet distribution with various tail indexes with random weight. For this second class, we consider $\theta(k)$ such that Δ approaches 1, as k approaches infinity, particularly, we select $\theta(k) = k$ using the Fréchet distribution for $m(50$ and $200)$ with $\gamma = 0.1, 0.5.1.0$ respectively.

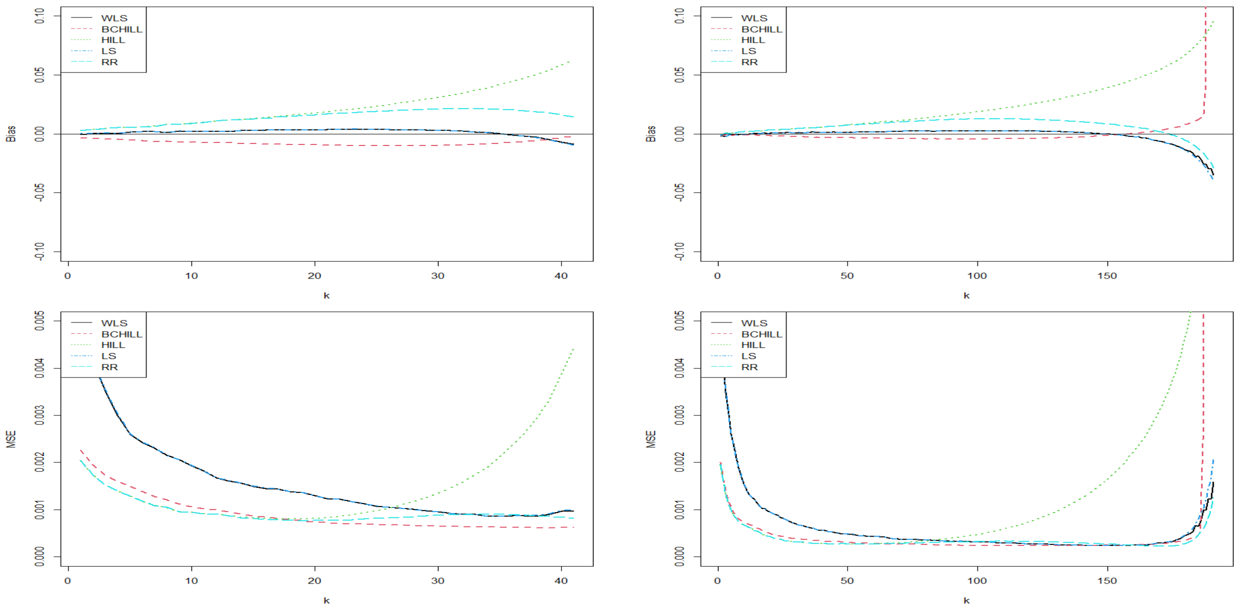

Figure 3.10: $\gamma = 0.1$

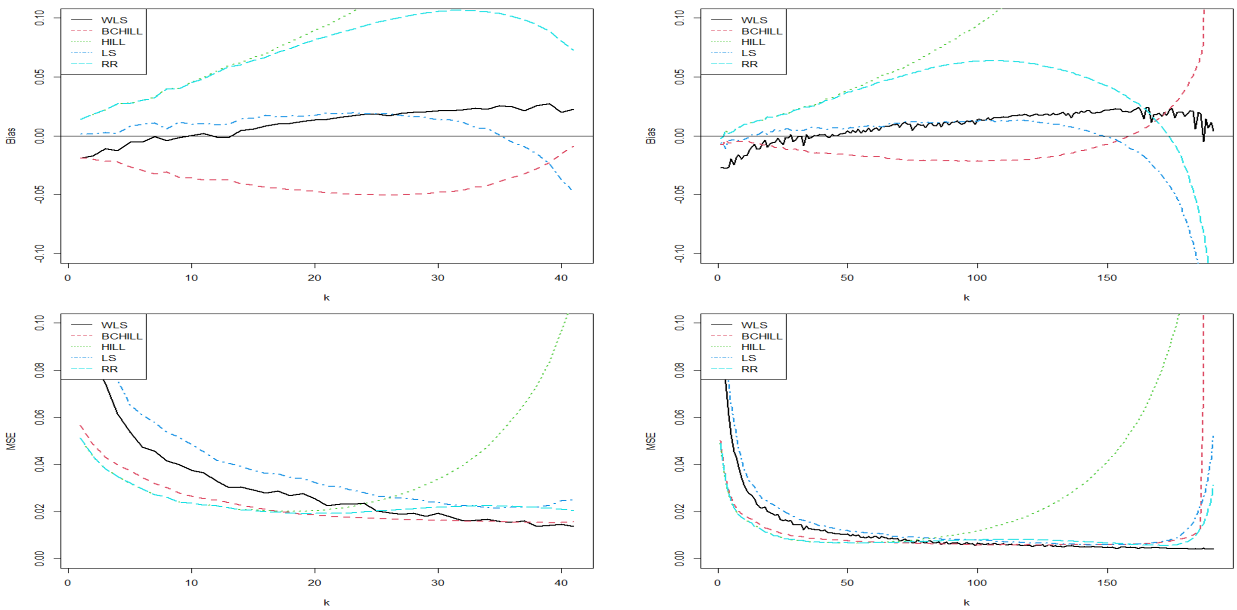

Figure 3.11: $\gamma = 0.5$

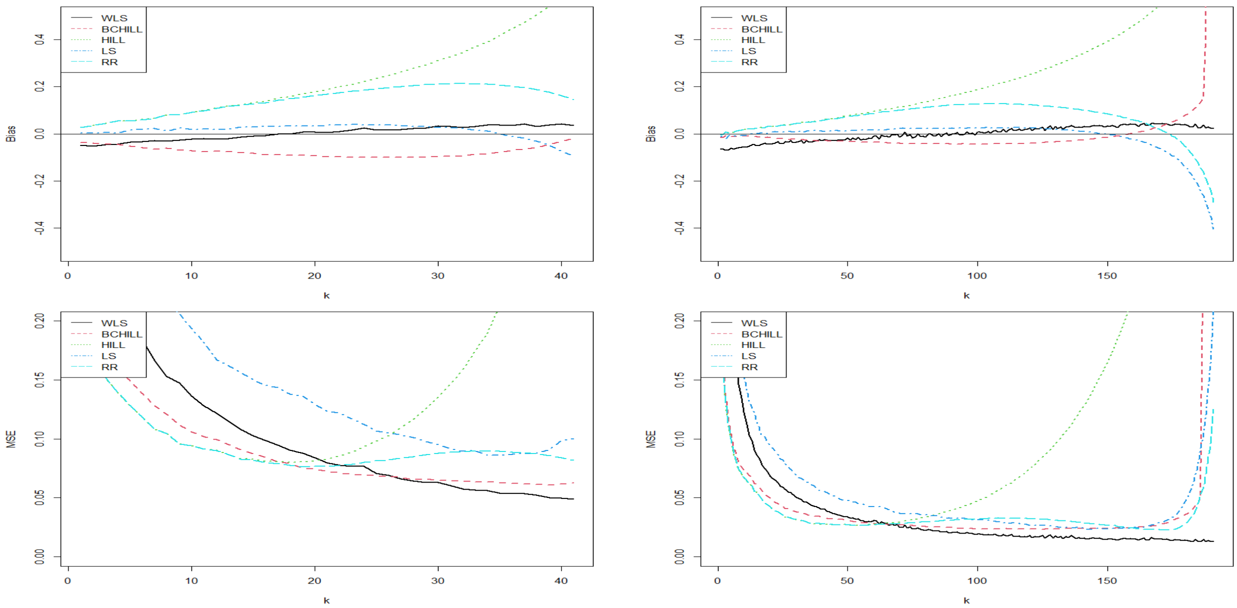

Figure 3.12: $\gamma = 1.0$

From Figure 3.10 − 3.12, It can be observed that WLS performs favourably well in terms of MSE and bias.This is almost horizontal The WLS, BCHILL and LS mostly have the same sample path for small to central values of k in terms of bias and in terms of MSE when $m = 200$. The WLS happens to be horizontal when $m = 200$ at medium to large values of k.

4. Application

In this section, the EVI estimators employed in the simulation studies are applied to real-life data sets to further demonstrate the advantages of our proposed estimators. In this practical illustration, we employ the Condroz data and the Secura Belgian reinsurance claim datasets.

- The Condroz data set contains calcium content $(mg/100g)$ of dry soil in the Condroz region located in Belgium with a data set of 1505 observations. The Condroz have been extensively studied by Beirlant et al. [5] and have been shown to belong to the Pareto domain of attraction.
- The Secura Belgian reinsurance claim data set contains 371 observations of automobile claims (in euros) that occurred from 1988 until 2001. Beirlant et al. [5] among other authors, have investigated this data set extensively and have shown that it has a heavy tail; thus, it is a member of the Pareto domain of attraction.

Figure 4.1 present the histogram of the two data sets. It can be observed from the respective histograms that the data sets are each skewed to the right. Hence, the data sets contain extremes.

The results of the EVI estimator employed in the simulation studies applied to the Secura reinsurance claim data are shown in Figure 4.2. The horizontal line through the graph represent median EVI estimate from our proposed estimators. The median value is 0.26. It can be observed that the HILL estimator is highly sensitive to k specifically at $k \geq 80$, and deviates from the linearity as k increases. The RR estimator shows stability for k values between 100 and 210. The BCHILL and LS show some moderate stability of the values of k. The WLS estimator displays stability in a wide range of values of k, for $k \geq 125$, which enables easier specification of the values of the EVI estimated for the Secura Belgian reinsurance claim data with an approximately deterministic weight function.

It can also be observed that, the proposed estimate with approximate deterministic weight is stable along the horizontal line over a large region of k. The ridge regression estimator also yields some stable estimates in the center region of k. In the case of our proposed EVI estimator with random weight, the EVI estimates are not stable due to it sensitivity to k and they follow the same path as the Least squares estimator.

In the case of the Condroz data set, Figure 4.3 presents the results of the EVI estimates from the various estimators employed. The median EVI value for the estimates is 0.27, and this is indicated on the horizontal line. The HILL and BCHILL estimators experience moderate stability for values of k between 250 and 500 but are sensitive to the changes in k especially for $k \geq 500$. The RR estimator displays stability across a large range of values of k between 350 and 1200. For $k \geq 1200$, it can observed that the ridge regression and the least squares regression estimators are also stable along the horizontal line. The WLS estimators show stability at the values of k between 750 and 1250. It can be observed that, the proposed estimator appears to be stable along the horizontal line.

Comparatively, it appears, our proposed estimator with random weight performs well when the sample size is large.

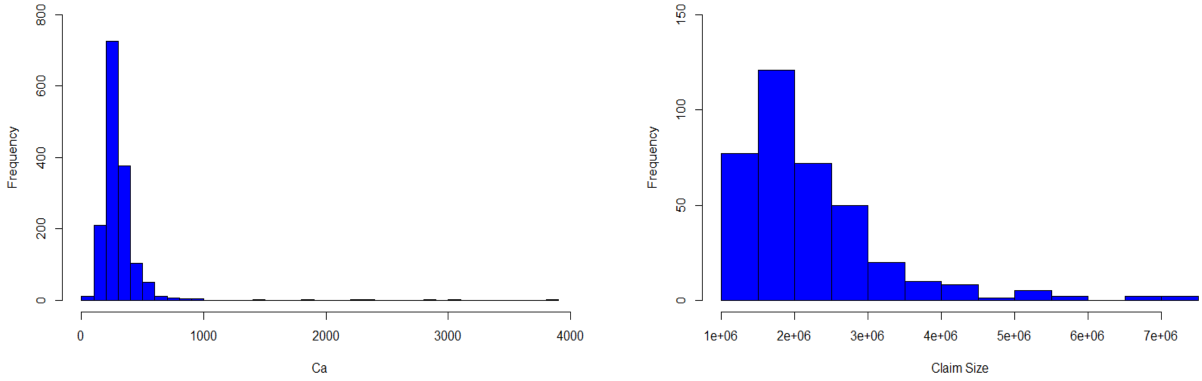

Figure 4.1: First column: Histogram of Condroz. Second column: Histogram of Secura Belgian reinsurance claims.

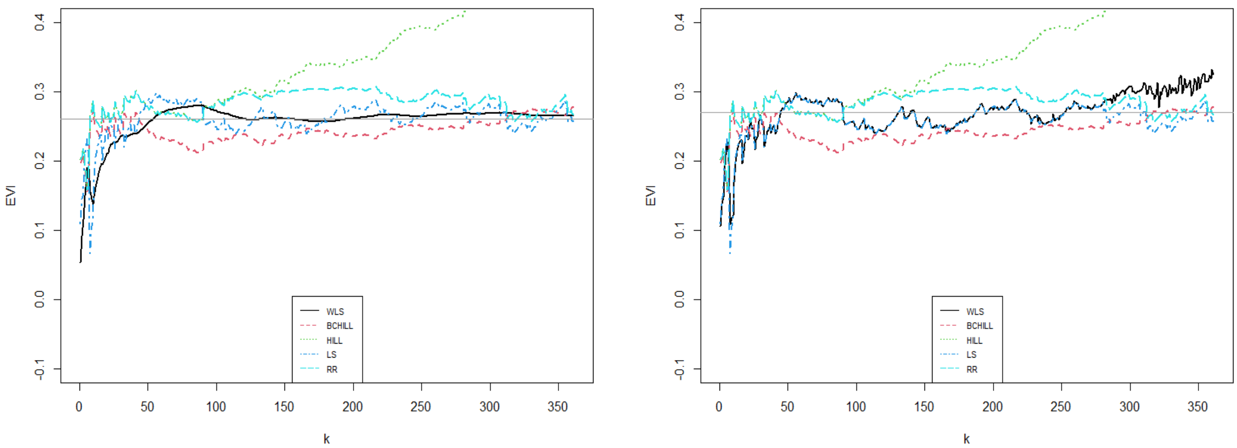

Figure 4.2: First column: Secura Belgian reinsurance claims with approximately deterministic weight. Second column: Secura Belgian reinsurance claims with a random weight.

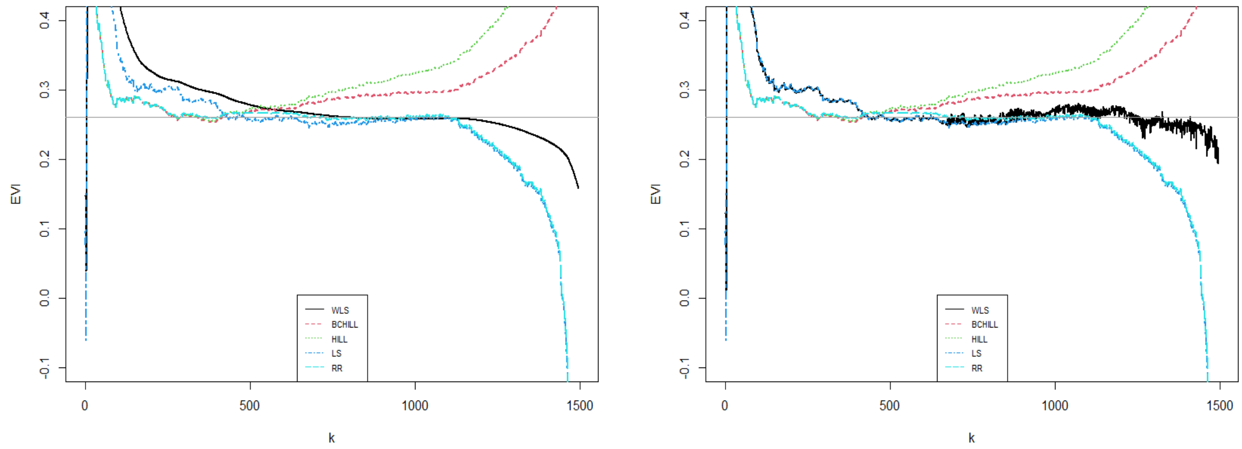

Figure 4.3: First column: Condroz with approximately deterministic weight. Second column: Condroz with a random weight.

4.1 Conclusion

In this essay, we proposed tail index estimators for the Pareto-type of distributions using the regression model. We proposed the weighted least squares estimator in two cases for the weight function: (i) when the weight function is approximately deterministic and (ii) a random weight function. In general, the WLS estimator generally yields lower bias and MSE for intermediate and large values of k, and therefore can be regarded as an appropriate estimator of the extreme value index for samples generated from the Pareto-type distribution. Again, the proposed estimator was found to be unbiased, consistent, and normal asymptotically.

Additionally, the WLS is the second-best performing estimator to the BCHILL estimator in terms of MSE for small sample size. However, the BCHILL, LS, HILL, and RR are highly sensitive to k for large sample size, but WLS is less sensitive as compared to the other estimators.

In most circumstances, the sample trajectories of the suggested estimators, the WLS and LS, have the same sample path for small values of k indicating that the two estimators are competitively close.

However, there's a fluctuation in the estimators with the random weight since the weight function depends on the cumulative distribution function, which is a step function. The fluctuations may be due to the discontinuity in the cumulative distribution function. The intensity of the fluctuations depends on the sample size for the random weight. Blanke and Bosq in [7] presented a polygonal smoothing and natural method for regularizing the empirical distribution function, which will help in smoothing the random weight.

References

[1] E. Afuecheta, C. Utazi, E. Ranganai, and C. Nnanatu. An application of Extreme value theory for Measuring Financial Fisk in Brics Economies. *Annals of Data Science*, pages 1–40, 2020.

[2] S. Alfarano and T. Lux. Extreme value theory as a theoretical background for power law behavior. 2010.

[3] J. Beirlant, G. Dierckx, Y. Goegebeur, and G. Matthys. Tail index estimation and an exponential regression model. *Extremes*, 2(2):177–200, 1999.

[4] J. Beirlant, G. Dierckx, A. Guillou, and C. Staăricaă. On exponential representations of log-spacings of extreme order statistics. *Extremes*, 5(2):157–180, 2002.

[5] J. Beirlant, Y. Goegebeur, J. Segers, and J. L. Teugels. *Statistics of extremes: theory and applications*, volume 558. John Wiley & Sons, 2004.

[6] J. Beran, D. Schell, and M. Stehlík. The harmonic moment tail index estimator: asymptotic distribution and robustness. *Annals of the Institute of Statistical Mathematics*, 66(1):193–220, 2014.

[7] D. Blanke and D. Bosq. Polygonal smoothing of the empirical distribution function. *Statistical Inference for Stochastic Processes*, 21(2):263–287, 2018.

[8] V. Brazauskas and R. Serfling. Robust and efficient estimation of the tail index of a single-parameter Pareto distribution. *North American Actuarial Journal*, 4(4):12–27, 2000.

[9] M. F. Brilhante, M. I. Gomes, and D. Pestana. A simple generalisation of the Hill estimator. *Computational Statistics & Data Analysis*, 57(1):518–535, 2013.

[10] S. Buitendag, J. Beirlant, and T. De Wet. Ridge regression estimators for the extreme value index. *Extremes*, 22(2):271–292, 2019.

[11] F. Caeiro, M. I. Gomes, and D. Pestana. Direct reduction of bias of the classical hill estimator. *Revstat*, 3(2):113–136, 2005.

[12] S. Csorgo, P. Deheuvels, and D. Mason. Kernel estimates of the tail index of a distribution. *The Annals of Statistics*, pages 1050–1077, 1985.

[13] J. Danielsson, D. W. Jansen, and C. G. De vries. The method of moments ratio estimator for the tail shape parameter. *Communications in Statistics-Theory and Methods*, 25(4): 711–720, 1996.

[14] 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*, pages 1833–1855, 1989.

[15] B. Finkenstadt and H. Rootzén. *Extreme values in finance, telecommunications, and the environment*. CRC Press, 2003.

[16] K. Gkillas and P. Katsiampa. An application of extreme value theory to cryptocurrencies. *Economics Letters*, 164:109–111, 2018.

[17] M. I. Gomesa and M. J. Martins. Asymptotically unbiased" estimators of the tail index based on external estimation of the second order parameter. *Extremes*, 5(1):5–31, 2002.

[18] B. M. Hill. A simple general approach to inference about the tail of a distribution. *The Annals of Statistics*, pages 1163–1174, 1975.

[19] M Ivette G., L. de Haan, and year=2008 Rodrigues, L. Tail index estimation for heavy-tailed models: accommodation of bias in weighted log-excesses series B statistical methodology.

[20] 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*, 2021, 2021.

[21] F. Longin. *Extreme events in finance: A handbook of extreme value theory and its applications*. John Wiley & Sons, 2016.

[22] G. Magnou. An application of extreme value theory for measuring financial risk in the Uruguayan Pension Fund. *Compendium: Cuadernos de Economía y Administración*, 4(7): 1–19, 2017.

[23] A. J. McNeil. Estimating the tails of loss severity distributions using extreme value theory. *ASTIN Bulletin: The Journal of the IAA*, 27(1):117–137, 1997.

[24] N. Mehrnia and S Coleri. Wireless channel modeling based on extreme value theory for ultra-reliable communications. *IEEE Transactions on Wireless Communications*, 2021.

[25] R. Minkah. Tail index estimation of the generalised Pareto distribution using a pivot from a transformed Pareto distribution. *Science and Development Journal*, 4(1):19–19, 2020.

[26] R. Minkah, D. E. Tertius, and K. DOKU-AMPONSAH. On extreme value index estimation under random censoring. *African Journal of Applied Statistics*, 5(2):419–445, 2018.

[27] R. Minkah, T. de Wet, and A. Ghosh. Robust estimation of Pareto-type tail index through an exponential regression model. *Communications in Statistics-Theory and Methods*, pages 1–19, 2021.

[28] E. Ocran, R. Minkah, and K. Doku-Amponsah. A Reduced-Bias Weighted least square estimation of the Extreme Value Index. *arXiv preprint arXiv:2110.08570*, 2021.

[29] E. Ocran, R. Minkah, G. Kallah-Dagadu, and K. Doku-Amponsah. Estimation of the tail index of Pareto-type distributions using regularisation. *arXiv preprint arXiv:2202.03319*, 2022.

[30] 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. *International Journal of Statistics and Probability*, 8(4):85–94, 2019.

[31] V. Paulauskas and M. Vaičiulis. On an improvement of Hill and some other estimators. *Lithuanian Mathematical Journal*, 53(3):336–355, 2013.

[32] V. Paulauskas and M. Vaičiulis. A class of new tail index estimators. *Annals of the Institute of Statistical Mathematics*, 69(2):461–487, 2017.

[33] S. Resnick and C. Stărică. Smoothing the Hill estimator. *Advances in Applied Probability*, 29(1):271–293, 1997.

[34] C. Rohrbeck, E. F. Eastoe, A. Frigessi, and J. A. Tawn. Extreme value modelling of water-related insurance claims. *The Annals of Applied Statistics*, 12(1):246–282, 2018.

[35] Y. M. Tripathi, S. Kumar, and C. Petropoulos. Improved estimators for parameters of a Pareto distribution with a restricted scale. *Statistical Methodology*, 18:1–13, 2014.

[36] J. B. P. Vynckier and J. L. Teugels. Tail index estimation, Pareto quantile plots, and regression diagnostics J. *Journal of the American Statistical Association*, 91:1659–1667, 1996.

[37] C. Wang and G. Chen. A new hybrid estimation method for the generalized Pareto distribution. *Communications in Statistics-Theory and Methods*, 45(14):4285–4294, 2016.

4.2 APPENDICES

Proof.

$$
\begin{aligned}
\sum_{j=1}^{k} \ell_j &= \sum_{j=1}^{k} \left(1 - U^{\theta(k)} \frac{j}{k+1}\right) && (4.2.1)\\
&= k - k \sum_{j=1}^{k} U^{\theta(k)} \frac{j}{k(k+1)} && (4.2.2)\\
&= k \left[\int_1^0 \left(1 - \mathbb{E}\left(U^{\theta(k)}\right)\right) du\right] \qquad \text{a.s} && (4.2.3)\\
&= k \left[\int_1^0 \left(1 - \left(\int_0^1 U^{\theta(k)} d\theta\right) u\right) du\right] \qquad \text{a.s} && (4.2.4)\\
&= k \left[\int_1^0 \left(1 - \frac{U^{\theta(k)+1}}{\theta(k)+1}\Bigg|_0^1 u\right) du\right] \qquad \text{a.s} && (4.2.5)\\
&= k \int_1^0 \left(1 - \frac{u}{\theta(k)+1}\right) du \qquad \text{a.s} && (4.2.6)\\
&= k \left[u - \frac{u^2}{2\theta(k)+2}\right]_0^1 \qquad \text{a.s} && (4.2.7)\\
&= k \left(1 - \frac{1}{2\theta(k)+2}\right) \qquad \text{a.s} && (4.2.8)\\
&= k \left(\frac{2\theta(k)+2-1}{2\theta(k)+2}\right) \qquad \text{a.s} && (4.2.9)\\
&= \frac{k}{2} \left(\frac{2\theta(k)+1}{\theta(k)+1}\right) \qquad \text{a.s} && (4.2.10)
\end{aligned}
$$

Therefore from Equation (4.2.1), we have

$$
\begin{aligned}
\overline{\ell_j} &\approxeq \frac{2}{k} \left(\frac{\theta(k)+1}{2\theta(k)+1}\right) \ell_j && (4.2.11)\\
\overline{\ell_j} &\approxeq \frac{2}{k} \pi(k) \ell_j \qquad \text{a.s} && (4.2.12)\\
\overline{\ell_j} &\approxeq \frac{2}{k} \pi(k) \quad k \to \infty \qquad \text{a.s} && (4.2.13)\\
\overline{\ell_j} &\approxeq \frac{1}{k} \qquad \text{a.s} && (4.2.14)\\
& && (4.2.15)
\end{aligned}
$$

where $\pi(k) = \left(\frac{\theta(k)+1}{2\theta(k)+1}\right)$, $-\mathcal{O}(-1) \leq F(x) \leq \mathcal{O}(1)$ and $\left(1 - \mathcal{O}(1) \frac{j}{k+1}\right) \leq \ell_j \leq \left(1 + \mathcal{O}(1) \frac{j}{k+1}\right)$ □

Proof of lemma 2.1.1

Proof. i. From Equation (2.0.19) and Equation (2.0.3), we have,

$$S_1(\rho,\theta) = \sum_{j=1}^{k} \bar{\ell}_j \delta_j \tag{4.2.16}$$

$$= \sum_{j=1}^{k} \frac{2}{k}\left(\frac{\theta(k)+1}{2\theta(k)+1}\right)\left(1 - U_j^{\theta(k)}\frac{j}{k+1}\right)\left(\frac{j}{k+1}\right)^{-\rho} \tag{4.2.17}$$

$$= \frac{2}{k}\left(\frac{\theta(k)+1}{2\theta(k)+1}\right)\sum_{j=1}^{k}\left(1 - U_j^{\theta(k)}\frac{j}{k+1}\right)\left(\frac{j}{k+1}\right)^{-\rho} \tag{4.2.18}$$

$$= \frac{2}{k}\left(\frac{\theta(k)+1}{2\theta(k)+1}\right)\left(k - \sum_{j=1}^{k}\mathbb{E}\left(U_j^{\theta(k)}\right)\frac{j}{k+1}\right)\left(\frac{j}{k+1}\right)^{-\rho} \quad \text{a.s} \tag{4.2.19}$$

$$= \frac{2}{k}\left(\frac{\theta(k)+1}{2\theta(k)+1}\right)\left(k - \sum_{j=1}^{k}\left(\int_0^1 U_j^{\theta(k)}dU\right)\frac{j}{k+1}\right)\left(\frac{j}{k+1}\right)^{-\rho} \quad \text{a.s} \tag{4.2.20}$$

$$= \frac{2}{k}\left(\frac{\theta(k)+1}{2\theta(k)+1}\right)\left(k - \sum_{j=1}^{k}\frac{1}{\theta(k)+1}\frac{j}{k+1}\right)\left(\frac{j}{k+1}\right)^{-\rho} \quad \text{a.s} \tag{4.2.21}$$

$$= 2\left(\frac{\theta(k)+1}{2\theta(k)+1}\right)\int_0^1\left(1 - \frac{u}{\theta(k)+1}\right)u^{-\rho}du + o(1) \quad \text{a.s} \tag{4.2.22}$$

$$= 2\left(\frac{\theta(k)+1}{2\theta(k)+1}\right)\left[\frac{u^{1-\rho}}{1-\rho} - \frac{u^{2-\rho}}{(2-\rho)(\theta(k)+1)}\right]_0^1 + o(1) \quad \text{a.s} \tag{4.2.23}$$

$$= \lim_{k\to\infty} 2\left(\frac{\frac{\theta(k)}{k}+\frac{1}{k}}{\frac{2\theta(k)}{k}+\frac{1}{k}}\right)\left(\frac{1}{1-\rho} - \frac{\frac{1}{k}}{(2-\rho)(\frac{\theta(k)}{k}+\frac{1}{k})}\right) + o(1) \quad \text{a.s} \tag{4.2.24}$$

$$S_1(\rho,\theta) = 2\left(\frac{\Delta}{2\Delta}\right)\left(\frac{1}{1-\rho}\right) = \frac{1}{1-\rho} \quad \text{a.s} \tag{4.2.25}$$

$$\tag{4.2.26}$$

where $\Delta = \lim_{k\to\infty}\frac{\theta(k)}{k}$ and therefore, as $k\to\infty,$ we have
$S_1(\theta) \xrightarrow{\text{a.s}} \frac{1}{(1-\rho)}$

ii.

$$S_2(\rho,\theta) = \sum_{j=1}^{k} \overline{\ell}_j \delta_j^2 - \left(\sum_{j=1}^{k} \overline{\ell}_j \delta_j \right)^2, \text{ from Equation (4.2.25), it follows that} \tag{4.2.27}$$

$$\tag{4.2.28}$$

$$S_2(\rho,\theta) = \sum_{j=1}^{k} \overline{\ell}_j \delta_j^2 - \left(\frac{1}{1-\rho} \right)^2 \tag{4.2.29}$$

Simplifying $\sum_{j=1}^{k} \overline{\ell}_j \delta_j^2$

$$\sum_{j=1}^{k}\bar{\ell}_j\delta_j^2 = \frac{2}{k}\left(\frac{\theta(k)+1}{2\theta(k)+1}\right)\sum_{j=1}^{k}\left(1-U_j^{\theta(k)}\frac{j}{k+1}\right)\left(\frac{j}{k+1}\right)^{-2\rho} \tag{4.2.30}$$

$$= \frac{2}{k}\left(\frac{\theta(k)+1}{2\theta(k)+1}\right)\left(k-\sum_{j=1}^{k}U_j^{\theta(k)}\frac{j}{k+1}\right)\left(\frac{j}{k+1}\right)^{-2\rho} \tag{4.2.31}$$

$$= \frac{2}{k}\left(\frac{\theta(k)+1}{2\theta(k)+1}\right)\left(k-\sum_{j=1}^{k}\mathbb{E}\left(U_j^{\theta(k)}\right)\frac{j}{k+1}\right)\left(\frac{j}{k+1}\right)^{-2\rho}+o(1) \qquad \text{a.s} \tag{4.2.32}$$

$$= \frac{2}{k}\left(\frac{\theta(k)+1}{2\theta(k)+1}\right)\left(k-\sum_{j=1}^{k}\left(\int_0^1 U_j^{\theta(k)}dU\right)\frac{j}{k+1}\right)\left(\frac{j}{k+1}\right)^{-2\rho}+o(1) \qquad \text{a.s} \tag{4.2.33}$$

$$= \frac{2}{k}\left(\frac{\theta(k)+1}{2\theta(k)+1}\right)\left(k-\sum_{j=1}^{k}\frac{1}{\theta(k)+1}\frac{j}{k+1}\right)\left(\frac{j}{k+1}\right)^{-2\rho}+o(1) \qquad \text{a.s} \tag{4.2.34}$$

$$= 2\left(\frac{\theta(k)+1}{2\theta(k)+1}\right)\int_0^1\cdot\left(1-\frac{u}{\theta(k)+1}\right)u^{-2\rho}du+o(1) \qquad \text{a.s} \tag{4.2.35}$$

$$= 2\left(\frac{\theta(k)+1}{2\theta(k)+1}\right)\int_0^1\left(u^{-2\rho}-\frac{u^{1-2\rho}}{\theta(k)+1}\right)du+o(1) \qquad \text{a.s} \tag{4.2.36}$$

$$= 2\left(\frac{\theta(k)+1}{2\theta(k)+1}\right)\left[\frac{u^{1-2p}}{1-2p}-\frac{u^{2-2p}}{(2-2\rho)(\theta(k)+1)}\right]_0^1+o(1) \qquad \text{a.s} \tag{4.2.37}$$

$$= 2\left(\frac{2(k)+1}{2\theta(k)+1}\right)\left(\frac{1}{1-2p}-\frac{1}{(2-2\rho)(\theta(k)+1)}\right)+o(1) \qquad \text{a.s} \tag{4.2.38}$$

$$= \lim_{k\to\infty}2\left(\frac{\frac{\theta(k)}{k}+\frac{1}{k}}{\frac{2\theta(k)}{k}+\frac{1}{k}}\right)\left(\frac{1}{1-2\rho}-\frac{\frac{1}{k}}{(2-2\rho)(\frac{\theta(k)}{k}+\frac{1}{k})}\right)+o(1) \qquad \text{a.s} \tag{4.2.39}$$

(4.2.40)

$$= 2\left(\frac{\Delta}{2\Delta}\right)\left(\frac{1}{1-2p}\right) = \frac{1}{1-2\rho} \qquad \text{a.s} \tag{4.2.41}$$

Now (4.2.42)

$$S_2(p,\theta) = \frac{1}{(1-2\rho)} - \frac{1}{(1-\rho)^2} \qquad \text{a.s, from Equation (4.2.29)} \tag{4.2.43}$$

$$= \frac{1-2\rho+\rho^2-1+2\rho}{(1-2\rho)(1-\rho)^2} = \frac{\rho^2}{(1-2\rho)(1-\rho)^2} \qquad \text{a.s} \tag{4.2.44}$$

(4.2.45)

where $\Delta = \lim_{k\to\infty} \frac{\theta(k)}{k}$ and therefore, as $k \to \infty$, we have, $S_2(\rho,\theta) \xrightarrow{\text{a.s}} \frac{\rho^2}{(1-2\rho)(1-\rho)^2}$

□