EUROPEAN JOURNAL OF PURE AND APPLIED MATHEMATICS
Vol. 16, No. 4, 2023, 2509-2543
ISSN 1307-5543 – ejpam.com
Published by New York Business Global

Kernel estimation of the Quintile Share Ratio index of inequality for heavy-tailed income distributions

Modou Kebe[1], El Hadji Deme[1,*], Tchilabalo Abozou Kpanzou[2], Solym Mawaki Manou-Abi[3], Ebrima Sisawo[1]

[1] *LERSTAD, UFR SAT, Université Gaston Berger, BP 234, Saint-Louis, Sénégal*
[2] *LaMMASD, Département de Mathématiques, Université de Kara, Togo*
[3] *Institut Montpelliérain Alexander Grothendieck,UMR CNRS 5149, Place Eugéne Bataillon, 34090,, Montpellier, France. Centre Universitaire de Formation et de Recherche, Mayotte, France*

Abstract. Evidence from micro-data shows that capital incomes are exceedingly volatile, which makes up a disproportionately high contribution to the overall inequality in populations with the heavy-tailed nature on the income distributions for many countries. The quintile share ratio (QSR) is a recently introduced measure of income inequality, also forming part of the European Laeken indicators and which cover four important dimensions of social inclusion (*health, education, employment* and *financial poverty*). In 2001, the European Council decided that income inequality in the European Union member states should be described using a number of indicators including the QSR. Non-parametric estimation has been developed on the QSR index for heavy-tailed capital incomes distributions. However, this method of estimation does not give satisfactory statistical performances, since it suffers badly from under coverage, and so we cannot rely on the non-parametric estimator. Hence, we need another estimator in the case of heavy-tailed populations. This is the reason why we introduce, in this paper, a class of semi-parametric estimators of the QSR index of economic inequality for heavy-tailed income distributions. Our methodology is based on the extreme value theory, which offers adequate statistical results for such distributions. We establish their asymptotic distribution, and through a simulation study, we illustrate their behavior in terms of the absolute bias and the median squared error. The simulation results clearly show that our estimators work well.

2020 Mathematics Subject Classifications: 62E20, 62G30, 62G32

Key Words and Phrases: Inequality Measures, QSR index, Kernel estimation, Heavy-tailed, Extreme value Theory.

*Corresponding author.
DOI: https://doi.org/10.29020/nybg.ejpam.v16i4.4765

Email addresses: modoukebe44@gmail.com (M. Kebe), elhadji.deme@egb.edu.sn (E. Deme) , kpanzout@gmail.com (T. A. Kpanzou), solym.manou-abi@univ-mayotte.fr (S. M.Manou-Abi), esso870@gmail.com (E. Sisawo)

1. Introduction

Inequality measurement is an attempt to make sense of comparisons of income distributions in terms of criteria that can be derived from ethical principles, appealing mathematical constructions or simple intuition. A serious approach to inequality measurement should begin with a consideration of the entities to which the tools of distributional judgement are applied. In the recent decades, the capital income is among the major incomes in a big number of countries, there are many important studies devoted to the capital income, mainly dealing with relationships between capital income taxation and welfare benefits, and in particular with the incidence and efficiency effects of taxes on incomes from capital in various scenarios of development sectors. For further details see, e.g., [1], [6], [20], [25], [28], [40].

Measuring and analyzing incomes, risks and other random outcomes has been an active and fruitful research field. Academics and governmental researchers have been developing measures that would aid them in understanding income and loss distributions, their differences with respect to geographic regions and changes over time periods. It is a fascinating area due to a number of reasons, one of them being the fact that different measures or indices are needed to reveal different features of capital income distributions. The Gini index has been widely used by economists and sociologists to measure economic inequality. It has been also studied extensively and its properties documented in a number of papers (see e.g. the survey papers [31], [43], ...). Also, [44] suggested the Zenga index of inequality measure, which aggregates the ratios of lower and upper conditional tail expectations, and by so doing takes into account the relative nature of the poor and the rich for a given population. It has been also explored from various points of view, see, eg. [22], [23], [24] and [25]. Another trend, somewhat different from Zenga's but equally interesting, which is based on the Palma index, is considered in, e.g. [7].

The quintile share ratio (QSR) is a recently introduced measure of income inequality, also forming part of the European Laeken indicators which cover four important dimensions of social inclusion (Financial poverty, employment, health and education). In 2001, the European Council decided that income inequality in the *European Union (EU)* member states should be described using a number of indicators including the Quintile share ratio index. Compared to the Gini index, relatively little research is available on the statistical inference of the QSR index. [32] investigated the QSR and established its variance in a complex sampling design framework. The authors upgraded on earlier work by [35] and [36]. As is to be expected from its definition, the influence function of the QSR is unbounded. The form of that influence function has been also derived by [30]. Also, [29] derived the asymptotic distribution of a non-parametric plug-in estimator for the QSR index. However, this estimation suffers badly from under coverage, and so we cannot rely on the non-parametric estimator. Hence, we need another estimator in the case of heavy-tailed populations. This is the reason why we introduce, in this paper, a class of semi-parametric estimators of the QSR index of inequality measure for heavy-tailed income distributions. Our consideration is based on the extreme value methodology, which offers adequate statistical tools for such distributions.

This paper is organised as follows. In Section 2, we shall recall the definition of the QSR index by presenting it in terms of upper and lower integrals. Also, we shall briefly consider a non-parametric estimator of the QSR index which is obtained by replacing the underlying cumulative distribution function (cdf) of the population by its empirical counterpart. As this estimator does not exhibit, satisfactory performance in heavy-tailed populations, we introduce, in Section 3, a class of semi-parametric estimators of the QSR index of economic inequality for heavy-tailed income distributions. Through the extreme value methodology, we establish their asymptotic distribution, and derived their improved results in Section 4. Under a simulation study, we illustrate their behavior in terms of the absolute bias and the median squared error in Section 5.

2. Definitions and empirical estimation

Suppose that we have at our disposal a sample $(X_1, \cdots, X_n)$, $n > 1$ of independent and identically distributed in a population represented by a non-negative random variable $X \geq 0$, with capital income distribution $F(x) = \mathbb{P}(X \leq x)$ and finite mean $\mu_F := \mathbb{E}[X]$. We assume that F is continuous and strictly increasing. The (generalized) inverse $Q : [0, 1) \mapsto [0, \infty)$ of the income distribution F, known in the literature as the quantile function, is defined for all $s \in [0, 1)$ by the formula $Q(s) = \inf\{x,\ F(x) \geq s\}$.
Considering two different levels α and β, such that $0 < \alpha < \beta < 1$ as illustrated in [29], the QSR index at levels α and β of the capital income X denoted by $\eta(Q, \alpha, \beta)$, is the ratio of an upper integral $U(Q, \beta)$ to a lower integral $L(Q, \alpha)$. More precisely, we have:

$$\eta(Q, \alpha, \beta) := \frac{U(Q, \beta)}{L(Q, \alpha)} = \frac{\int_\beta^1 Q(s)ds}{\int_0^\alpha Q(s)ds}. \tag{1}$$

The QSR index is then given by: $\eta(Q, 0.2, 0.8)$. In what follows, we will consider the more general QSR index $\eta(Q, \alpha, \beta)$ and the results will follow directly from that.

The empirical estimator of the distribution F is defined by $F_n(x) = n^{-1} \sum_{i=1}^n \mathbb{I}_{\{X_i \leq x\}}$ and its corresponding empirical quantile function is expressed by $Q_n(s) = \inf\{x; F_n(x) \geq s\}$, where $\mathbb{I}_S$ is the indicator function of the set S. Denote by $X_{1,n} \leq ... \leq X_{n,n}$ the order statistics associated with the sample $(X_1, \ldots, X_n)$. Thus, $Q_n(s)$ is equal to the *i-th* order statistic $X_{i,n}$ for all $s \in ((i-1)/n, i/n]$, and for all $i = 1, ..., n$. For this, one natural candidate for the empirical estimator of $\eta(Q, \alpha, \beta)$ is obtained by replacing in (1) the true quantile $Q(\cdot)$ with the sample quantiles $Q_n(\cdot)$. We arrive at the following 'traditional' QSR index estimator (see, e.g., [29]):

$$\widehat{\eta}_n(\alpha, \beta) := \eta(Q_n, \alpha, \beta) = \frac{\int_\beta^1 Q_n(s)ds}{\int_0^\alpha Q_n(s)ds}. \tag{2}$$

Clearly, the estimator $\widehat{\eta}_n(\alpha,\beta)$ can be rewritten as:

$$\widehat{\eta}_n(\alpha,\beta) = \left\{ n^{-1} \sum_{j=[n\beta]+1}^{n} X_{j,n} \right\} \Big/ \left\{ n^{-1} \sum_{j=1}^{[n\alpha]} X_{j,n} \right\}, \tag{3}$$

where $[x]$ is the integer part of x. Note that the QSR estimator $\widehat{\eta}_n(\alpha,\beta)$ is the ratio of a U-statistic to a L-statistic. According to [29]), for a given capital income distribution F with finite variance, the following result holds:

$$\sqrt{n}\,\left(\widehat{\eta}_n(\alpha,\beta) - \eta(Q,\alpha,\beta)\right) \xrightarrow{d} \mathcal{N}\left(0, \sigma_\eta^2(\alpha,\beta)\right), \text{ as } n \to \infty,$$

where

$$\sigma_\eta^2(\alpha,\beta) := \sigma^2(\beta,1) + (\eta(Q,\alpha,\beta))^2\,\sigma^2(0,\alpha) - 2\eta(Q,\alpha,\beta)\int_0^\alpha s dQ(s)\int_\beta^1 (1-t)dQ(t),$$

with for $0 \le s < t \le 1$ and

$$\sigma^2(s,t) = \int_s^t\int_s^t (\min(u;v) - uv)dQ(u)dQ(v).$$

This result is violated when capital incomes have heavy-tailed distributions with infinite variance, since the asymptotic variance $\sigma_\eta^2(\alpha,\beta)$ is also infinite. For more detail; refer to [29]).

However, micro-data show that capital incomes account for a large part of disparity in populations. Furthermore, in some countries, capital incomes have been making up a disproportionately high contribution to the overall inequality (see, e.g., [17]). The present research has been motivated by the need for better understanding the distribution and inequality of capital incomes, which in many countries appear to be heavy-tailed (see, eg. [25]).

To this aim, we assume that the income distribution F is heavy-tailed. This is equivalent to the fact that the survival function $\overline{F} := 1 - F$ associated to F is regularly varying at infinity with index $-1/\gamma < 0$. More precisely,

$$\overline{F}(x) = x^{-1/\gamma}\ell_F(x),\ x > 0, \tag{4}$$

where ℓ_F is a slowly varying function at infinity, *i.e* for all $x > 0$, $\ell_F(tx)/\ell_F(t) \to 1$, as $t \to \infty$. The relation (4) is also equivalent to $Q(1-s) = s^{-\gamma}\ell_Q(s),\ s \in (0,1)$, where $\ell_Q(zs)/\ell_Q(s) \to 1$, as $s \to 0$, for all $z \in (0,1)$. From (4), one can easily see that for all $x > 0$ and $z \in (0,1)$:

$$\lim_{t\to\infty}\frac{\overline{F}(tx)}{\overline{F}(t)} = x^{-1/\gamma} \quad \text{and} \quad \lim_{s\to 0}\frac{Q(1-zs)}{Q(1-s)} = z^{-\gamma}. \tag{5}$$

The relation in (5) is namely called the first order regularly varying condition. The parameter γ is called the tail index (or the extreme value index) and governs the tail behavior, with larger values indicating heavier tails. Its estimation has received a great attention in the extreme value literature (see, e.g., [10]). This kind of models and its unidentified parameters have been previously used by various authors such as [15], [25], [29] to assess inequality measure of capital incomes.

Next, we also note that:

- When $\gamma > 1$, the QSR index $\eta(Q, \alpha, \beta)$ and thus its estimator $\widehat{\eta}_n(\alpha, \beta)$ are not defined.
- When $0 < \gamma \leq 1/2$ (the lower half of the unit interval), then $\mathbb{E}[X^{2+\epsilon}] < \infty$ for some $\epsilon > 0$, and so we can use the available asymptotically normal estimator $\widehat{\eta}_n(\alpha, \beta)$, (see, [29]).
- When $1/2 < \gamma < 1$ (The upper half of the unit interval), then the second moment is infinite, and so the asymptotic normality of the estimator $\widehat{\eta}_n(\alpha, \beta)$ is violated (see, [29]).

The last situation motivate the need of a specific estimator of the QSR index for heavy-tailed income distributions with infinite second moments. The class of heavy-tailed distributions (the so-called Pareto-type distributions) includes distributions such as Pareto, Burr, Student, Lévy-stable, and log-gamma which are known to be appropriate models in Extreme Value Theory for fitting large insurance claims, large fluctuations of prices, log-returns, incomes of countries with very high economic inequality, etc. (see, e.g., [2]; [3]; [4]; [11] ; [13]; [14]; [15]; [25]; [34]; [39]; [41]).

To better understand the heavy-tailed distribution and the inequality of capital incomes, which are governed by the unknown extreme value index γ, we make use, in this paper, of the extreme value methodology and propose asymptotically normal estimators of the QSR index $\eta(Q, \alpha, \beta)$. The following section concerns a class of semi-parametric estimators of the QSR index $\eta(Q, \alpha, \beta)$ for heavy-tailed income distributions with infinite second order moments.

3. Kernel estimation of the Quintile Share Ratio index

In the rest of this paper, we shall be concerned with heavy-tailed capital income distributions with index in the upper half of the unit interval. More precisely, we will deal with the case where F satisfies

$$\overline{F}(x) = x^{-1/\gamma}\ell_F(x), \ x > 0, \ 1/2 < \gamma < 1. \tag{6}$$

We have mentioned above that the tail index γ controls the behavior of income distribution F and its finite variance. In this spirit, we shall take into account the estimation of γ in the construction of our class of semi-parametric estimators for QSR index $\eta(Q,\alpha,\beta)$.

Now, Let $k = k(n)$ be an intermediate sequence of integers, i.e., a sequence such that:

$$1 < k < n, \quad k \to \infty \quad \text{and} \quad k/n \to 0 \quad \text{as} \quad n \to \infty. \tag{7}$$

Using the same arguments as in [34], the QSR index can be decomposed for $0 < \alpha < \beta < 1 - k/n$ as follows:

$$\eta(Q), \alpha, \beta) := \frac{1}{L(Q,\alpha)} \left\{ \int_{\beta}^{1-k/n} Q(s)ds + \int_{1-k/n}^{1} Q(s)ds \right\}. \tag{8}$$

As mentioned above, one can estimate the moderate quantile $Q(s)$, $\beta \leq s \leq 1 - k/n$ by its empirical estimator $Q_n(s)$. But the case where $1 - k/n < s < 1$ corresponds to high quantiles i.e $Q(s)$, $s \to 1$ and it is not possible to use the empirical estimation $Q_n(s)$. Under the first order regularly varying condition (5), we have $Q(1-zs) \approx z^{-\gamma} Q(1-s),\ s \to 0$. By setting $zs = 1 - u$ and $s = k/n$, we obtain the following approximation:

$$Q(u) \approx (n(1-u)/k)^{-\gamma} Q(1-k/n),\ u \to 1. \tag{9}$$

This leads to the following Weissman's type estimator ([42]) of high quantile $Q(u)$, $u \to 1$:

$$Q_{n,k}^{(K)}(u) = \left(\frac{n}{k}(1-u)\right)^{-\widehat{\gamma}_{n,k}^{(K)}} X_{n-k,n}, \tag{10}$$

with $\widehat{\gamma}_{n,k}^{K}$, the kernel class of estimators for the tail index γ, introduced in [9] and given by:

$$\widehat{\gamma}_{n,k}^{(K)} = \frac{1}{k} \sum_{j=1}^{k} jK\left(\frac{j}{k+1}\right) \log\left(\frac{X_{n-j+1,n}}{X_{n-j,n}}\right), \tag{11}$$

where K is a kernel integrating to one. Note that in the particular case where $K = \underline{K} := \mathbb{I}_{(0,1)}$, the estimator $\widehat{\gamma}_{n,k}^{\underline{K}}$ corresponds to the well-known Hill's estimator ([27]) of the tail index γ defined by:

$$\widehat{\gamma}_{n,k} := \widehat{\gamma}_{n,k}^{(\underline{K})} = \frac{1}{k} \sum_{j=1}^{k} j \log\left(\frac{X_{n-j+1,n}}{X_{n-j,n}}\right). \tag{12}$$

The estimator $\widehat{\gamma}_{n,k}$ is the most popular estimator of the tail index γ in the framework of heavy-tailed distributions. The Weissman's estimator ([42]) of high quantile is thus defined as $Q_{n,k}^{(\underline{K})}(u)$.

Next, replacing in (8), $Q(s)$ by its empirical quantile estimator $Q_n(s)$, for $\beta \leq s \leq 1-k/n$ and by its high quantiles estimator $Q_{n,k}^{(K)}(s)$, for $1-k/n < s < 1$, we arrive at the following kernel-type estimators of $\eta(Q,\alpha,\beta)$, $0 < \alpha < \beta < 1 - k/n$:

$$\widehat{\eta}_{n,k}^{(K)}(\alpha,\beta) = \frac{1}{L_n(\alpha)} \left\{ \int_{\beta}^{1-k/n} Q_n(s)ds + \int_{1-k/n}^{1} Q_{n,k}^{(K)}(s)ds \right\}, \tag{13}$$

where $L_n(\alpha) = n^{-1}\sum_{j=1}^{[n\alpha]} X_{j,n}$ is the above empirical estimator of the lower integral $L(\alpha)$, which can be rewritten as:

$$\widehat{\eta}_{n,k}^{(K)}(\alpha,\beta) = \frac{1}{L_n(\alpha)}\left\{\sum_{j=1}^{n-k}\left[\left(\frac{j}{n}-\beta\right)_+ - \left(\frac{j-1}{n}-\beta\right)_+\right]X_{j,n} + \frac{(k/n)}{1-\widehat{\gamma}_{n,k}^{(K)}}X_{n-k,n}\right\}, \quad (14)$$

where $(s-t)_+$ is the classical notation for the positive part of $(s-t)$. The estimator $\widehat{\eta}_{n,k}^{(K)}(\alpha,\beta)$ generalizes the one proposed in [15], when we use a general kernel instead of $\underline{K}$. In the above definitions and in what follows we indicate by $Q(\cdot)$ and $Q_n(\cdot)$ the quantile function and its empirical counterpart (both functions are left-continuous).

Let us now state an asymptotic normality result for $\widehat{\eta}_{n,k}^{(K)}(\alpha,\beta)$. As it exhibit a bias, we will introduce a bias reduction method to estimate the QSR index.

4. Main results

In extreme value analysis, one can easily achieve asymptotic normality results by imposing a second order regularly varying condition $(\mathcal{R}_{\mathbb{U}})$, (see, e.g., [10], Page 48), which is necessary to quantify the speed of convergence in (5). This condition can be formulated in different ways, below we state it in terms of the tail quantile functions $\mathbb{U}(x) = Q(1-1/x)$:
$(\mathcal{R}_{\mathbb{U}})$: *There exist a function $a(x) \to 0$ as $x \to \infty$ of constant sign for large values of x and a second order parameter $\rho < 0$ such that, for any $x > 0$,*

$$\lim_{t\to\infty}\frac{\log\mathbb{U}(tx) - \log\mathbb{U}(t) - \gamma\log x}{a(t)} = \frac{x^\rho - 1}{\rho}. \quad (15)$$

Note that the condition $(\mathcal{R}_{\mathbb{U}})$ implies that $|a|$ is regularly varying with index ρ (see, e.g., [10];[18]; [33]). As an example of heavy-tailed distributions satisfying the second order regularly varying condition $(\mathcal{R}_{\mathbb{U}})$, we have the so called and frequently used Hall's model which is a class of cdf's, such that $\mathbb{U}(t) = ct^\gamma(1 + d\rho^{-1}a(t) + o(t^\rho))$, as $t \to \infty$, where $\gamma > 0$, $\rho \le 0$, $c > 0$, and $d \in \mathbb{R}^*$. This sub-class of heavy-tailed models contains the distributions such as Pareto, Burr, Fréchet and Student-t. For statistical inference concerning the second-order parameter ρ, we refer, for example, to [12] and [38].

Section 4.1 below gives the asymptotic normality of our proposed estimator $\widehat{\eta}_{n,k}^{(K)}(\alpha,\beta)$.

4.1. Asymptotic normality of the kernel estimator $\widehat{\eta}_{n,k}^{(K)}(\alpha,\beta)$.

To establish the asymptotic normality of the kernel-type estimator $\widehat{\eta}_{n,k}^{(K)}(\alpha,\beta)$, some classical assumptions about the kernel K are needed.

Condition ($\mathbb{K}$). Let K be a function defined on $(0,1]$ such that
(i) $K(s) \geq 0$, whenever, $0 < s \leq 1$ and $K(1) = 0$; **(ii)** $K(\cdot)$ is differentiable, non increasing and right continuous on $(0,1]$; **(iii)** K and K' are bounded; **(iv)** $\int_0^1 K(u)du = 1$; **(v)** $\int_0^1 u^{-1/2}K(u)du < 1$.

These conditions are not restrictive but are satisfied by the usual weight functions used in the literature, including the power kernel $K(s) = (1+\tau)s^\tau \mathbb{I}_{\{0<s<1\}}$, $\tau \geq 0$, and the log-weight function $K(s) = (-\log s)^\tau / \Gamma(\tau+1)\mathbb{I}_{\{0<s<1\}}$, $\{\tau \geq 1\}$. In particular, we note that the classical Hill's estimator in (12) can be viewed as a particular case of our power kernel-type estimator corresponding to $\tau = 0$ and $K(s) := \underline{K}(s) = \mathbb{I}_{\{0<s<1\}}$.

Under the second order regularly varying assumption $(\mathcal{R}_{\mathbb{U}})$ and the condition $(\mathbb{K})$, [13] showed that

$$\widehat{\gamma}_{n,k}^{(K)} \overset{d}{=} \gamma + a(n/k)\int_0^1 s^{-\rho}K(s)ds + k^{-1/2}\ \xi_{n,k} + o_{\mathbb{P}}(k^{-1/2}), \tag{16}$$

where $\xi_{n,k}$ is asymptotically a centred normal distribution with variance $\gamma^2 \int_0^1 K^2(s)ds$. In this spirit, we establish in Theorem 1 below the asymptotic normality of the class of kernel type estimators $\widehat{\eta}_{n,k}^{(K)}(\alpha,\beta)$ for the QSR index.

Theorem 1. *Let K be a kernel satisfying $(\mathbb{K})$ and assume that the distribution F satisfies $(\mathcal{R}_{\mathbb{U}})$ with $\gamma \in (1/2,1)$. Then for any sequence of integers $k = k(n)$ satisfying $k \to \infty$, $k/n \to 0$ and $\sqrt{k}a(n/k) \to \lambda \in \mathbb{R}$ as $n \to \infty$, we have, for $0 < \alpha < \beta < 1 - k/n$,*

$$\frac{\sqrt{n}\ \left(\widehat{\eta}_{n,k}^{(K)}(\alpha,\beta) - \eta(Q,\alpha,\beta)\right)}{(k/n)^{1/2}X_{n-k,n}} \overset{d}{\to} \mathcal{N}\left(\lambda \mathcal{AB}_K^{(\eta)}(\gamma,\rho,\alpha), \mathcal{AV}_K^{(\eta)}(\gamma,\alpha)\right),$$

where

$$\mathcal{AB}_K^{(\eta)}(\gamma,\rho,\alpha) := \frac{1}{(1-\gamma)\,L(Q,\alpha)}\left(\frac{1}{\gamma+\rho-1} + \frac{1}{1-\gamma}\int_0^1 s^{-\rho}K(s)ds\right)$$

and

$$\mathcal{AV}_K^{(\eta)}(\gamma,\alpha) := \frac{\gamma^2}{(1-\gamma)^2\ L^2(Q,\alpha)}\left(\frac{1}{2\gamma-1} + \frac{1}{(1-\gamma)^2}\int_0^1 K^2(s)ds\right).$$

Proof of Theorem 1.
Let $Y_1,...,Y_n$ be independent and identically distributed random variables from the unit Pareto distribution G, defined as $G(y) = 1 - y^{-1}$, $y \geq 1$. For each $n \geq 1$, let $Y_{1,n} \leq ... \leq Y_{n,n}$ be the order statistics pertaining to $Y_1,...,Y_n$. Clearly $X_{j,n} \overset{d}{=} \mathbb{U}(Y_{j,n})$, $j = 1,...,n$. In order to use the results from [8], a probability space $(\Omega, \mathbb{A}, \mathbb{P})$ is constructed carrying a sequence $\xi_1, \xi_2, ...$ of independent random variables uniformly distributed on $(0,1)$ and a

sequence of Brownian bridges $\mathbb{B}_n(s)$, $0 \leq s \leq 1, n = 1, 2...$ such that for all $0 \leq \nu < 1/2$ and $\lambda_1 > 0$

$$\sup_{\lambda_1/n \leq s \leq 1-\lambda_1/n} \frac{|\beta_n(s) - \mathbb{B}_n(s)|}{(s(1-s))^{1/2-\nu}} = O_{\mathbb{P}}(n^{-\nu}), \tag{17}$$

where β_n is the resulting empirical quantile function denoted by:

$$\beta_n(t) = \sqrt{n}\ (t - \mathbb{V}_n(t))$$

with $\mathbb{V}_n(s) = \xi_{j,n}$, $\frac{j-1}{n} < s \leq \frac{j}{n}$, $j = 1, ..., n$ and $\mathbb{V}_n(0) = 0$.

Before we establish the asymptotic results in Theorem ‘1, let’s introduce the following notations. Next, from (8), (13), (14) and (42), the QSR index $\eta(Q, \alpha, \beta)$ and its biased estimator $\widehat{\eta}_{n,k}^{(K)}(\alpha, \beta)$, $0 < \alpha < \beta < 1 - k/n$, can be respectively rewritten as:

$$\eta(Q, \alpha, \beta) := \frac{U_{n,k,1}(Q, \beta)}{L(Q, \alpha)} + \frac{U_{n,k,2}(Q)}{L(Q, \alpha)},$$

and

$$\widehat{\eta}_{n,k}^{(K)}(\alpha, \beta) := \frac{U_{n,k,1}(Q_n, \beta)}{L_n(\alpha)} + \frac{U_{n,k,2}\left(Q_{n,k}^{(K)}\right)}{L_n(\alpha)},$$

where $L_n(\alpha) := \int_0^\alpha Q_n(s)ds = n^{-1}\sum_{j=1}^{[n\alpha]} X_{j,n}$ is the empirical estimator of the lower integral $L(Q, \alpha) = \int_0^\alpha Q(s)ds$ and the U-functional integrals are defined as:

$$\begin{aligned}
U_{n,k,1}(Q, t) &:= \int_t^{1-k/n} Q(s)ds, \quad \text{for} \quad 0 \leq t < 1 - k/n \\
U_{n,k,2}(Q) &:= \int_{1-k/n}^1 Q(s)ds, \\
U_{n,k,1}(Q_n, t) &:= \int_t^{1-k/n} Q_n(s)ds, \quad \text{for} \quad 0 \leq t < 1 - k/n, \\
U_{n,k,2}\left(Q_{n,k}^{(K)}\right) &:= \int_{1-k/n}^1 Q_{n,k}^{(K)}(s)ds = \frac{(k/n)X_{n-k,n}}{1 - \widehat{\gamma}_{n,k}^{(K)}}.
\end{aligned}$$

with $Q_n(\cdot)$ (respectively, $Q_{n,k}^{(K)}$ is the empirical estimator (respectively, the Weissman's type estimator) of the quantile function $Q(\cdot)$ and $k = k(n)$ is a sequence of integers satisfying $k \to \infty$, $k/n \to 0$ and as $n \to \infty$.

To simplify the proof, we need the following preliminary results whose proofs are given after this one.

Lemma 1. *Assume that the distribution F satisfies the regularly varying condition (5) with $\gamma \in (1/2, 1)$. If $k = k(n)$ is a sequence of integers satisfying $k \to \infty$, $k/n \to 0$, as*

$n \to \infty$, then for $0 < \alpha < 1 - k/n$, we have:

$$\frac{\sqrt{n}\Big(L_n(\alpha) - L(Q,\alpha)\Big)}{(k/n)^{1/2} X_{n-k,n}} \quad = \quad o_{\mathbb{P}}(1), \quad \text{as } n \to \infty. \tag{18}$$

Lemma 2. *Assume that the distribution F satisfies the regularly varying condition (5) with $\gamma \in (1/2, 1)$. If $k = k(n)$ is a sequence of integers satisfying $k \to \infty$, $k/n \to 0$, as $n \to \infty$, then for $0 < \alpha < \beta < 1 - k/n$, we have:*

$$\frac{\sqrt{n}\left\{\dfrac{U_{n,k,1}(Q_n,\beta)}{L_n(\alpha)} - \dfrac{U_{n,k,1}(Q,\beta)}{L(Q,\alpha)}\right\}}{(k/n)^{1/2} X_{n-k,n}} \overset{d}{=} W_{n,\alpha,1} + o_{\mathbb{P}}(1), \tag{19}$$

as $n \to \infty$, where

$$W_{n,\alpha,1} := -\frac{\int_0^{1-k/n} \mathbb{B}_n(s) dQ(s)}{L(Q,\alpha)\ (k/n)^{1/2} Q(1-k/n)}.$$

Lemma 3. *Under the assumptions of Theorem 1, we have for $0 < \alpha < 1 - k/n$:*

$$\frac{\sqrt{n}\left\{\dfrac{U_{n,k,2}\left(Q_{n,k}^{(K)}\right)}{L_n(\alpha)} - \dfrac{U_{n,k,2}(Q)}{L(Q,\alpha)}\right\}}{(k/n)^{1/2} X_{n-k,n}} \ \overset{d}{=}\ \lambda \mathcal{AB}_K^{(\eta)}(\gamma,\rho,\alpha) + W_{n,\alpha,2} + W_{n,\alpha,3} + o_{\mathbb{P}}(1), \tag{20}$$

as $n \to \infty$, where $\mathcal{AB}_K^{(\eta)}(\gamma,\rho,\alpha)$ is defined in Theorem 1 and

$$\begin{cases} W_{n,\alpha,2} := -\dfrac{\gamma}{(1-\gamma)\ L(Q,\alpha)} \sqrt{\dfrac{n}{k}}\, \mathbb{B}_n(1-k/n), \\ W_{n,\alpha,3} := \dfrac{\gamma}{(1-\gamma)^2\ L(Q,\alpha)} \sqrt{\dfrac{n}{k}} \displaystyle\int_0^1 s^{-1} \mathbb{B}_n(1-sk/n) d(sK(s)). \end{cases}$$

Now, coming back to the proof of the theorem, under assumptions, we have:

$$\begin{aligned} \widehat{\eta}_{n,k}^{(K)}(\alpha,\beta) - \eta(Q,\alpha,\beta) \quad &= \quad \left\{\frac{U_{n,k,1}(Q_n,\beta)}{L_n(\alpha)} - \frac{U_{n,k,1}(Q,\beta)}{L(Q,\alpha)}\right\} \\ &\quad + \left\{\frac{U_{n,k,2}\left(Q_{n,k}^{(K)}\right)}{L_n(\alpha)} - \frac{U_{n,k,2}(Q)}{L(Q,\alpha)}\right\} \\ &:= \quad A_{n,1} + A_{n,2}. \end{aligned} \tag{21}$$

For all values of n large enough, we get respectively from Lemma 2 and Lemma 3:

$$\frac{\sqrt{n}\ A_{n,1}}{(k/n)^{1/2} X_{n-k,n}} \overset{d}{=} W_{n,\alpha,1} + o_{\mathbb{P}}(1),$$

and

$$\frac{\sqrt{n}A_{n,2}}{(k/n)^{1/2}X_{n-k,n}} \stackrel{d}{=} \lambda\mathcal{A}\mathcal{B}_K^{(\eta)}(\gamma,\rho,\alpha) + W_{n,\alpha,2} + W_{n,\alpha,3} + o_{\mathbb{P}}(1).$$

This leads to

$$\frac{\sqrt{n}\,\left(\widehat{\eta}_{n,k}^{(K)}(\alpha,\beta) - \eta(Q,\alpha,\beta)\right)}{(k/n)^{1/2}X_{n-k,n}} \stackrel{d}{=} \lambda\mathcal{A}\mathcal{B}_K^{(\eta)}(\gamma,\rho,\alpha) + W_{n,\alpha,1} + W_{n,\alpha,2} + W_{n,\alpha,3} + o_{\mathbb{P}}(1). \tag{22}$$

Now, our next step is to compute the asymptotic variance of the process $W_{n,\alpha,1} + W_{n,\alpha,2} + W_{n,\alpha,3}$. The computations quite direct and we give below the main arguments, i.e.

$$\begin{aligned}
\mathbb{E}W_{n,\alpha,1}^2 &= \frac{\int_0^{1-k/n}(1-t)\left(\int_0^t s dQ(s)\right)dQ(t)}{L^2(Q,\alpha)\,k/n\,Q^2(1-k/n)} + \frac{\int_0^{1-k/n} t\left(\int_t^{1-k/n}(1-s)dQ(s)\right)dQ(t)}{L^2(Q,\alpha)\,k/n\,Q^2(1-k/n)} \\
&= \frac{\int_{k/n}^1 u\left(\int_u^1 dQ(1-v)\right)dQ(1-u)}{L^2(Q,\alpha)\,k/n\,Q^2(1-k/n)} - \frac{\int_{k/n}^1 u\left(\int_u^1 v dQ(1-v)\right)dQ(1-u)}{L^2(Q,\alpha)\,k/n\,Q^2(1-k/n)} \\
&+ \frac{\int_{k/n}^1 \left(\int_{k/n}^u v dQ(1-v)\right)dQ(1-u)}{L^2(Q,\alpha)\,k/n\,Q^2(1-k/n)} - \frac{\int_{k/n}^1 u\left(\int_{k/n}^u v dQ(1-v)\right)dQ(1-u)}{L^2(Q,\alpha)\,k/n\,Q^2(1-k/n)} \\
&=: Q_{n,\alpha,1} + Q_{n,\alpha,2} + Q_{n,\alpha,3} + Q_{n,\alpha,4}.
\end{aligned}$$

Recall now that $Q(1-s) = s^{-\gamma}\ell_Q(s)$ with ℓ_Q a slowly varying function at 0. By integration by parts and using Lemma 6 in [13]

$$Q_{n,\alpha,1} = \frac{1}{2\,L^2(Q,\alpha)}\left[1 + \frac{\int_{k/n}^1 Q^2(1-u)du}{k/nQ^2(1-k/n)}\right] \longrightarrow \frac{\gamma}{(2\gamma-1)\,L^2(Q,\alpha)}.$$

Remark that $d\left(\int_u^1 v dQ(1-v)\right) = -u\,dQ(1-u)$ which implies that

$$Q_{n,\alpha,2} = -\frac{1}{2\,L^2(Q,\alpha)}\,\frac{k}{n}\left[\frac{\int_{k/n}^1 v dQ(1-v)}{k/n\,Q(1-k/n)}\right]^2 = o(1) \tag{23}$$

this last result coming from the fact that, according to Proposition 1.3.6 in [5] for all $\varepsilon > 0$, $x^{-\varepsilon}\ell(x) \longrightarrow \infty$ as $x \to 0$. Thus, choosing $0 < \varepsilon < \gamma - \frac{1}{2}$ entails

$$\begin{aligned}
0 \le s\left(\frac{\int_s^1 t d(Q(1-t))}{sQ(1-s)}\right)^2 &= s\left(1 + \frac{\int_s^1 t^{-\gamma}\ell_Q(t)dt}{s^{1-\gamma}\ell_Q(s)}\right)^2 \\
&\le s\left(1 + Cs^{\gamma-1-\varepsilon}\right)^2 = O\left(s^{1+2[\gamma-1-\varepsilon]}\right) = o(1)
\end{aligned}$$

where C is a suitable constant. Consequently, $Q_{n,\alpha,2} \longrightarrow 0$. The two others terms, $Q_{n,\alpha,3}$ and $Q_{n,\alpha,4}$, can be treated similarly, leading to

$$Q_{n,\alpha,3} = Q_{n,\alpha,1} \quad \longrightarrow \quad \frac{\gamma}{(2\gamma-1)\,L^2(Q,\alpha)}$$

$$Q_{n,\alpha,4} = Q_{n,\alpha,2} \quad \longrightarrow \quad 0.$$

Finally,

$$\mathbb{E}W_{n,\alpha,1}^2 \longrightarrow \frac{2\gamma}{(2\gamma-1)\,L^2(Q,\alpha)},$$

and direct computations now lead to

$$\begin{aligned}
\mathbb{E}(W_{n,\alpha,2}^2) &\longrightarrow \frac{\gamma^2}{(1-\gamma)^2\,L^2(Q,\alpha)} \\
\mathbb{E}W_{n,\alpha,3}^2 &\longrightarrow \frac{\gamma^2}{(1-\gamma)^4\ L^2Q,\alpha)}\int_0^1 K^2(s)ds \quad \text{by Corollary 1 in [13]} \\
\mathbb{E}(W_{n,\alpha,1}W_{n,\alpha,2}) &\longrightarrow \frac{\gamma}{(1-\gamma)L^2(Q,\alpha)} \quad \text{by (23)}, \\
E(W_{n,\alpha,1}W_{n,\alpha,3}) &= 0 \qquad \text{and} \qquad E(W_{n,\alpha,2}W_{n,\alpha,3}) = 0.
\end{aligned}$$

Combining all these results, Theorem 1 follows. ■

Proof of Lemma 1. Let $t \in |0, 1-k/n)$, we have:

$$\frac{\sqrt{n}\Big(U_{n,k,1}(Q_n,t) - U_{n,k,1}(Q,t)\Big)}{(k/n)^{1/2}X_{n-k,n}} = \frac{\int_t^{1-k/n}\sqrt{n}\Big(Q_n(s)-Q(s)\Big)ds}{(k/n)^{1/2}X_{n-k,n}}.$$

Since $Q(1-\cdot)$ is a regularly varying function at zero with index $-\gamma$, then from Theorem 2.4.1 in [10], $X_{n-k,n} = Q_n(1-k/n) = Q(1-k/n)(1+o_{\mathbb{P}}(1))$, as $n \to \infty$. Using the approach in (17) and the Vervaat process (see [45]), [34] showed in Statement 4.3, p. 8, for all $t \in (0;1)$, that

$$\frac{\int_t^{1-k/n}\sqrt{n}(Q_n(s)-Q(s))ds}{(k/n)^{1/2}X_{n-k,n}} \stackrel{d}{=} -\frac{\int_0^{1-k/n}\mathbb{B}_n(s)dQ(s)}{(k/n)^{1/2}Q(1-k/n)} + o_{\mathbb{P}}(1). \tag{24}$$

Also note that this result in (24) [when $t=0$] is equivalent to that of [37]. More precisely, we have

$$\frac{\int_0^{1-k/n}\sqrt{n}(Q_n(s)-Q(s))ds}{(k/n)^{1/2}X_{n-k,n}} \stackrel{d}{=} -\frac{\int_0^{1-k/n}\mathbb{B}_n(s)dQ(s)}{(k/n)^{1/2}Q(1-k/n)} + o_{\mathbb{P}}(1). \tag{25}$$

This concludes that for $0 \le t < 1-k/n$, we have, as $n \to \infty$:

$$\frac{\sqrt{n}\Big(U_{n,k,1}(Q_n,t) - U_{n,k,1}(Q,t)\Big)}{(k/n)^{1/2}X_{n-k,n}} \stackrel{d}{=} -\frac{\int_0^{1-k/n}\mathbb{B}_n(s)dQ(s)}{(k/n)^{1/2}Q(1-k/n)} + o_{\mathbb{P}}(1). \tag{26}$$

Next, we remark also that, $L(Q,\alpha) := U_{n,k,1}(Q,0) - U_{n,k,1}(Q,\alpha)$ and $L_n(\alpha) := U_{n,k,1}(Q_n,0) - U_{n,k,1}(Q,\alpha)$. This leads to:

$$\frac{\sqrt{n}\Big(L_n(\alpha) - L(Q,\alpha)\Big)}{(k/n)^{1/2}X_{n-k,n}} = \frac{\sqrt{n}\Big(U_{n,k,1}(Q_n,0) - U_{n,k,1}(Q,0)\Big)}{(k/n)^{1/2}X_{n-k,n}}$$

$$- \frac{\sqrt{n}\Big(U_{n,k,1}(Q_n,\alpha) - U_{n,k,1}(Q,\alpha)\Big)}{(k/n)^{1/2}X_{n-k,n}}.$$

From (26), we get for all large values of n:

$$\frac{\sqrt{n}\Big(L_n(\alpha) - L(Q,\alpha)\Big)}{(k/n)^{1/2}X_{n-k,n}} = o_{\mathbb{P}}(1). \tag{27}$$

The Lemma 1 follows. ■

Proof of Lemma 2. We first have:

$$\begin{aligned}\frac{U_{n,k,1}(Q_n,\beta)}{L_n(\alpha)} - \frac{U_{n,k,1}(Q,\beta)}{L(Q,\alpha)} &= \frac{U_{n,k,1}(Q_n,\beta)}{L_n(\alpha)} - \frac{U_{n,k,1}(Q_n,\beta)}{L(Q,\alpha)} \\ &\quad + \frac{U_{n,k,1}(Q_n,\beta)}{L(Q,\alpha)} - \frac{U_{n,k,1}(Q,\beta)}{L(Q,\alpha)} \\ &= \frac{1}{L(Q,\alpha)}\Big(U_{n,k,1}(Q_n,\beta) - U_{n,k,1}(Q,\beta)\Big) \\ &\quad - \frac{U_{n,k,1}(Q_n,\beta)}{L_n(\alpha)L(Q,\alpha)}\Big(L_n(\alpha) - L(Q,\alpha)\Big).\end{aligned}$$

This leads to:

$$\begin{aligned}\frac{\sqrt{n}\left\{\frac{U_{n,k,1}(Q_n,\beta)}{L_n(\alpha)} - \frac{U_{n,k,1}(Q,\beta)}{L(Q,\alpha)}\right\}}{(k/n)^{1/2}X_{n-k,n}} &= \frac{1}{L(Q,\alpha)} \times \frac{\sqrt{n}\,(U_{n,k,1}(Q_n,\beta) - U_{n,k,1}(Q,\beta))}{(k/n)^{1/2}X_{n-k,n}} \\ &\quad - \frac{U_{n,k,1}(Q_n,\beta)}{L_n(\alpha)L(Q,\alpha)} \times \frac{\sqrt{n}\,(L_n(\alpha) - L(Q,\alpha))}{(k/n)^{1/2}X_{n-k,n}}.\end{aligned}$$

Next, from (26), we have for all large values of n,

$$\frac{\sqrt{n}\,(U_{n,k,1}(Q_n,\beta) - U_{n,k,1}(Q,\beta))}{L(Q,\alpha)(k/n)^{1/2}X_{n-k,n}} \stackrel{d}{=} \frac{\int_0^{1-k/n}\mathbb{B}_n(s)dQ(s)}{L(Q,\alpha)(k/n)^{1/2}Q(1-k/n)} + o_{\mathbb{P}}(1). \tag{28}$$

Since the right term in (28) is bounded in probability, we get for all large values of n,

$$U_{n,k,1}(Q_n,\beta) = U_{n,k,1}(Q,\beta) + o_{\mathbb{P}}(1).$$

Remarking that $U_{n,k,1}(Q,\beta) = \int_\beta^{1-k/n} Q(s)ds$ and $k/n \to 0$, as $n \to \infty$, we have

$$U_{n,k,1}(Q_n,\beta) = \int_\beta^1 Q(s)ds\Big\{1 + o_{\mathbb{P}}(1)\Big\}.$$

In the other hand, from Lemma 1, we have $L_n(\alpha) = L(Q,\alpha) + o_{\mathbb{P}}(1)$, as $n \to \infty$. Therefore, using again the Lemma 1 and the fact that the lower integral Ł(Q,α) and the upper integral $\int_\beta^1 Q(s)ds$ are finite, we get for all n large enough:

$$\frac{U_{n,k,1}(Q_n,\beta)}{L_n(\alpha)L(Q,\alpha)} \times \frac{\sqrt{n}\left(L_n(\alpha) - L(Q,\alpha)\right)}{(k/n)^{1/2}X_{n-k,n}} = o_{\mathbb{P}}(1). \tag{29}$$

Finally, combining (28) and (29), the Lemma 2 follows. ■

Proof of Lemma 3.
We use the following decomposition:

$$\begin{aligned}\frac{U_{n,k,2}\left(Q_{n,k}^{(K)}\right)}{L_n(\alpha)} - \frac{U_{n,k,2}(Q)}{L(Q,\alpha)} =& \frac{U_{n,k,2}\left(Q_{n,k}^{(K)}\right)}{L_n(\alpha)} - \frac{U_{n,k,2}\left(Q_{n,k}^{(K)}\right)}{L(Q,\alpha)} \\ &+ \frac{U_{n,k,2}\left(Q_{n,k}^{(K)}\right)}{L(Q,\alpha)} - \frac{U_{n,k,2}(Q)}{L(Q,\alpha)} \\ =& \frac{1}{L(Q,\alpha)}\Big(U_{n,k,2}\left(Q_{n,k}^{(K)}\right) - U_{n,k,2}(Q)\Big) \\ &- \frac{U_{n,k,2}\left(Q_{n,k}^{(K)}\right)}{L_n(\alpha)L(Q,\alpha)}\Big(L_n(\alpha) - L(Q,\alpha)\Big).\end{aligned}$$

This implies that:

$$\begin{aligned}\frac{\sqrt{n}\left\{\dfrac{U_{n,k,2}\left(Q_{n,k}^{(K)}\right)}{L_n(\alpha)} - \dfrac{U_{n,k,2}(Q)}{L(Q,\alpha)}\right\}}{(k/n)^{1/2}X_{n-k,n}} =& \frac{1}{L(Q,\alpha)} \times \frac{\sqrt{n}\left(U_{n,k,2}\left(Q_{n,k}^{(K)}\right) - U_{n,k,2}(Q)\right)}{(k/n)^{1/2}X_{n-k,n}} \\ &- \frac{U_{n,k,2}\left(Q_{n,k}^{(K)}\right)}{L_n(\alpha)L(Q,\alpha)} \times \frac{\sqrt{n}\Big(L_n(\alpha) - L(Q,\alpha)\Big)}{(k/n)^{1/2}X_{n-k,n}}\end{aligned} \tag{30}$$

Recall that

$$U_{n,k,2}\left(Q_{n,k}^{(K)}\right) = \frac{k/n}{1-\widehat{\gamma}_{n,k}^{(K)}}X_{n-k,n}.$$

According to Theorem 1 in [13], we have as $n \to \infty$:

$$\sqrt{k}\left(\widehat{\gamma}_{n,k}^{(K)} - \gamma\right) \stackrel{d}{=} \sqrt{k}\, a\left(n/k\right)\int_0^1 s^{-\rho}K(s)ds + \gamma\sqrt{\frac{n}{k}}\int_0^1 s^{-1}\mathbb{B}_n\left(1 - s\frac{k}{n}\right)d\left(sK(s)\right) + o_{\mathbb{P}}(1) \tag{31}$$

This leads to the weak consistency of $\widehat{\gamma}_{n,k}^{(K)}$ to γ. Since $Q(1-\cdot)$ is a regularly varying function at zero with index $-\gamma$, then from Theorem 2.4.1 in [10], $X_{n-k,n} = Q(1-k/n)(1+o_{\mathbb{P}}(1))$, as $n \to \infty$ and $(k/n)Q(1-k/n) = (k/n)^{1-\gamma}\ell_Q(k/n)$. Since $\gamma \in (1/2,1)$, we have from Proposition 1.3.6 in [5], $(k/n)^{1-\gamma}\ell_Q(k/n) \to 0$, as $n \to \infty$. Therefore $U_{n,k,2}\left(Q_{n,k}^{(K)}\right) \overset{\mathbb{P}}{\to} 0$, as $n \to \infty$.

Finally, according to the Lemma 1 in this paper, the second right term of the Equation 30 is equal to $o_{\mathbb{P}}(1)$.

Now, it allows us to look at the first right term of the Equation 30. Clearly $X_{n-k,n} \overset{d}{=} \mathbb{U}(Y_{n-k,n})$ with $\mathbb{U}(x) = Q(1-1/x)$ and

$$U_{n,k,2}\left(Q_{n,k}^{(K)}\right) \overset{d}{=} \frac{k/n}{1-\widehat{\gamma}_{n,k}^{(K)}}\mathbb{U}\left(Y_{n-k,n}\right).$$

By remarking that $X_{n-k,n} = \mathbb{U}(n/k)(1+o_{\mathbb{P}}(1))$ with $\mathbb{U}\left(n/k\right) = Q(1-k/n)$, we have:

$$\frac{\sqrt{n}\left(U_{n,k,2}\left(Q_{n,k}^{(K)}\right) - U_{n,k,2}(Q)\right)}{L(Q,\alpha)(k/n)^{1/2}X_{n-k,n}} \overset{d}{=} \frac{\sqrt{n}\left(U_{n,k,2}\left(Q_{n,k}^{(K)}\right) - U_{n,k,2}(Q)\right)}{L(Q,\alpha)(k/n)^{1/2}\mathbb{U}\left(n/k\right)}\Big\{1+o_{\mathbb{P}}(1)\Big\}.$$

As a consequence, the following expansion holds:

$$\frac{\sqrt{n}\left(U_{n,k,2}\left(Q_{n,k}^{(K)}\right) - U_{n,k,2}(Q)\right)}{L(Q,\alpha)(k/n)^{1/2}\mathbb{U}\left(n/k\right)} \overset{d}{=} \sum_{j=1}^{4} T_{n,j},$$

where

$$\begin{aligned}
T_{n,1} &:= \frac{\sqrt{k}}{L(Q,\alpha)\left(1-\widehat{\gamma}_{n,k}^{(K)}\right)}\left[\frac{\mathbb{U}\left(Y_{n-k,n}\right)}{\mathbb{U}(n/k)} - \left(\frac{k}{n}Y_{n-k,n}\right)^{\gamma}\right],\\
T_{n,2} &:= \frac{\sqrt{k}}{L(Q,\alpha)(1-\widehat{\gamma}_{n,k}^{(K)})}\left[\left(\frac{k}{n}Y_{n-k,n}\right)^{\gamma} - 1\right],\\
T_{n,3} &:= \frac{1}{L(Q,\alpha)\ (1-\widehat{\gamma}_{n,k}^{(K)})(1-\gamma)}\sqrt{k}\left(\widehat{\gamma}_{n,k}^{(K)} - \gamma\right),\\
T_{n,4} &:= \frac{\sqrt{n}}{L(Q,\alpha)\,(k/n)^{1/2}\,\mathbb{U}(n/k)}\left[\frac{k/n}{1-\gamma}\mathbb{U}(n/k) - U_{n,k,2}(Q,\beta)\right].
\end{aligned}$$

We study each term separately.
<u>Term $T_{n,1}$</u>. According to the Theorem 2.3.9 in [10], for any $\delta > 0$, we have

$$\sqrt{k}\left(\frac{\mathbb{U}\left(Y_{n-k,n}\right)}{\mathbb{U}(n/k)} - \left(\frac{k}{n}Y_{n-k,n}\right)^{\gamma}\right) = \sqrt{k}\ a\left(\frac{n}{k}\right)\left\{\left(\frac{k}{n}Y_{n-k,n}\right)^{\gamma}\frac{\left(\frac{k}{n}Y_{n-k,n}\right)^{\rho}-1}{\rho} + o_{\mathbb{P}}(1)\left(\frac{k}{n}Y_{n-k,n}\right)^{\gamma+\rho\pm\delta}\right\},$$

We study each term separately.
<u>Term $T_{n,1}$</u>. According to [10] Theorem 2.3.9) , for any $\delta > 0$, we have

$$\sqrt{k}\left(\frac{\mathbb{U}\left(Y_{n-k,n}\right)}{\mathbb{U}(n/k)} - \left(\frac{k}{n}Y_{n-k,n}\right)^{\gamma}\right) = \sqrt{k}\ a\left(\frac{n}{k}\right)\left\{\left(\frac{k}{n}Y_{n-k,n}\right)^{\gamma}\frac{\left(\frac{k}{n}Y_{n-k,n}\right)^{\rho}-1}{\rho} + o_{\mathbb{P}}(1)\left(\frac{k}{n}Y_{n-k,n}\right)^{\gamma+\rho\pm\delta}\right\},$$

Thus, since $kY_{n-k,n}/n \to 1$, $\sqrt{k}\ a(n/k) \to \lambda \in \mathbb{R}$ and $\widehat{\gamma}_{n,k}^{(K)} \overset{\mathbb{P}}{\to} \gamma$, as $n \to \infty$, it readily follows that

$$T_{n,1} = o_{\mathbb{P}}(1). \tag{32}$$

Term $T_{n,2}$. The equality $Y_{n-k,n} \overset{d}{=} (1-\xi_{n-k,n})^{-1}$ yields

$$\begin{aligned}
\sqrt{k}\left[\left(\frac{k}{n}Y_{n-k,n}\right)^{\gamma} - 1\right] &\overset{d}{=} \sqrt{k}\left(\left(\frac{n}{k}(1-\xi_{n-k,n})\right)^{-\gamma} - 1\right) \\
&= -\gamma\sqrt{k}\left(\frac{n}{k}(1-\xi_{n-k,n}) - 1\right)(1+o_{\mathbb{P}}(1)) \quad \text{by a Taylor expansion} \\
&= -\gamma\sqrt{\frac{n}{k}}\beta_n\left(1-\frac{k}{n}\right)(1+o_{\mathbb{P}}(1)) \\
&= -\gamma\sqrt{\frac{n}{k}}\left(\mathbb{B}_n\left(1-\frac{k}{n}\right) + O_{\mathbb{P}}(n^{-\nu})\left(\frac{k}{n}\right)^{1/2-\nu}\right)(1+o_{\mathbb{P}}(1)),
\end{aligned}$$

for $0 \le \nu < 1/2$, by [8]. Thus, using again the fact that $\widehat{\gamma}_{n,k}^{(K)} \overset{\mathbb{P}}{\to} \gamma$, it follows that

$$T_{n,2} \overset{d}{=} -\frac{\gamma}{L(Q,\alpha)(1-\gamma)}\sqrt{\frac{n}{k}}\mathbb{B}_n\left(1-\frac{k}{n}\right)(1+o_{\mathbb{P}}(1)) = W_{n,\alpha,2} + o_{\mathbb{P}}(1). \tag{33}$$

Term $T_{n,3}$. By using again the weak consistency of $\widehat{\gamma}_{n,k}^{(K)}$ to γ and the equation 31, we get

$$\begin{aligned}
T_{n,3} &\overset{d}{=} \frac{1}{L(Q,\alpha)(1-\gamma)^2}\left\{\sqrt{k}\ a(n/k)\int_0^1 s^{-1}K(s)ds + \gamma\sqrt{\frac{n}{k}}\int_0^1 s^{-1}\mathbb{B}_n\left(1-s\frac{k}{n}\right)d\left(sK(s)\right)\right\} + o_{\mathbb{P}}(1) \\
&= \frac{1}{L(Q,\alpha)\ (1-\gamma)^2}\sqrt{k}\ a(n/k)\int_0^1 s^{-1}K(s)ds + W_{n,\alpha,3} + o_{\mathbb{P}}(1). \qquad (34)
\end{aligned}$$

Term $T_{n,4}$. A change of variables and an integration by parts yield

$$\begin{aligned}
T_{n,4} &= \frac{\sqrt{k}}{L(Q,\alpha)}\left\{\frac{1}{1-\gamma} - \int_1^\infty x^{-2}\frac{\mathbb{U}(nx/k)}{\mathbb{U}(n/k)}dx\right\} \\
&= -\frac{\sqrt{k}}{L(Q,\alpha)}\int_1^\infty x^{-2}\left(\frac{\mathbb{U}(nx/k)}{\mathbb{U}(n/k)} - x^{\gamma}\right)dx.
\end{aligned}$$

Theorem 2.3.9 in [10] entails that, for $\gamma \in (1/2, 1)$,

$$\begin{aligned}
T_{n,4} &= -\frac{\sqrt{k}a\left(\frac{n}{k}\right)}{L(Q,\alpha)}\int_1^\infty x^{\gamma-2}\frac{x^{\rho}-1}{\rho}dx\,(1+o_{\mathbb{P}}(1)) \\
&= \frac{\sqrt{k}a\left(\frac{n}{k}\right)}{L(Q,\alpha)}\frac{1}{(1-\gamma)(\gamma+\rho-1)}(1+o_{\mathbb{P}}(1)). \qquad (35)
\end{aligned}$$

Combining (32)-(35), Lemma 3 follows. ■

Remark 1. *Since $Q(1-\cdot)$ is a regularly varying function at zero with index $-\gamma$, then from Theorem 2.4.1 in [10], $X_{n-k,n} = Q(1-k/n)(1+o_{\mathbb{P}}(1))$, as $n \to \infty$ and $(k/n)^{1/2}Q(1-k/n) = (k/n)^{1/2-\gamma}\ell_Q(k/n)$, for $k \to \infty$, $k/n \to 0$, as $n \to \infty$. Remarking that $n^{1/2}/(k/n)^{1/2}Q(1-k/n)$ is equal to $k^{1/2}/(k/n)^{1-\gamma}\ell_Q(k/n)$ and since $\gamma \in (1/2, 1)$, we have from Proposition 1.3.6 in [5], $(k/n)^{1-\gamma}\ell_Q(k/n) \to 0$, as $n \to \infty$. and the rate of convergence in Theorem 1 tends to infinity as n goes to infinity.*

From Theorem 1, it is clear that the estimator $\widehat{\eta}_{n,k}^{(K)}(\alpha,\beta)$ exhibits a bias due to the fact that we use in its construction a symptomatic estimator of $Q(\cdot)$ derived from the Weissman's type estimator $Q_{n,k}^{(K)}(\cdot)$, which is known to have such a problem. To solve this issue, we propose in the next section to use a bias reduction method and to introduce an improved estimator of the QSR index $\eta(Q,\alpha,\beta)$.

4.2. Reduced bias estimation of the QSR index

In this section, we propose to substitute in (13), the Weissman's estimator $Q_{n,k}^{(K)}$ with an asymptotically unbiased estimator of the extreme quantile. Our approach is similar to the bias reduction procedure introduced in [19] and [26].

In order to find an asymptotically unbiased estimator of the extreme quantile, we use the second order condition $(\mathcal{R}_{\mathbb{U}})$, for which the following approximation holds:

$$Q(u) \approx \left(\frac{n}{k}(1-u)\right)^{-\gamma} Q(1-k/n)\left\{1 - \frac{a(n/k)}{\rho}\left[1-\left(\frac{n}{k}(1-u)\right)^{-\rho}\right]\right\},\ u \to 1, \tag{36}$$

where γ, $a(\cdot)$ and ρ are unknown. The first part $\left(\frac{n}{k}(1-u)\right)^{-\gamma} Q(1-k/n)$ in the right side of (36) is exactly estimated by the Weissman's type estimator $Q_{n,k}^{(K)}(u)$ and defined in (9). Clearly, the estimator $Q_{n,k}^{(K)}$ exhibits a potential bias because it depends on the Kernel type estimator $\widehat{\gamma}_{n,k}^{(K)}$ of the tail index γ, which from (16) has such problem. The expression $1-\rho^{-1}a(n/k_n)[1-\left(\frac{n}{k}(1-u)\right)^{-\rho}]$ can be viewed as a correcting term since $a(n/k_n)$ tends to 0. This leads to the necessity to find good estimators for γ, $a(n/k_n)$ and ρ.

We first propose to introduce an asymptotically unbiased estimator for γ by following an approach similar to [19]. To this end, consider two kernel functions K_1 and K_2 satisfying $(\mathbb{K})$ and define a mixture of them in the form $K_\Delta(s) = \Delta K_1(s) + (1-\Delta)K_2(s)$, for $\Delta \in \mathbb{R}$. Clearly K_Δ also satisfies the condition $(\mathbb{K})$ and hence by the result given in (16), the asymptotic bias $\lambda\int_0^1 s^{-\rho}K_\Delta(s)ds$ of $\widehat{\gamma}_{n,k}^{(K_\Delta)}$ is such that

$$\lambda\int_0^1 s^{-\rho}K_\Delta(s)ds = \lambda\Delta\int_0^1 s^{-\rho}K_1(s)ds + \lambda(1-\Delta)\int_0^1 s^{-\rho}K_2(s)ds.$$

Equating the right-hand side of the above equation to zero leads to the value of eliminating the asymptotic bias

$$\Delta^* = \frac{\int_0^1 s^{-\rho} K_2(s) ds}{\int_0^1 s^{-\rho} \Big\{ K_2(s) - K_1(s) \Big\} ds}, \tag{37}$$

provided $\int_0^1 s^{-\rho} \Big\{ K_2(s) - K_1(s) \Big\} ds \neq 0$. Clearly, the tail index estimator $\widehat{\gamma}_{n,k}^{(K_{\Delta^*})}$ is shown to be asymptotically unbiased in the sense that the mean of its limiting distribution is zero, whatever the value of λ. More precisely, we have from (16):

$$k^{1/2} \left(\widehat{\gamma}_{n,k}^{(K_{\Delta^*})} - \gamma \right) \xrightarrow{d} \mathcal{N} \left(0, \gamma^2 \int_0^1 K_{\Delta^*}^2(s) ds \right). \tag{38}$$

An open problem is to determine whether among this class of unbiased estimators $\widehat{\gamma}_{n,k}^{(K_{\Delta^*})}$, we can find the asymptotically unbiased estimator with minimum variance. Clearly, the asymptotic variance $\gamma^2 \int_0^1 K_{\Delta^*}^2(s) ds$ is minimal for a minimum value of $\int_0^1 K_{\Delta^*}^2(s) ds$. According to [19] and [13], the minimum of $\int_0^1 K_{\Delta^*}^2(s) ds$ is obtained at the "optimal" function given by:

$$K_{\Delta^*_{\text{opt}}}(s) = \left(\frac{1-\rho}{\rho} \right)^2 - \frac{(1-\rho)(1-2\rho)}{\rho^2} s^{-\rho}, \text{ for } \quad s \in (0,1), \tag{39}$$

and $K_{\Delta^*_{\text{opt}}}(s) = 0$ otherwise. Note that this unction can be viewed as a mixture between two power kernels: $K_1(s) := \underline{K}(s) = \mathbb{I}_{(0<s<1)}$ and $K_2(s) := K_{2,\rho}(s) := (1-\rho)\, s^{-\rho} \mathbb{I}_{(0<s<1)}$ and $\Delta^* = (1-\rho)^2/\rho^2$ is as in (37). In that case, the minimal variance $\gamma^2 \int_0^1 K_{\Delta^*_{\text{opt}}}^2(s) ds$ equals to $\gamma^2(1-\rho)^2/\rho^2$.

From a practical point of view, the unbiased tail index estimator with minimum variance $\widehat{\gamma}_{n,k}^{(K_{\Delta^*_{\text{opt}}})}$ cannot be obtained directly, since it depends on the unknown parameters and expressions: γ, ρ, $a(n/k)$ and $K_{\Delta^*_{\text{opt}}}$ are unknown. To solve this issue, we propose to replace ρ by $\widehat{\rho}$, where $\widehat{\rho}$ is either a canonical negative value $\widehat{\rho} = \rho = \rho_0$ or an external estimator $\widehat{\rho} = \widehat{\rho}_{k_\rho}$, consistent in probability to ρ, with $k_\rho := k_\rho(n)$ an intermediate sequence of integers greater than k, satisfying $k_\rho \to \infty$ and $k_\rho/n \to 0$, as $n \to \infty$. Finally, as in (11), we arrive to the following tail index estimator:

$$\widehat{\gamma}_{n,k}^{(K_{\widehat{\Delta}^*_{opt}})} = \frac{1}{k} \sum_{j=1}^{k} j K_{\widehat{\Delta}^*_{opt}} \left(\frac{j}{k+1} \right) \log \left(\frac{X_{n-j+1,n}}{X_{n-j,n}} \right),$$

where $K_{\widehat{\Delta}^*_{opt}}$ is defined as $K_{\Delta^*_{opt}}$ in (39) with ρ replaced by $\widehat{\rho}$.

Next, for the estimation of the rate $a(\cdot)$, we use the result in (16) from which we have, as $n \to \infty$,

$$\widehat{\gamma}_{n,k}^{(\underline{K})} - \widehat{\gamma}_{n,k}^{(K_{2,\rho})} = -a(n/k) \frac{\rho^2}{(1-\rho)(1-2\rho)} + o_{\mathbb{P}}(1).$$

Thus, we can approximate

$$a(n/k)\frac{\rho^2}{(1-\rho)(1-2\rho)} \qquad \text{by} \qquad -\left\{\widehat{\gamma}_{n,k}^{(\underline{K})} - \widehat{\gamma}_{n,k}^{(K_{2,\rho})}\right\},$$

which mean that $a(n/k)$ can be estimated by;

$$\widehat{a}_{n,k}(\widehat{\rho}\,) := -\frac{(1-\widehat{\rho})(1-2\widehat{\rho})}{\widehat{\rho}^{\,2}}\left\{\widehat{\gamma}_{n,k}^{(\underline{K})} - \widehat{\gamma}_{n,k}^{(K_{2,\widehat{\rho}})}\right\}.$$

Clearly, the estimators $\widehat{\gamma}_{n,k}^{(K_{\widehat{\Delta}^*_{opt}})}$ and $\widehat{a}_{n,k}(\widehat{\rho}\,)$ can be easily viewed as the least squared based estimators of γ and $a(n/k)$ studied in [16]; [3]; [4] and [13]. This approach is based on the following exponential regression model:

$$j\log\left(\frac{X_{n-j+1,n}}{X_{n-j,n}}\right) \sim \left(\gamma + A(n/k)\left(\frac{j}{k+1}\right)^{-\rho}\right) + \varepsilon_{j,k},\ 1\leq j\leq k, \tag{40}$$

where $\varepsilon_{j,k}$ are zero-centered error terms and in which ρ is substituted by $\widehat{\rho}$.

Finally, using the relation in (36), we arrive at the following unbiased estimator of the extreme quantile $Q(u)$, $u\to 1$:

$$Q_{n,k,\widehat{\rho}}^{(K_{\widehat{\Delta}^*_{opt}})}(u) = \left(\frac{n}{k}(1-u)\right)^{-\widehat{\gamma}_{n,k}^{(K_{\widehat{\Delta}^*_{opt}})}} X_{n-k,n}\left\{1-\frac{\widehat{a}_{n,k}}{\rho}\left[1-\left(\frac{n}{k}(1-u)\right)^{-\widehat{\rho}}\right]\right\}. \tag{41}$$

In the spirit of (2), substituting the extreme quantile $Q(u)$ with $Q_{n,k,\widehat{\rho}}^{(K_{\widehat{\Delta}^*_{opt}})}(u)$, we obtain the following unbiased estimator of the QSR index

$$\begin{aligned}\widetilde{\eta}_{n,k,\widehat{\rho}}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}(\alpha,\beta) : &= \frac{1}{L_n(\alpha)}\sum_{j=1}^{n-k}\left[\left(\frac{j}{n}-\beta\right)_+ - \left(\frac{j-1}{n}-\beta\right)_+\right]X_{j,n} \\ &\quad + \frac{(k/n)\ X_{n-k},n}{L_n(\alpha)\left(1-\widehat{\gamma}_{n,k}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}\right)}\left\{1-\frac{\widehat{a}_{n,k}(\widehat{\rho}\,)}{\widehat{\gamma}_{n,k}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}+\widehat{\rho}-1}\right\}.\end{aligned} \tag{42}$$

A possible choice for $\widehat{\rho}_{k_\rho}$ is one of the most performed estimator among those studied in [21], generalized in [12]) and defined by:

$$\widehat{\rho}_{k_\rho} = \frac{6S_{k_\rho}^{(2)} - 4 + \sqrt{3S_{k_\rho}^{(2)} - 2}}{4S_{k_\rho}^{(2)} - 3}, \quad \text{provided} \quad S_{k_\rho}^{(2)} \in \left(\frac{2}{3},\frac{3}{4}\right), \tag{43}$$

where

$$S_{k_\rho}^{(2)} = \frac{3}{4} \frac{\left[M_{k_\rho}^{(4)} - 24\left(M_{k_\rho}^{(1)}\right)^4\right]\left[M_{k_\rho}^{(2)} - 2\left(M_{k_\rho}^{(1)}\right)^2\right]}{\left[M_{k_\rho}^{(3)} - 6\left(M_{k_\rho}^{(1)}\right)^3\right]^2},$$

and

$$M_{k_\rho}^{(r)} := \frac{1}{k_\rho} \sum_{j=1}^{k_\rho} \left(\log \frac{X_{n-j+1,n}}{X_{n-k_\rho,n}}\right)^r, \ r > 0.$$

The consistency of $\widehat{\rho}_{k_\rho}^{(*)}$ to ρ have been established in [21] and [12]) under the second order condition $(\mathcal{R}_{\mathbb{U}})$ and the assumptions $k_\rho \to \infty$, $k_\rho/n \to 0$ and $k_\rho^{1/2} a(n/k_\rho) \to \infty$, as $n \to \infty$.

Our next goal is to establish, under suitable assumptions, the asymptotic normality of $\widetilde{\eta}_{n,k,\widehat{\rho}}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}(\alpha, \beta)$. This is done in the following theorem.

Theorem 2. *Under the assumptions of Theorem 1, if $\widehat{\rho}$ is either a canonical negative value $\widehat{\rho} = \rho = \rho_0$ or an external estimator $\widehat{\rho} = \widehat{\rho}_{k_\rho}$, consistent in probability to ρ, with $k_\rho := k_\rho(n)$, an intermediate sequence of integers greater than k, satisfying $k_\rho \to \infty$ and $k_\rho/n \to 0$, as $n \to \infty$, then we have:*

$$\frac{\sqrt{n}\left(\widetilde{\eta}_{n,k,\widehat{\rho}}^{(K_{\widehat{\Delta}^*_{opt}})}(\alpha, \beta) - \eta(Q, \alpha, \beta)\right)}{(k/n)^{1/2} X_{n-k,n}} \xrightarrow{d} \mathcal{N}\left(0, \widetilde{\mathcal{AV}}(\gamma, \rho, \alpha)\right),$$

where

$$\widetilde{\mathcal{AV}}(\gamma, \rho, \alpha) = \frac{\gamma^4(\gamma - \rho)^2}{(2\gamma - 1)(1 - \gamma)^4(\gamma + \rho - 1)^2 L^2(Q, \alpha)}.$$

Proof of Theorem 2

For simplify the proof, we introduce the following Lemmas, whose proofs are given just after this one.

Lemma 4. *Suppose that the distribution F satisfies the second order condition $(\mathcal{R}_{\mathbb{U}})$. If $k \to \infty$, $k/n \to 0$ and $\sqrt{k}\, a(n/k) \to \lambda \in \mathbb{R}$, as $n \to \infty$ and $\widehat{\rho}$ is either a canonical negative value $\widehat{\rho} = \rho = \rho_0$ or an external estimator $\widehat{\rho} = \widehat{\rho}_{k_\rho}$, consistent in probability to ρ, with $k_\rho := k_\rho(n)$ an intermediate sequence of integers greater than k, satisfying $k_\rho \to \infty$ and $k_\rho/n \to 0$, as $n \to \infty$, then we have*

$$\sqrt{k}\left(\widehat{\gamma}_{n,k}^{\left(K_{\widehat{\Delta}^*_{opt}}\right)} - \gamma\right) \stackrel{d}{=} \gamma\sqrt{n/k} \int_0^1 s^{-1}\mathbb{B}_n(1 - sk/n) d(sK_{\Delta^*_{opt}}(s)) + o_{\mathbb{P}}(1)$$

and

$$\sqrt{k}\left(\widehat{a}_{n,k}(\widehat{\rho}) - a(n/k)\right) \stackrel{d}{=} \gamma(1 - \rho)\sqrt{n/k} \int_0^1 s^{-1}\mathbb{B}_n(1 - sk/n) d(s(K_1(s) - K_{\Delta^*_{opt}}(s))) + o_{\mathbb{P}}(1).$$

Lemma 5. *Under the assumptions of Theorem 2, we have for $0 < \alpha < 1 - k/n$:*

$$\frac{\sqrt{n}\left\{\dfrac{U_{n,k,2}\Big(Q_{n,k,\widehat{\rho}}^{(K_{\widehat{\Delta}^*_{opt}})}\Big)}{L_n(\alpha)} - \dfrac{U_{n,k,2}(Q)}{L(Q,\alpha)}\right\}}{(k/n)^{1/2}X_{n-k,n}} \stackrel{d}{=} W_{n,\alpha,1} + W_{n,\alpha,2} + W_{n,\alpha,4} + W_{n,\alpha,5} + o_{\mathbb{P}}(1), \quad (44)$$

as $n \to \infty$, where

$$\begin{cases} W_{n,\alpha,4} := \dfrac{\rho\gamma^2}{(\gamma+\rho-1)(1-\gamma)^2\ L(Q,\alpha)}\sqrt{n/k}\displaystyle\int_0^1 s^{-1}\mathbb{B}_n(1-sk/n)d(sK_{\Delta^*_{opt}}(s)) \\ W_{n,\alpha,5} := -\dfrac{(1-\gamma)(1-\rho)}{(\gamma+\rho-1)}W_{n,\alpha,3}. \end{cases}$$

Now coming back to this proof, under the assumptions, we have:

$$\begin{aligned} \widehat{\eta}_{n,k}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}(\alpha,\beta) - \eta(Q,\alpha,\beta) &= \left\{\frac{U_{n,k,1}(Q_n,\beta)}{L_n(\alpha)} - \frac{U_{n,k,1}(Q,\beta)}{L(Q,\alpha)}\right\} \\ &\quad + \left\{\frac{U_{n,k,2}\left(Q_{n,k}^{K_{\widehat{\Delta}^*_{\text{opt}}})}\right)}{L_n(\alpha)} - \frac{U_{n,k,2}(Q)}{L(Q,\alpha)}\right\} \\ &:= B_{n,1} + B_{n,2}. \end{aligned} \quad (45)$$

For all values of n large enough, we get respectively from Lemma 4 and Lemma 5:

$$\frac{\sqrt{n}\ B_{n,1}}{(k/n)^{1/2}X_{n-k,n}} \stackrel{d}{=} W_{n,\alpha,1} + o_{\mathbb{P}}(1),$$

and

$$\frac{\sqrt{n}B_{n,2}}{(k/n)^{1/2}X_{n-k,n}} \stackrel{d}{=} W_{n,\alpha,2} + W_{n,\alpha,4} + W_{n,\alpha,5} + o_{\mathbb{P}}(1).$$

This leads to

$$\frac{\sqrt{n}\ \left(\widehat{\eta}_{n,k}^{(K)}(\alpha,\beta) - \eta(Q,\alpha,\beta)\right)}{(k/n)^{1/2}X_{n-k,n}} \stackrel{d}{=} W_{n,\alpha,1} +\ W_{n,\alpha,2} + W_{n,\alpha,4} + W_{n,\alpha,5} + o_{\mathbb{P}}(1). \quad (46)$$

We only have to compute the asymptotic variance of the sum of process in the right term of (46).

As in Theorem 1, the computations are quite direct and the desired asymptotic variance can be obtained by noticing that

$$\mathbb{E}W_{n,\alpha,5}^2 \quad \longrightarrow \quad \frac{\gamma^2(1-\rho)^2}{(1-\gamma)^2(\gamma+\rho-1)^2\,L^2(Q,\alpha)}$$

$$\begin{aligned}
\mathbb{E}(W_{n,\alpha,1}W_{n,\alpha,4}) &= 0\\
\mathbb{E}(W_{n,\alpha,5}) &= 0\\
\mathbb{E}W_{n,\alpha,4}^2 &= \frac{\gamma^4(1-\rho)^2}{(1-\gamma)^4(\gamma+\rho-1)^2\ L^2(Q,\alpha)}\\
\mathbb{E}(W_{n,\alpha,2}W_{n,\alpha,5}) &= 0\\
\mathbb{E}(W_{n,\alpha,}W_{n,\alpha,4}) &= 0\\
\mathbb{E}(W_{n,\alpha,4}W_{n,\alpha,5}) &= -\frac{\rho\gamma^3(1-\rho)}{(1-\gamma)^3(\gamma+\rho-1)^2\, L^2(Q,\alpha)}.
\end{aligned}$$

Combining all these results, Theorem 2 follows. ■

Proof of Lemma 4.

Note that the first quantity of interest can be expanded as

$$\begin{aligned}
\sqrt{k}\left(\widehat{\gamma}_{n,k}^{\left(K_{\widehat{\Delta}^*_{\text{opt}}}\right)}-\gamma\right) &= \sqrt{k}\left(\widehat{\gamma}_{n,k}^{\left(K_{\widehat{\Delta}^*_{\text{opt}}}\right)}-\widehat{\gamma}_{n,k}^{\left(K_{\Delta^*_{\text{opt}}}\right)}\right)+\sqrt{k}\left(\widehat{\gamma}_{n,k}^{\left(K_{\Delta^*_{\text{opt}}}\right)}-\gamma\right)\\
&= \sqrt{k}\frac{1}{k}\sum_{j=1}^{k}\left\{K_{\widehat{\Delta}^*_{\text{opt}}}\left(\frac{j}{k+1}\right)-K_{\Delta^*_{\text{opt}}}\left(\frac{j}{k+1}\right)\right\}j\log\left(\frac{X_{n-j+1,n}}{X_{n-j,n}}\right)\\
&\quad +\sqrt{k}\left(\widehat{\gamma}_{n,k}^{\left(K_{\Delta^*_{\text{opt}}}\right)}-\gamma\right),
\end{aligned}$$

where

$$K_{\Delta^*_{\text{opt}}}(s)=\left(\frac{(1-\rho)^2}{\rho^2}\right)^2-\frac{(1-\rho)(1-2\rho)}{\rho^2}s^{-\rho},\ t\in(0,1),$$

and $K_{\Delta^*_{\text{opt}}}(s)=0$ otherwise,

$K_{\widehat{\Delta}^*_{\text{opt}}}$ is defined as $K_{\Delta^*_{\text{opt}}}$ with ρ replaced by $\widehat{\rho}$. We have all ready mentioned that the function $K_{\Delta^*_{\text{opt}}}$ is viewed as a mixture between two power kernels: $K_1(s):=\mathbb{I}_{(0<s<1)}$ and $K_{2,\rho}(s):=(1-\rho)\,s^{-\rho}\mathbb{I}_{(0<s<1)}$ with and $\Delta^*=(1-\rho)^2/\rho^2$. Thus, according to the proof of Theorem 3.2 of [4], we have

$$\sqrt{k}\frac{1}{k}\sum_{j=1}^{k}\left\{K_{\widehat{\Delta}^*_{\text{opt}}}\left(\frac{j}{k+1}\right)-K_{\Delta^*_{\text{opt}}}\left(\frac{j}{k+1}\right)\right\}j\log\left(\frac{X_{n-j+1,n}}{X_{n-j,n}}\right)=o_{\mathbb{P}}(1),$$

and

$$\sqrt{k}\left(\widehat{\gamma}_{n,k}^{\left(K_{\widehat{\Delta}^*_{\text{opt}}}\right)}-\gamma\right)=\sqrt{k}\left(\widehat{\gamma}_{n,k}^{\left(K_{\Delta^*_{\text{opt}}}\right)}-\gamma\right)+o_{\mathbb{P}}(1).$$

Recall now that

$$\widehat{\gamma}_{n,k}^{\left(K_{\Delta^*_{\text{opt}}}\right)}=\frac{(1-\rho)^2}{\rho^2}\widehat{\gamma}_{n,k}^{(K_1)}-\frac{(1-2\rho)}{\rho^2}\widehat{\gamma}_{n,k}^{(K_{2,\rho})}$$

We use the following decomposition,

$$\sqrt{k}\left(\widehat{\gamma}_{n,k}^{\left(K_{\Delta^*_{\text{opt}}}\right)}-\gamma\right)=\frac{(1-\rho)^2}{\rho^2}\sqrt{k}\left(\widehat{\gamma}_{n,k}^{(K_1)}-\gamma\right)-\frac{(1-2\rho)}{\rho^2}\sqrt{k}\left(\widehat{\gamma}_{n,k}^{(K_{2,\rho})}-\gamma\right)$$

From (31), it is clear that

$$\sqrt{k}\left(\widehat{\gamma}_{n,k}^{(K_1)}-\gamma\right)=\sqrt{k}\ \frac{a\,(n/k)}{1-\rho}+\gamma\sqrt{\frac{n}{k}}\int_0^1 s^{-1}\mathbb{B}_n\left(1-s\frac{k}{n}\right)d\left(sK_1(s)\right)+o_{\mathbb{P}}(1)$$

and

$$\sqrt{k}\left(\widehat{\gamma}_{n,k}^{(K_{2,\rho})}-\gamma\right)=\frac{1-\rho}{1-2\rho}\sqrt{k}a\,(n/k)+\gamma\sqrt{\frac{n}{k}}\int_0^1 s^{-1}\mathbb{B}_n\left(1-s\frac{k}{n}\right)d\left(sK_{2,\rho}(s)\right)+o_{\mathbb{P}}(1).$$

Finally, combining these two previews expansions, we get:

$$\begin{aligned}\sqrt{k}\left(\widehat{\gamma}_{n,k}^{\left(K_{\Delta^*_{\text{opt}}}\right)}-\gamma\right) &= \gamma\sqrt{\frac{n}{k}}\int_0^1 s^{-1}\mathbb{B}_n\left(1-s\frac{k}{n}\right)d\left(s\Big\{\frac{(1-\rho)^2}{\rho^2}K_1(s)\Big\}\right)\\ &\quad -\gamma\sqrt{\frac{n}{k}}\int_0^1 s^{-1}\mathbb{B}_n\left(1-s\frac{k}{n}\right)d\left(s\Big\{\frac{1-2\rho}{\rho^2}K_{2,\rho}(s)\Big\}\right)+o_{\mathbb{P}}(1)\\ &= \gamma\sqrt{\frac{n}{k}}\int_0^1 s^{-1}\mathbb{B}_n\left(1-s\frac{k}{n}\right)d\left(sK_{\Delta^*_{\text{opt}}}(s)\right)+o_{\mathbb{P}}(1).\end{aligned}$$

The first part of Lemma 4 follow.

Focussing on the second part and we have

$$\widehat{a}_{n,k}(\widehat{\rho}):=-\frac{(1-\widehat{\rho})(1-2\widehat{\rho})}{\widehat{\rho}^{\,2}}\left\{\widehat{\gamma}_{n,k}^{(K_1)}-\widehat{\gamma}_{n,k}^{(K_{2,\widehat{\rho}})}\right\}.$$

Thus,

$$\begin{aligned}\sqrt{k}\left(\widehat{a}_{n,k}(\widehat{\rho})-a(n/k)\right) &= (1-\widehat{\rho})\sqrt{k}\left(\widehat{\gamma}_{n,k}^{(K_1)}-\gamma-\frac{a(n/k)}{1-\rho}\right)-(1-\widehat{\rho})\sqrt{k}\left(\widehat{\gamma}_{n,k}^{\left(K_{\widehat{\Delta}^*_{\text{opt}}}\right)}-\gamma\right)\\ &\quad +\sqrt{k}\ a(n/k)\left(\frac{(1-\widehat{\rho})}{(1-\rho)}-1\right).\end{aligned}$$

Since $\widehat{\rho}$ is a consistent estimator of ρ, this leads to the desired result. ■

Proof of Lemma 5.
Following the same approach as in the Proof of Lemma 3, we have

$$\begin{aligned}\frac{U_{n,k,2}\left(Q_{n,k,\widehat{\rho}}^{(K_{\widehat{\Delta}^*_{\mathrm{opt}}})}\right)}{L_n(\alpha)}-\frac{U_{n,k,2}(Q)}{L(Q,\alpha)}=&\frac{U_{n,k,2}\left(Q_{n,k,\widehat{\rho}}^{(K_{\widehat{\Delta}^*_{\mathrm{opt}}})}\right)}{L_n(\alpha)}-\frac{U_{n,k,2}(Q)}{L(Q,\alpha)}\\&+\frac{U_{n,k,2}\left(Q_{n,k,\widehat{\rho}}^{(K_{\widehat{\Delta}^*_{\mathrm{opt}}})}\right)}{L(Q,\alpha)}-\frac{U_{n,k,2}(Q)}{L(Q,\alpha)}\\=&\frac{1}{L(Q,\alpha)}\Big(U_{n,k,2}\left(Q_{n,k,\widehat{\rho}}^{(K_{\widehat{\Delta}^*_{\mathrm{opt}}})}\right)-U_{n,k,2}(Q)\Big)\\&-\frac{U_{n,k,2}\left(Q_{n,k,\widehat{\rho}}^{(K_{\widehat{\Delta}^*_{\mathrm{opt}}})}\right)}{L_n(\alpha)L(Q,\alpha)}\Big(L_n(\alpha)-L(Q,\alpha)\Big).\end{aligned}$$

This implies that:

$$\begin{aligned}\frac{\sqrt{n}\left\{\dfrac{U_{n,k,2}\left(Q_{n,k,\widehat{\rho}}^{(K_{\widehat{\Delta}^*_{\mathrm{opt}}})}\right)}{L_n(\alpha)}-\dfrac{U_{n,k,2}(Q)}{L(Q,\alpha)}\right\}}{(k/n)^{1/2}X_{n-k,n}}=&\frac{1}{L(Q,\alpha)}\times\frac{\sqrt{n}\left(U_{n,k,2}\left(Q_{n,k,\widehat{\rho}}^{(K_{\widehat{\Delta}^*_{\mathrm{opt}}})}\right)-U_{n,k,2}(Q)\right)}{(k/n)^{1/2}X_{n-k,n}}\\&-\frac{U_{n,k,2}\left(Q_{n,k,\widehat{\rho}}^{(K_{\widehat{\Delta}^*_{\mathrm{opt}}})}\right)}{L_n(\alpha)L(Q,\alpha)}\times\frac{\sqrt{n}\Big(L_n(\alpha)-L(Q,\alpha)\Big)}{(k/n)^{1/2}X_{n-k,n}}\end{aligned}\qquad(47)$$

Recall that

$$U_{n,k,2}\Big(Q_{n,k,\widehat{\rho}}^{(K_{\widehat{\Delta}^*_{\mathrm{opt}}})}\Big)=\frac{(k/n)\ X_{n-k,n}}{1-\widehat{\gamma}_{n,k}^{(K_{\widehat{\Delta}^*_{\mathrm{opt}}})}}\left\{1-\frac{\widehat{a}_{n,k}(\widehat{\rho})}{\widehat{\gamma}_{n,k}^{(K_{\widehat{\Delta}^*_{\mathrm{opt}}})}+\widehat{\rho}-1}\right\}.$$

For a given $\widehat{\rho}$ be either a canonical negative value $\widehat{\rho}=\rho=\rho_0$ or an external estimator $\widehat{\rho}=\widehat{\rho}_{k_\rho}$, consistent in probability to ρ, with $k_\rho:=k_\rho(n)$ an intermediate sequence of integers greater than k, satisfying $k_\rho\to\infty$ and $k_\rho/n\to 0$, as $n\to\infty$, we have from Lemma 4, $\widehat{\gamma}_{n,k}^{(K_{\widehat{\Delta}^*_{\mathrm{opt}}})}\xrightarrow{\mathbb{P}}\gamma$ and $\widehat{a}_{n,k}(\widehat{\rho})\xrightarrow{\mathbb{P}}0$, as $n\to\infty$. Since $(k/n)\ X_{n-k,n}\xrightarrow{\mathbb{P}}0$ (see the Proof of Lemma 3). Therefore, $U_{n,k,2}\Big(Q_{n,k,\widehat{\rho}}^{(K_{\widehat{\Delta}^*_{\mathrm{opt}}})}\Big)\xrightarrow{\mathbb{P}}0$, as $n\to 0$. Consequently, according to the Lemma 1, the second right term of the Equation 47 is equal to $o_{\mathbb{P}}(1)$.

Now, it allows us to look at the first right term of the Equation 47. Thus, we have the following decomposition:

$$\frac{\sqrt{n}}{L(Q,\alpha)\ (k/n)^{1/2}\mathbb{U}\left(n/k\right)}\left(U_{n,k,2}\Big(Q_{n,k,\widehat{\rho}}^{(K_{\widehat{\Delta}^*_{\mathrm{opt}}})}\Big)-U_{n,k,2}(Q)\right)=\sum_{i=1}^{6}S_{n,i}$$

where

$$
\begin{aligned}
S_{n,1} &= \frac{1}{L(Q,\alpha)\Big(1-\widehat{\gamma}_{n,k}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}\Big)}\left(1-\frac{\widehat{a}_{n,k}(\widehat{\rho})}{\widehat{\gamma}_{n,k}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}+\widehat{\boldsymbol{\rho}}-1}\right)\sqrt{k}\left[\frac{\mathbb{U}\left(Y_{n-k,n}\right)}{\mathbb{U}(n/k)}-\left(\frac{k}{n}Y_{n-k,n}\right)^{\gamma}\right]\\
S_{n,2} &= \frac{1}{L(Q,\alpha)\Big(1-\widehat{\gamma}_{n,k}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}\Big)}\left(1-\frac{\widehat{a}_{n,k}(\widehat{\rho})}{\widehat{\gamma}_{n,k}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}+\widehat{\boldsymbol{\rho}}-1}\right)\sqrt{k}\left[\left(\frac{k}{n}Y_{n-k,n}\right)^{\gamma}-1\right]\\
S_{n,3} &= \frac{1}{L(Q,\alpha)\Big(1-\widehat{\gamma}_{n,k}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}\Big)\Big(1-\gamma\Big)}\sqrt{k}\left(\widehat{\gamma}_{n,k}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}-\gamma\right)\\
S_{n,4} &= \frac{\sqrt{k}a(n/k)}{L(Q,\alpha)}\left[\frac{1}{(1-\gamma)(\gamma+\rho-1)}-\frac{1}{\left(1-\widehat{\gamma}_{n,k}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}\right)\left(\widehat{\gamma}_{n,k}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}+\widehat{\boldsymbol{\rho}}-1\right)}\right]\\
S_{n,5} &= -\frac{1}{L(Q,\alpha)\left(1-\widehat{\gamma}_{n,k}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}\right)\left(\widehat{\gamma}_{n,k}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}+\widehat{\boldsymbol{\rho}}-1\right)}\sqrt{k}\left(\widehat{a}_{n,k}(\widehat{\rho})-a(n/k)\right)\\
S_{n,6} &= \frac{\sqrt{n}}{L(Q,\alpha)(k/n)^{1/2}\mathbb{U}(n/k)}\left[\frac{k/n}{1-\gamma}\ \left(1-\frac{a(n/k)}{\gamma+\rho-1}\right)\mathbb{U}(n/k)-U_{n,k,2}(Q)\right].
\end{aligned}
$$

Next, we are going to study separately the terms $S_{n,1},...,S_{n,6}$.
Term $S_{n,1}$. Note that

$$S_{n,1}=\frac{1-\widehat{\gamma}_{n,k}^{(K)}}{1-\widehat{\gamma}_{n,k}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}}\left(1-\frac{\widehat{a}_{n,k}(\widehat{\rho})}{\widehat{\gamma}_{n,k}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}+\widehat{\boldsymbol{\rho}}-1}\right)T_{n,1}$$

where $T_{n,1}$ is defined in the Proof of Lemma 3. Thus combining Lemma 4 with the consistency of $\widehat{\boldsymbol{\rho}}$ and (32), we obtain that

$$S_{n,1}=o_{\mathbb{P}}(1). \tag{48}$$

Term $S_{n,2}$. Similarly, we observe that $S_{n,2}=T_{n,2}(1+o_{\mathbb{P}}(1))$ where $T_{n,2}$ is defined in the proof of Lemma 3. Thus according to (33), we have

$$S_{n,2} \stackrel{d}{=} W_{n,\alpha,2}+o_{\mathbb{P}}(1). \tag{49}$$

Term $S_{n,3}$. Combining Lemma 4 with the consistency of $\widehat{\gamma}_{n,k}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}$, we infer that

$$S_{n,3} \stackrel{d}{=} \frac{\gamma+\rho-1}{\rho\gamma}W_{n,\alpha,4}+o_{\mathbb{P}}(1). \tag{50}$$

Term $S_{n,4}$. Under the assumption that $\sqrt{k}\ a(n/k) \to \lambda \in \mathbb{R}$, as $n \to \infty$ and by the consistency of $\widehat{\boldsymbol{\rho}}$ and $\widehat{\gamma}_{n,k}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}$, we have

$$S_{n,4} = o_{\mathbb{P}}(1). \tag{51}$$

Term $S_{n,5}$. Using again the Lemma 4, we get

$$\begin{aligned} S_{n,5} &\overset{d}{=} -\frac{\gamma(1-\rho)}{(1-\gamma)(\gamma+\rho-1)}\sqrt{\frac{n}{k}}\int_0^1 s^{-1}\mathbb{B}_n\left(1-\frac{sk}{n}\right)d(s(K(s)-K_{\widehat{\Delta}^*_{\text{opt}}}(s)))+o_{\mathbb{P}}(1) \\ &= -\frac{(1-\rho)(1-\gamma)}{\gamma+\rho-1}\left(W_{n,\alpha,3}-\frac{\gamma+\rho-1}{\rho\gamma}W_{n,\alpha,4}\right)+o_{\mathbb{P}}(1) \\ &= W_{n,\alpha,5}+\frac{(1-\rho)(1-\gamma)}{\gamma\rho}W_{n,\alpha,4}+o_{\mathbb{P}}(1). \end{aligned} \tag{52}$$

Term $S_{n,6}$. Remark that

$$S_{n,6} = -\frac{\sqrt{k}a(n/k)}{(1-\gamma)(\gamma+\rho-1)} + T_{n,4},$$

where $T_{n,4}$ is defined in the proof of Lemma 3. Thus using (35) and the assumption that $\sqrt{k}\ a(n/k) \to \lambda \in \mathbb{R}$, as $n \to \infty$. We deduce that

$$S_{n,6} = o_{\mathbb{P}}(1). \tag{53}$$

Combining (48)-(53), Lemma 5 follows. ■

5. Simulation study

In this section, the class of biased estimator $\widehat{\eta}_{n,k}^{(K)}(0.2, 0.8)$ and the reduced-bias estimator $\widetilde{\eta}_{n,k,\widehat{\rho}}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}(0.2, 0.8)$ of the QSR index $\eta(Q, 0.2, 0.8)$ are compared in a simulation study. To this end, $N = 500$ samples of size $n := 1000;\ 1500;\ 2000$ are generated from a Burr distribution defined as $\overline{F}(x) = (1 + x^{-\rho/\gamma})^{1/\rho}$, with $\gamma = 2/3$ and different values of $\rho := -0.5;\ -0.75;\ -1$. It is known that this distribution is heavy-tailed and satisfies the second order condition $(\mathcal{R}_{\mathbb{U}})$ with $a(t) = \gamma t^{\rho}$. This kind of Burr distribution and its unidentified parameters were previously used by various authors such as [13], [14] and [11] to assess risk measures for heavy-tailed losses. [29] and [15] also used this kind of distribution to estimate the QSR index for heavy-tailed capital incomes.

Now, for computation and the comparison of the estimators, we adopt the following steps:

- The estimator $\widehat{\eta}_{n,k}^{(K)}(0.2, 0.8)$ is computed with the tail index estimators $\widehat{\gamma}_{n,k}^{(K)}$, for different sample fractional numbers of top order statistics $k = 10, ..., m_n$, where m_n is the integer part of $0.2 \times n$, which ensures the validity of the condition $\beta = 0.8 < 1 - k/n$. For the choice of the kernels K, we use the power kernel, which satisfies the assumption ($\mathbb{K}$) and is defined by $K(s) = (1+\tau)s^{\tau}\mathbb{I}_{\{0<s<1\}}$, with $\tau := 0,\ 1$. In the case where $\tau = 0$, we denote $K := K_1 = \underline{K}$ and $\widehat{\eta}_{n,k}^{(\underline{K})}(0.2, 0.8)$ corresponds to the QSR index estimator associated with the Hill's estimator $\widehat{\gamma}_{n,k}^{(\underline{K})}$. For $\tau = 1$, the corresponding kernel is exactly the above mentioned $K := K_{2,\bar{\rho}}$, with $\bar{\rho} = -1$.

• The estimator $\widetilde{\eta}_{n,k,\widehat{\rho}}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}(0.2, 0.8)$ is computed with the tail index estimators $\widehat{\gamma}_{n,k}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}$, for $k = 10, ..., m_n$ and $\widehat{\rho} := \widehat{\rho}_{k^*_\rho}$ defined in (43), where k^*_ρ is selected as follows:

$$k^*_\rho := \sup\left\{k_\rho :\ k_\rho \leq \min\left(n-1, \frac{2n}{\log\log n}\right) \quad \text{and} \quad \widehat{\rho}_{k_\rho} \quad \text{exists}\right\}.$$

• Next, we compare on the one hand the performance of the mentioned QSR index estimators by computing the absolute value of the median together with the median squared errors (MSE) based on the N samples, and defined as:

$$\text{ABias}(\eta, k) := \left|\text{median}\left\{\frac{\widehat{\eta}^{(1)}}{\eta}, ..., \frac{\widehat{\eta}^{(N)}}{\eta}\right\} - 1\right|$$

and

$$\text{MSE}(\eta, k) := \text{median}\left\{\left(\frac{\widehat{\eta}^{(1)}}{\eta} - 1\right)^2, ..., \left(\frac{\widehat{\eta}^{(N)}}{\eta} - 1\right)^2\right\},$$

where $\eta := \eta(Q, 0.2, 0.8)$ is the true value of the QSR index and $\widehat{\eta}^{(i)}$ is the i-th value ($i = 1, ..., N$) of an estimator of $\eta(Q, 0.2, 0.8)$ evaluated at different sample fractional numbers of top order statistics k as mentioned above.

Figure 1 resp. Figure 2 show the Absolute bias of the median resp. the Median Squared Error of $\widehat{\eta}_{n,k}^{(\underline{K})}(0.2, 0.8)$ (black line), $\widehat{\eta}_{n,k}^{(K_{2,\bar{\rho}})}(0.2, 0.8)$ (blue line) and $\widetilde{\eta}_{n,k,\widehat{\rho}}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}(0.2, 0.8)$ (red line) as a function of k based on $N = 500$ samples of size 1000 (top), 1500 (middle) and 2000 (down) for QSR index $\eta(Q, 0.2, 0.8)$ from the underlying Burr distribution. From the left to the right: ($\rho = -0.5,\ \ \eta(Q, 0.2, 0.8) = 292.93$), ($\rho = -0.75,\ \ \eta(Q, 0.2, 0.8) = 73.47$) and ($\rho = -1,\ \ \eta(Q, 0.2, 0.8) = 37.70$).

To compute the confidence intervals of the estimators under simulation, we need an optimal number of k, whose choice is a serious challenge. The algorithm of [39], Page 137, gives an automatic choice of the number of top extremes k for tail index estimators in $\widehat{\gamma}_{n,k}^{\bullet}$. According to these authors, an automatic choice of top extremes used in $\widehat{\gamma}_{n,k}^{\bullet}$ is as the value k^* that minimizes

$$\frac{1}{k}\sum_{j=1}^{k} j^{\delta}\left|\widehat{\gamma}^{\bullet}_{n,j} - \text{median}\left(\widehat{\gamma}^{\bullet}_{n,1},...,\widehat{\gamma}^{\bullet}_{n,k}\right)\right|,\ 10 \le k \le m_n, \tag{54}$$

where $0 \le \delta < 1/2$. By the way, choosing $\delta = 1/4$, we compute the optimal values k^* as in (54) for each tail index estimator used in the computation of their associated QSR index estimators. In the Table 1, Table 2 and Table 3, we present the results of the estimated values of the above mentioned QSR index estimators with respect to the sample size. Remarking that from Theorem 1 and Theorem 2, the asymptotic variances of the QSR index estimators under study depend on unknown parameters, we opt to use a block bootstrapping method to construct a 95% confidence interval for the QSR index estimates. The block bootstrapping follows the routine boot of the package **boot** in R software. By repeating such bootstrapping procedure $T = 10,000$ times, we obtain T bootstrapped estimates for each QSR index estimator. The sample standard deviation across the T estimates gives an estimate of the standard deviation of the underlying QSR index estimators for given $k \in \{10,...,m_n\}$. We construct the 95% confidence interval using the point estimate and the estimated standard deviation. This procedure is applied to all values of k of each estimator. The point estimates of QSR index at its optimal value k^* as well as the lower and upper bounds of the confidence intervals are given in Table 1, Table 2 and Table 3.

Based on these simulations, we can draw the following conclusions:

• It appears on Figure 2 that the closer ρ is to 0, the more important is the bias of $\widetilde{\eta}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}_{n,k,\widehat{\rho}}(0.2, 0.8)$ with a longer stability as a function of k. The bias is also less variable than the two others for the lowest values of k. Also, The effect of the bias correction on the MSE is well illustrated on Figure 2. We can observe that the MSE of the reduced-bias estimator $\widetilde{\eta}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}_{n,k,\widehat{\rho}}(0.2, 0.8)$ is almost constant with respect to k, especially when bias of $\widehat{\eta}^{(\underline{K})}_{n,k}(0.2, 0.8)$ and $\widehat{\eta}^{(K_{2,\bar{\rho}})}_{n,k}(0.2, 0.8)$ are strong, *i.e*, when ρ is close to 0.

• After the inspection of the tables, two conclusions can be drawn regardless of the situation. First, we notice that the absolute bias of both estimators increases as ρ goes to 0. Second, the reduced bias estimator is more efficient than the biased estimators regardless to the absolute bias, the median squared errors and the cover values when ρ is closer to 0. That illustrates well our conclusions drawn from the graphical analysis.

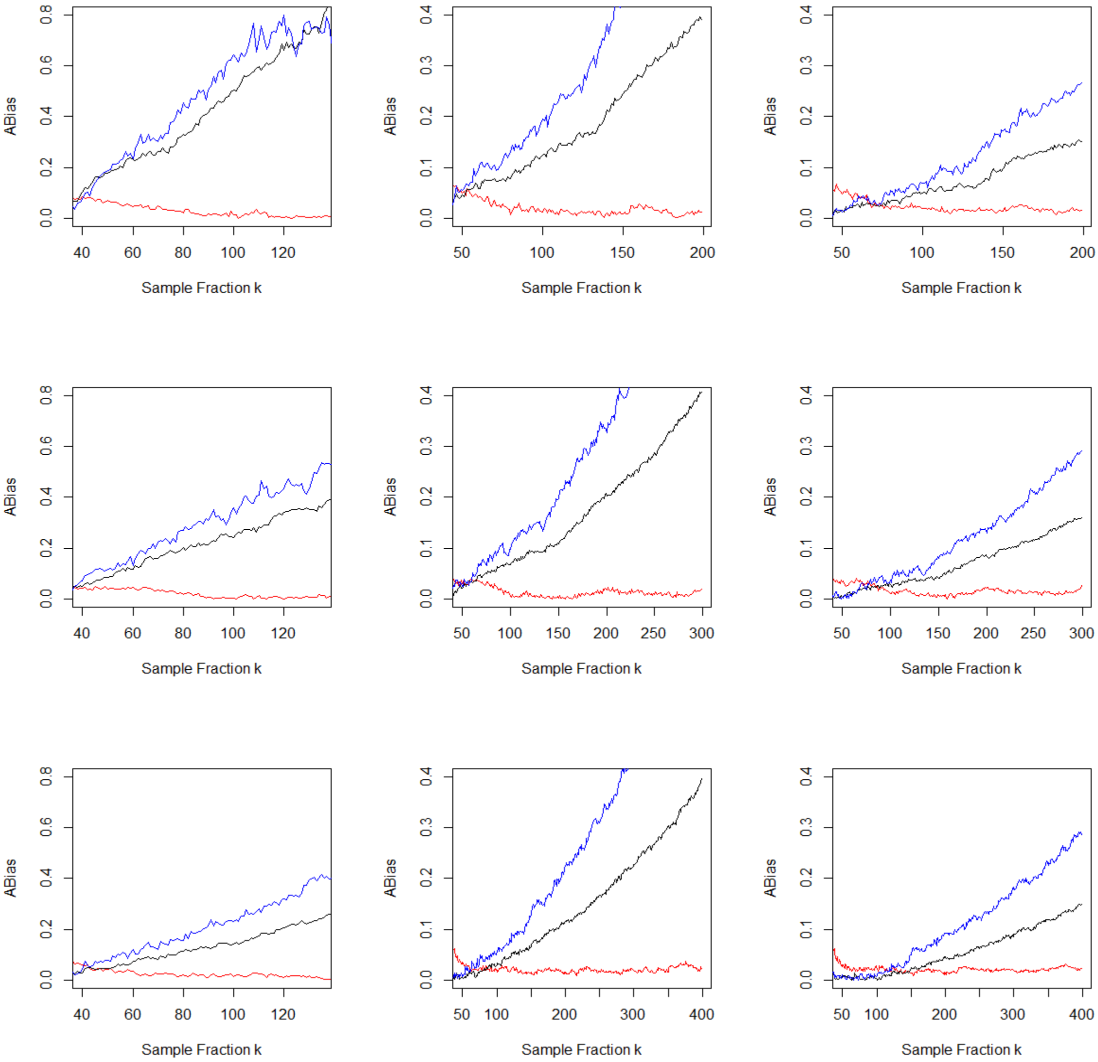

Figure 1: Absolute bias of the median of $\widehat{\eta}^{(K)}_{n,k}(0.2,0.8)$ (black line), $\widehat{\eta}^{(K_2,\bar{\rho})}_{n,k}(0.2,0.8)$ (blue line) and $\widetilde{\eta}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}_{n,k,\widehat{\rho}}(0.2,0.8)$ (red line) as a function of k based on $N = 500$ samples of size 1000 (top), 1500 (middle) and 2000 (down) for QSR index $\eta(Q,0.2,0.8)$ from a Burr distribution defined as $\overline{F}(x) = (1 + x^{-\frac{3\rho}{2}})^{1/\rho}$. From the left to the right: $(\rho = -0.5,\ \ \eta(Q,0.2,0.8) = 292.93)$, $(\rho = -0.75,\ \ \eta(Q,0.2,0.8) = 73.47)$ and $(\rho = -1,\ \ \eta(Q,0.2,0.8) = 37.70)$.

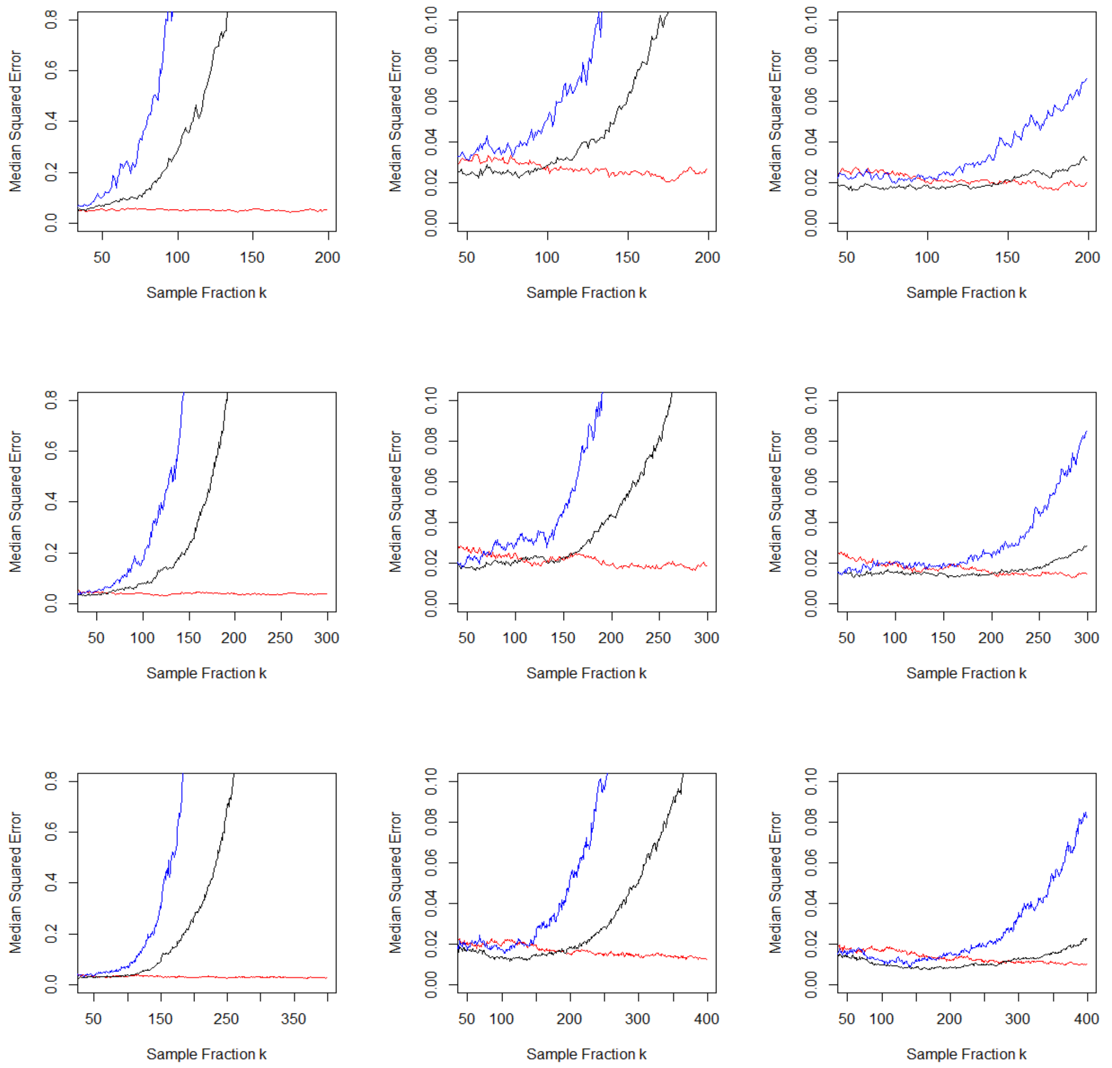

Figure 2: Median Squared Errors (MSE) of $\widehat{\eta}^{(K)}_{n,k}(0.2,0.8)$ (black line), $\widehat{\eta}^{(K_{2,\bar{\rho}})}_{n,k}(0.2,0.8)$ (blue line) and $\widetilde{\eta}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}_{n,k,\widehat{\rho}}(0.2,0.8)$ (red line) as a function of k based on $N = 500$ samples of size 1000 (top), 1500 (middle) and 2000 (down) for QSR index $\eta(Q,0.2,0.8)$ from a Burr distribution defined as $\overline{F}(x) = (1 + x^{-\frac{3\rho}{2}})^{1/\rho}$. From the left to the right: ($\rho = -0.5$, $\eta(Q,0.2,0.8) = 292.93$), ($\rho = -0.75$, $\eta(Q,0.2,0.8) = 73.47$) and ($\rho = -1$, $\eta(Q,0.2,0.8) = 37.70$).

6. Conclusion

In this paper, we introduced a large class of asymptotically normal estimators of the Quintile Share Ratio (QSR) index for heavy tailed income distributions. From that class, we derived a bias reduction procedure and we proposed an unbiased estimator with minimal variance of the

Table 1: Estimation results of the QSR index estimators $\widetilde{\eta}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}_{n,k,\widehat{\rho}}(0.2,0.8)$, $\widehat{\eta}^{(\underline{K})}_{n,k^*}(0.2,0.8)$ and $\widehat{\eta}^{(K_{2,\bar{\rho}})}_{n,k^*}(0.2,0.8)$ with their 95% confidence intervals, computed with their associated optimal numbers of top statistics k^*, based on $N=500$ samples of size $n=1000$, from a Burr distribution defined as $\overline{F}(x)=(1+x^{-\frac{3\rho}{2}})^{1/\rho}$. The true values of the QSR index are $\eta(Q,0.2,0.8)=292.93$ for $\rho=-0.5$, $\eta(Q,0.2,0.8)=73.47$, for $\rho=-0.75$ and $\eta(Q,0.2,0.8)=37.70$ for $\rho=-1$.

$n=1000$								
ρ	γ-estimates		QSR-estimates		Abias	MSE	95%-Conf. Int	Cover
-0.5	$\widehat{\gamma}^{(\underline{K})}_{n,k^*}$	0.831	$\widehat{\eta}^{(\underline{K})}_{n,k^*}(0.2,0.8)$	396.559	0.353	0.165	(287.176; 698.188)	411.012
	$\widehat{\gamma}^{(K_{2,\bar{\rho}})}_{n,k^*}$	0.864	$\widehat{\eta}^{(K_{2,\bar{\rho}})}_{n,k^*}(0.2,0.8)$	404.973	0.382	0.367	(280.834; 712.542)	431.708
	$\widehat{\gamma}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}_{n,k^*}$	0.733	$\widetilde{\eta}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}_{n,k^*,\widehat{\rho}}(0.2,0.8)$	297.701	0.016	0.055	(257.609; 310.992)	53.382
-0.75	$\widehat{\gamma}^{(\underline{K})}_{n,k^*}$	0.749	$\widehat{\eta}^{(\underline{K})}_{n,k^*}(0.2,0.8)$	84	0.143	0.032	(70.909; 104.751)	33.841
	$\widehat{\gamma}^{(K_{2,\bar{\rho}})}_{n,k^*}$	0.768	$\widehat{\eta}^{(K_{2,\bar{\rho}})}_{n,k^*}(0.2,0.8)$	89.821	0.222	0.059	(69.444; 131.834)	62.390
	$\widehat{\gamma}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}_{n,k^*}$	0.687	$\widetilde{\eta}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}_{n,k^*,\widehat{\rho}}(0.2,0.8)$	75.140	0.009	0.028	(66.043; 75.648)	9.604
-1	$\widehat{\gamma}^{(\underline{K})}_{n,k^*}$	0.736	$\widehat{\eta}^{(\underline{K})}_{n,k^*}(0.2,0.8)$	42.859	0.136	0.024	(36.466; 44.092)	7.626
	$\widehat{\gamma}^{(K_{2,\bar{\rho}})}_{n,k^*}$	0.754	$\widehat{\eta}^{(K_{2,\bar{\rho}})}_{n,k^*}(0.2,0.8)$	46.517	0.203	0.056	(36.106; 49.632)	13.526
	$\widehat{\gamma}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}_{n,k^*}$	0.672	$\widetilde{\eta}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}_{n,k^*,\widehat{\rho}}(0.2,0.8)$	37.810	0.002	0.018	(34.252; 38.446)	4.193

Table 2: Estimation results of the QSR index estimators $\widetilde{\eta}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}_{n,k,\widehat{\rho}}(0.2,0.8)$, $\widehat{\eta}^{(\underline{K})}_{n,k^*}(0.2,0.8)$ and $\widehat{\eta}^{(K_{2,\bar{\rho}})}_{n,k^*}(0.2,0.8)$ with their 95% confidence intervals, computed with their associated optimal numbers of top statistics k^*, based on $N=500$ samples of size $n=1500$, from a Burr distribution defined as $\overline{F}(x)=(1+x^{-\frac{3\rho}{2}})^{1/\rho}$. The true values of the QSR index are $\eta(Q,0.2,0.8)=292.93$ for $\rho=-0.5$, $\eta(Q,0.2,0.8)=73.47$, for $\rho=-0.75$ and $\eta(Q,0.2,0.8)=37.70$ for $\rho=-1$.

$n=1500$								
ρ	γ-estimates		QSR-estimates		Abias	MSE	95%-Conf. Int	Cover
-0.5	$\widehat{\gamma}^{(\underline{K})}_{n,k^*}$	0.817	$\widehat{\eta}^{(\underline{K})}_{n,k^*}(0.2,0.8)$	380.375	0.298	0.108	(284.818; 667.245)	382.427
	$\widehat{\gamma}^{(K_{2,\bar{\rho}})}_{n,k^*}$	0.850	$\widehat{\eta}^{(K_{2,\bar{\rho}})}_{n,k^*}(0.2,0.8)$	392.540	0.369	0.267	(281.455; 702.841)	421.386
	$\widehat{\gamma}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}_{n,k^*}$	0.717	$\widetilde{\eta}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}_{n,k^*,\widehat{\rho}}(0.2,0.8)$	295.998	0.010	0.038	(258.821; 303.800)	44.978
-0.75	$\widehat{\gamma}^{(\underline{K})}_{n,k^*}$	0.735	$\widehat{\eta}^{(\underline{K})}_{n,k^*}(0.2,0.8)$	81.594	0.132	0.029	(69.898, 100.133)	30.235
	$\widehat{\gamma}^{(K_{2,\bar{\rho}})}_{n,k^*}$	0.762	$\widehat{\eta}^{(K_{2,\bar{\rho}})}_{n,k^*}(0.2,0.8)$	86.718	0.213	0.038	(69.231; 124.037)	54.806
	$\widehat{\gamma}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}_{n,k^*}$	0.674	$\widetilde{\eta}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}_{n,k^*,\widehat{\rho}}(0.2,0.8)$	73.599	0.006	0.018	(66.998; 73.983)	6.985
-1	$\widehat{\gamma}^{(\underline{K})}_{n,k^*}$	0.728	$\widehat{\eta}^{(\underline{K})}_{n,k^*}(0.2,0.8)$	42.206	0.119	0.018	(37.105; 43.399)	6.293
	$\widehat{\gamma}^{(K_{2,\bar{\rho}})}_{n,k^*}$	0.749	$\widehat{\eta}^{(K_{2,\bar{\rho}})}_{n,k^*}(0.2,0.8)$	45.521	0.198	0.038	(36.469; 47.377)	10.907
	$\widehat{\gamma}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}_{n,k^*}$	0.669	$\widetilde{\eta}^{(K_{\widehat{\Delta}^*_{\text{opt}}})}_{n,k^*,\widehat{\rho}}(0.2,0.8)$	37.369	0.001	0.014	(34.357, 37.751)	3.394

QSR index. Comparing the bias reduction procedure to the alternative estimators, our unbiased estimator provides, in addition to lower absolute bias and median squared error in general, more stability over the number of top statistics k, especially when bias of the alternative estimators

Table 3: Estimation results of the QSR index estimators $\widetilde{\eta}_{n,k,\widehat{\rho}}^{(K_{\widehat{\Delta}^*_{\mathrm{opt}}})}(0.2,0.8)$, $\widehat{\eta}_{n,k^*}^{(\underline{K})}(0.2,0.8)$ and $\widehat{\eta}_{n,k^*}^{(K_{2,\bar{\rho}})}(0.2,0.8)$ with their 95% confidence intervals, computed with their associated optimal numbers of top statistics k^*, based on $N=500$ samples of size $n=2000$, from a Burr distribution defined as $\overline{F}(x)=(1+x^{-\frac{3\rho}{2}})^{1/\rho}$. The true values of the QSR index are $\eta(Q,0.2,0.8)=292.93$ for $\rho=-0.5$, $\eta(Q,0.2,0.8)=73.47$, for $\rho=-0.75$ and $\eta(Q,0.2,0.8)=37.70$ for $\rho=-1$.

$n=2000$								
ρ	γ-estimates		QSR-estimates		Abias	MSE	95%-Conf. Int	Cover
-0.5	$\widehat{\gamma}_{n,k^*}^{(\underline{K})}$	0.782	$\widehat{\eta}_{n,k^*}^{(\underline{K})}(0.2,0.8)$	352.200	0.223	0.102	(286.512; 574.367)	287.855
	$\widehat{\gamma}_{n,k^*}^{(K_{2,\bar{\rho}})}$	0.804	$\widehat{\eta}_{n,k^*}^{(K_{2,\bar{\rho}})}(0.2,0.8)$	371.115	0.279	0.212	(282.517; 621.934)	339.417
	$\widehat{\gamma}_{n,k^*}^{(K_{\widehat{\Delta}^*_{\mathrm{opt}}})}$	0.692	$\widetilde{\eta}_{n,k^*,\widehat{\rho}}^{(K_{\widehat{\Delta}^*_{\mathrm{opt}}})}(0.2,0.8)$	293.616	0.008	0.018	(271.012; 297.014)	26.002
-0.75	$\widehat{\gamma}_{n,k^*}^{(\underline{K})}$	0.727	$\widehat{\eta}_{n,k^*}^{(\underline{K})}(0.2,0.8)$	79.682	0.117	0.012	(71.515; 101.509)	29.994
	$\widehat{\gamma}_{n,k^*}^{(K_{2,\bar{\rho}})}$	0.758	$\widehat{\eta}_{n,k^*}^{(K_{2,\bar{\rho}})}(0.2,0.8)$	82.763	0.185	0.024	(70.208; 130.772)	60.563
	$\widehat{\gamma}_{n,k^*}^{(K_{\widehat{\Delta}^*_{\mathrm{opt}}})}$	0.668	$\widetilde{\eta}_{n,k^*,\widehat{\rho}}^{(K_{\widehat{\Delta}^*_{\mathrm{opt}}})}(0.2,0.8)$	73.416	0.0007	0.010	(68.456; 73.943)	5.486
-1	$\widehat{\gamma}_{n,k^*}^{(\underline{K})}$	0.714	$\widehat{\eta}_{n,k^*}^{(\underline{K})}(0.2,0.8)$	40.918	0.085	0.009	(36.921; 43.617)	6.695
	$\widehat{\gamma}_{n,k^*}^{(K_{2,\bar{\rho}})}$	0.731	$\widehat{\eta}_{n,k^*}^{(K_{2,\bar{\rho}})}(0.2,0.8)$	43.246	0.146	0.019	(35.998; 48.017)	12.018
	$\widehat{\gamma}_{n,k^*}^{(K_{\widehat{\Delta}^*_{\mathrm{opt}}})}$	0.657	$\widetilde{\eta}_{n,k^*,\widehat{\rho}}^{(K_{\widehat{\Delta}^*_{\mathrm{opt}}})}(0.2,0.8)$	37.461	0.0005	0.0007	(35.149; 37.706)	2.557

are strong. The comparison are also made at their optimal point of top statistics and with their 95% confidence intervals, constructed from a Bootstrap methodology. The results show that, the reduced bias estimator is more efficient than alternative estimators regardless to the absolute bias, the median squared errors and the coverage. An important feature expected in this type of of bias reduction approach to be applicable in practice. In application, the unbiased estimator can be proposed to any heavy-tailed income distributions for which QSR index needs to be calculated.

Acknowledgements

The authors would like to thank thank the Reviewers and editors for their valuable comments and suggestions that helped improve considerably the paper.

References

[1] Andrew B. Abel. Optimal capital income taxation. Technical Report 13354, National Bureau of Economic Research, 2007.

[2] Ahmed Bari, Adnene Rassoul, and Hatem Ouarda Rouis. Estimating the gini index for heavy-tailed distributions. *South African Statistical Journal*, 55(1):15–28, 2021.

[3] Jan Beirlant, Goedele Dierckx, Yuri Goegebeur, and Gert Matthys. Tail index estimation and an exponential regression model. *Extremes*, 2(2):177–200, 1999.

[4] Jan Beirlant, Goedele Dierckx, Armelle Guillou, and Catalin Starica. On exponential representations of log-spacings of extreme order statistics. *Extremes*, 5(2):157–180, 2002.

[5] N. H. Bingham, C.M. Goldie, and J.L. Teugels. *Regular variation.* Cambridge University Press, 1987.

[6] Christophe Chamley. Optimal taxation of capital income in general equilibrium with infinite lives. *Econometrica*, 54(3):607–622, 1986.

[7] Alex Cobham and Andy Sumner. Is it all about the tails? the palma measure of income inequality. Available at the address http://www.ecineq.org/milano/WP/ECINEQ2013-308.pdf, 2013.

[8] M. Csörgö, S. Csörgö, L. Horváth, and D.M. Mason. Weighted empirical and quantile processes. *Annals of Probability*, 14(1):31–85, 1986.

[9] S. Csörgő, P. Deheuvels, and D.M. Mason. Kernel estimates of the tail index of a distribution. *Annals of Statistics*, 13(3):1050–1077, 1985.

[10] L. de Haan and A. Ferreira. *Extreme value theory: an introduction.* Springer, 2006.

[11] E. H. Deme, M. Allaya, S. Deme, A. Dhaker, and A.S. Dabye. Estimation of risk measures from heavy-tailed distributions. *East Journal of Theoretical Statistics*, 62(1):35–80, 2021.

[12] E. H. Deme, L. Gardes, and S. Girard. On the estimation of the second order parameter for heavy-tailed distributions. *REVSTAT - Statistical Journal*, 11(3):277–299, 2013.

[13] E. H. Deme, S. Girard, and A. Guillou. Reduced-bias estimator of the proportional hazard premium for heavy-tailed distributions. *Insurance Mathematics & Economics*, 52(3):550–559, 2013.

[14] E. H. Deme, S. Girard, and A. Guillou. Reduced-biased estimators of the conditional tail expectation for heavy-tailed distributions. In *Mathematical Statistics and Limit Theorems*, pages 105–123. Springer, 2015.

[15] E.H. Deme, T. Kpanzou, and E. Sisawo. Semi-parametric estimation of the quintile share ratio index of inequality measure for heavy-tailed income distributions with index in the upper half of the unit interval. *Afrika Statistika*, 17(1):3095–3114, 2022.

[16] Andrey Feuerverger and Peter Hall. Estimating a tail exponent by modelling departure from a pareto distribution. *Annals of Statistics*, 27(2):760–781, 1999.

[17] Anna Fräßdorf, Markus M. Grabka, and Johannes Schwarze. The impact of household capital income on income inequality: A factor decomposition analysis for the uk, germany and the usa. *Journal of Economic Inequality*, 9(1):35–56, 2011.

[18] Jan L. Geluk and Laurens de Haan. *Regular Variation, Extensions and Tauberian Theorems.* Number 40 in CWI Tract. Center for Mathematics and Computer Science, Amsterdam, The Netherlands, 1987.

[19] Yuri Goegebeur and Guillou Armelle. Asymptotically unbiased estimation of the coefficient of tail dependence. *Scandinavian Journal of Statistics*, 40:174–189, 2013.

[20] Mikhail Golosov, Narayana R Kocherlakota, and Aleh Tsyvinski. Optimal indirect and capital taxation. *Review of Economic Studies*, 70(3):569–587, 2003.

[21] M. Ivette Gomes, Francisco Figueiredo, and Sandra Mendonca. Asymptotically best linear unbiased tail estimators under a second-order regular variation condition. *Journal of Statistical Planning and Inference*, 134(2):409–433, 2005.

[22] Francesca Greselin and Leo Pasquazzi. Asymptotic confidence intervals for a new inequality measure. *Communications in Statistics - Computational and Simulation*, 38(1):17–42, 2009.

[23] Francesca Greselin, Leo Pasquazzi, and Ricardas Zitikis. Zenga's new index of economic inequality, its estimation, and an analysis of incomes in italy. *Journal of Probability and Statistics*, 2010(Article ID 718905):26, 2010.

[24] Francesca Greselin, Leo Pasquazzi, and Ricardas Zitikis. Contrasting the gini and zenga indices of economic inequality. *Journal of Applied Statistics*, 40(2):282–297, 2013.

[25] Francesca Greselin, Leo Pasquazzi, and Ricardas Zitikis. Heavy tailed capital incomes: Zenga index, statistical inference, and echp data analysis. *Extremes*, 17(1):127–155, 2014.

[26] Armelle Guillou and Val'erie Chavez-Demoulin. Extreme quantile estimation for β-mixing time series and applications. *Insurance: Mathematics and Economics*, 80:59–74, 2018.

[27] Bruce M Hill. A simple approach to inference about the tail of a distribution. *Annals of Statistics*, pages 1136–1174, 1975.

[28] Kenneth L Judd. Capital-income taxation with imperfect competition. *American Economic Review*, 92(2):417–421, 2002.

[29] Tchilabalo A Kpanzou. Asymptotic distribution of the quintile share ratio estimator. *Afrika Statistika*, 9(1):659–670, 2014.

[30] Tchilabalo A Kpanzou. On the influence function of the quintile share ratio. *Communications in Statistics-Simulation and Computation*, 44(8):2492–2499, 2015.

[31] M. Langel and Y. Tillé. Variance estimation of the gini index: revisiting a result several times published. *Journal of the Royal Statistical Society*, 176:521–540, 2013.

[32] Matti Langel and Yves Tillé. Statistical inference for the quintile share ratio. *Journal of Statistical Planning and Inference*, 141(10):2976–2985, 2011.

[33] G.S. Lo and A. M. Fall. Another look at second order condition in extreme value theory. *Afrika Statistika*, 6:346–370, 2011.

[34] A. Necir, A. Rassoul, and R. Zitikis. Estimating the conditional tail expectation in the case of heavy-tailed losses. *Journal of Probability and Statistics*, ID 596839:17 pages, 2010.

[35] G. Osier. Variance estimation: the linearization approach applied by eurostat to the 2004 silc operation. *Technical Report, Eurostat and Statistics Finland Methodological Workshop on EU-SILC*, 2006.

[36] G. Osier. Variance estimation for complex indicators of poverty and inequality using linearization techniques. *Survey Research Methods*, 3:167–195, 2009.

[37] L. Peng. Estimating the mean of a heavy tailed distribution. *Stat. Prob. Lett.*, 52:255–264, 2001.

[38] L. Peng and Y. Qi. Estimating the first- and second-order parameters of a heavy-tailed distribution. *Aust. N. Z. J. Stat.*, 46(2):305–312, 2004.

[39] R. D. Reiss and M. Thomas. *Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields.* Birkhäuser, 3rd edition, 2007.

[40] P. B. Sørensen. Can capital income taxes survive? and should they? *CESifo Econ. Stud.*, 53(2):172–228, 2007.

[41] O. Tami, A. Rassoul, and Ould R.H. An improved estimator of the zenga index for heavy-tailed distributions. *J. Stat. Appl. Pro.*, 8(2):91–102, 2019.

[42] I. Weissman. Estimation of parameters and larges quantiles based on the k largest observations. *Journal of American Statistical Association*, 73(364):812–815, 1978.

[43] K. Xu. How has the literature on gini's index evolved in the past 80 years? Technical report, Department of Economics, Dalhousie University, 2004.

[44] M. Zenga. Inequality curve and inequality index based on the ratios between lower and upper arithmetic means. *Stat. Appl.*, 5:3–27, 2007.

[45] R. Zitikis. The vervaat process. In *Asymptotic Methods in Probability and Statistics*, page 667–694. North-Holland, 1998.