Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.18576/jsap/080203

An Improved Estimator of the Zenga Index for Heavy-Tailed Distributions

Tami Omar[1], *Rassoul Abdelaziz*[2,*] *and Ould Rouis Hamid* [1]

[1] LAMDARO Laboratory, Departement of Mathematics, University of Blida 1, Algeria
[2] GEE Laboratory, Ecole Nationale Superieure dHydraulique, Blida, Algeria

Received: 17 Feb. 2019, Revised: 06 May 2019, Accepted: 22 May 2019
Published online: 1 Jul. 2019

Abstract: In the present paper, we focus on the Zenga index, the asymptotic normality of the classical estimators has been established in the literature under the classical assumption that the second moment of the loss variable is finite, this condition is very restrictive in practical applications. Such a result has been extended by Greselin et al. (2014) [31] in the case of distributions with infinite second moment. Thus, we base on this framework and propose a reduced-bias estimator for the classical estimators. Finally, we illustrate the efficiency of our approach by some results on a simulation study and compare its performance with other estimators.

Keywords: Zenga index, Heavy-tailed distribution, kernel-type estimator, Hill estimator, extreme quantile, reduced-bias

1 Introduction and motivation

Inequality measurement is an attempt to give meaning to comparisons of income distributions in terms of criteria which may be derived from ethical principles, appealing mathematical constructs or simple intuition. A serious approach to inequality measurement should begin with a consideration of the entities to which the tools of distributional judgment are applied. In the recent decades, economic thinking about the income shows that, the capital income is among the major incomes in a big number of countries, there are many important econometrics devoted to the income capital, 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 economic scenarios, for further details consult the following references (cf., e.g., Chamley, (1986)[5]; Judd, (2002) [34]; Sörensen, (2007) [40]; Abel, (2007) [1]; Golosov et al., (2003)[27], Greselin, et al. (2014) [31]).
For more details about the analysis of the impact of the capital incomes on the income inequality we refer to the following references (cf., e.g., Lerman and Yitzhaki, (1985)[35]; Saez, (2005) [41]). In recent years, micro-data shows that capital revenues are extremely volatile, and that their share of disposable income has increased. Moreover, in some countries, capital incomes have been making up an insufficient income for high contribution to the overall inequality (cf., e.g., Frädorf et al., (2011) [18]).

The present work has been motivated by the need for better understanding of the distribution and inequality of capital incomes, which in many cases appear to be heavy tailed. We focus on capital incomes which are income flows from financial assets actually received during the reference year. Capital gains and losses are thus excluded from the definition. Since there are many individuals with no capital incomes, we restrict our attention to only those with positive capital incomes.
The rest of the paper is organized as follows. In Section 2 we formulate the main quantities of interest: the Zenga curve and index, their empirical estimators, and we formulate the kernel-type estimator. In Section 3 we present the main results of statistical inference for the kernel-type estimators of the Zenga index and the construction of a reduced-bias estimator and formulate the different results of its asymptotic behaviours. In Section 4 we illustrate and show the performance of our results by some results of simulation study. The proof of different results are presented in Section 5.

* Corresponding author e-mail: a.rassoul@ensh.dz

Note that throughout this paper, the standard notations $\overset{\mathbb{P}}{\to}$, $\overset{d}{\to}$ and $\overset{d}{=}$ respectively stand for convergence in probability, convergence in distribution and equality in distribution, $\mathcal{N}(a,b^2)$ denotes the normal distribution with mean a and variance b^2, and $\{\mathbf{B}_n(s), 0<s<1\}_{n\geq 1}$ is a sequence of standard Brownian bridges.

2 Zenga curve and index: Definitions

M. Zenga (2007) [43] proposed a measure of inequality index which aggregates the ratios of lower and upper conditional tail expectations, and by doing so it takes into account the relative nature of the poor and the rich. Namely, let F denote the cumulative distribution function (cdf) of the population represented by a non-negative random variable $X \geq 0$. We assume that F is continuous and strictly increasing. The (generalized) inverse $\mathbb{Q}:(0,1)\to[0,\infty)$ of the cdf F, known in the literature as the quantile function, is defined for all $t\in(0,1)$ by the formula $\mathbb{Q}(t)=\inf\{x: F(x)\geq t\}$.

The upper and lower conditional tail expectations are $\mathbb{E}[X|X>\mathbb{Q}(t)]$ and $\mathbb{E}[X|X\leq\mathbb{Q}(t)]$, respectively. Since F is continuous, they coincide with the upper and lower Tail-Values-at-Risk (TVaR) (e.g., Denuit et al., (2005) [14]):

$$TVaR_F(t)=\frac{1}{1-t}\int_t^1 \mathbb{Q}(s)ds \text{ and } TVaR_F^*(t)=\frac{1}{t}\int_0^t \mathbb{Q}(s)ds$$

respectively. M. Zenga (2007) [43] defined the index Z_F of (relative) inequality and the corresponding curve, which are given by, respectively,

$$Z_F=\int_0^1 z_F(s)ds \text{ where } z_F(s)=1-\frac{TVaR_F^*(t)}{TVaR_F(t)}.$$

To interpret that, we first note that the Zenga curve $z_F(s)$ measures the inequality between the poorest $t\times 100\%$ part of the population and the remaining richer $(1-t)\times 100\%$ part. Note also that $z_F(t)\in[0,1]$. Hence, the Zenga index Z_F takes on values closer to 1 when, averaging over all $t\in[0,1]$, the mean income of the poorest $t\times 100\%$ sub-population is small compared to the mean of the remaining richer sub-population. On the other hand, the index Z_F takes on values closer to 0 when the difference between the mean incomes of the aforementioned two sub-populations is small, in average. For additional thoughts on the topic, we refer to Zenga, (2007) [43], Greselin et al., (2013) [30], and references therein.
To compare the Zenga index and curve with the classical Gini index and curve, we write the latter pair as follows:

$$G_F(t)=\int_0^1 g_F(t)(2t)dt \text{ and } g_F(t)=1-\frac{TVaR_F^*(t)}{TVaR_F(0)}.$$

Now, we contrast the definitions of the curves $z_F(t)$ and $g_F(t)$, we see that while $z_F(t)$ compares the mean incomes of two disjoint sub-populations, the poor and the rich, the Gini curve $g_F(t)$ compares overlapping parts of the population. Furthermore, we see that, due to the weighting function $2t$, the Gini index underestimates the comparisons measured by $g_F(t)$ between the very poor and the whole population and emphasizes those comparisons that involve almost identical population subgroups.
The Zenga index, hence, detects with the same sensibility all deviations from equality in any part of the distribution, measured by $z_F(t)$. For more details on contrasting the Zenga and Gini indices, we refer to Greselin et al. (2013) [30] and references therein.

2.1 Empirical estimator of the Zenga index

Replacing the population cdf F by its empirical counterpart F_n, where $F_n(x)=\frac{1}{n}\sum_{i=1}^n 1_{\{X_i\leq x\}}$, where $1_C(t)$ denotes the indicator function ($1_C(t)=1$ if $t\in C$, and equal to 0 otherwise), and $X_1,X_2,...,X_n$ are independent and identically distributed (i.i.d.) observations drawn from the distribution function F, we arrive at the 'traditional' Zenga estimator (cf., e.g., Greselin and Pasquazzi, (2009) [28] and Greselin et al., (2013) [30]

$$\widehat{Z}_n=1-\frac{1}{n-1}\sum_{j=1}^{n-1}\frac{j^{-1}\sum_{i=1}^j X_{i,n}}{(n-j)^{-1}\sum_{i=j+1}^n X_{i,n}}, \tag{1}$$

where $X_{1,n}\leq X_{2,n}\leq ...\leq X_{n,n}$ are the order statistics of $X_1,X_2,...,X_n$. Note that the estimator is neither a ratio of L-statistics nor a ratio of U-statistics. For this reason, Greselin and Pasquazzi (2009) [28] and Greselin et al. (2010) [29] have developed an asymptotic theory for the traditional Zenga estimator, assuming that the underlying i.i.d. random variables $X_1,X_2,...,X_n$ have finite $(2+\varepsilon)$ moments for some $\varepsilon>0$ as small as desired.

2.2 Kernel-type estimators

The latter moment assumption plays a crucial role. To illustrate the performance of $\widehat{Z}_n$, we choose the Pareto distribution

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

for some $\gamma>0$, which is called the tail index.

When $\gamma>1$, then $TVaRF(t)$, thus $z_F(t)$ and Z_F are not defined.
When $\gamma<0.5$, then $E[X^{2+\varepsilon}]<\infty$ for some $\varepsilon>0$, and so we can use the available estimator of Z_F.
When $\gamma\in(0.5,1)$ the second moment is infinite, and so we cannot rely on the estimator (1). Greselin, et al. (2014) [30] motivated the need of a specific estimator for heavy-tailed distributions, motivated that the traditional estimator $\widehat{Z}_n$ suffers badly from undercoverage by drawing samples from Pareto model with index $\gamma=0.75$.

In the next part of this paper, we consider a class of heavy-tailed distributions, we assume that

$$\lim_{t\to\infty}\frac{1-F(tx)}{1-F(x)}=x^{-1/\gamma},\ x>0. \tag{3}$$

We restrict ourself to the case $\gamma\in(0.5,1)$, hence, we indeed need another estimator in the case of heavy-tailed populations. The estimation of γ has been extensively studied in the literature and the most famous estimator is the Hill (1975) [33] estimator defined as

$$\widehat{\gamma}^{H}_{n,k}=\frac{1}{k}\sum_{i=1}^{k}\log X_{n-i+1,n}-\log X_{n-k,n}, \tag{4}$$

or, equivalently

$$\widehat{\gamma}^{H}_{n,k}=\frac{1}{k}\sum_{i=1}^{k}i\left(\log X_{n-i+1,n}-\log X_{n-i,n}\right), \tag{5}$$

for an intermediate sequence k, i.e. a sequence such that $k\to\infty$ and $k/n\to 0$ as $n\to\infty$.
More generally, Csörgő et al. (1985) [9] extended the Hill estimator (5) into a kernel class of estimators as follows:

$$\widehat{\gamma}^{K}_{n,k}=\frac{1}{k}\sum_{i=1}^{k}K\left(\frac{i}{k+1}\right)Y_i, \tag{6}$$

where K is a kernel integrating to one and $Y_i=i\left(\log X_{n-i+1,n}-\log X_{n-i,n}\right)$.
Note that the Hill estimator corresponds to the particular case where $K=\underline{K}:=1_{(0,1)}$.

In this spirit, we can construct a new estimator of Z_F for a heavy-tailed distribution satisfying the condition (3) as follows:

$$\widetilde{Z}^{K}_{n,k}=1-\int_0^1\left(\frac{\widetilde{TVaR}^{*}_{n,k}(t)}{\widetilde{TVaR}_{n,k}(t)}\right)dt, \tag{7}$$

where $\widetilde{TVaR}_{n,k}(t)$ is a semi-parametric estimator of $TVaR_{n,k}(t)$, we can rewrite $TVaR_{n,k}(t)$ as follows:

$$\begin{aligned}TVaR_{n,k}(t) &= \frac{1}{(1-t)}\int_t^{1-k/n}\mathbb{Q}(s)ds+\frac{1}{(1-t)}\int_0^{k/n}\mathbb{Q}(1-s)ds\\ &= TVaR^{(1)}_{n,k}(t)+TVaR^{(2)}_{n,k}(t),\end{aligned}$$

then, an estimator of $TVaR_{n,k}(t)$ is defined by the formula:

$$\widetilde{TVaR}_{n,k}(t) = \widetilde{TVaR}^{(1)}_{n,k}(t)+\widetilde{TVaR}^{(2)}_{n,k}(t) \tag{8}$$

$$= \frac{1}{(1-t)}\int_t^{1-k/n}\mathbb{Q}_n(s)ds+\frac{(k/n)X_{n-k,n}}{(1-t)\left(1-\widehat{\gamma}^{K}_{n,k}\right)}, \tag{9}$$

where $\mathbb{Q}_n(s)$ is a trimmed empirical estimaor of the quantile for $s \in (t, 1-k/n)$. To estimate $TVaR^{(2)}_{n,k}(t)$ we use an extreme quantile, known by Weissman-type estimator [42] for $\mathbb{Q}$, such that:

$$\hat{\mathbb{Q}}^W_n(1-s) := X_{n-k,n}\ (k/n)^{\hat{\gamma}^K_{n,k}}\ s^{-\hat{\gamma}^K_{n,k}},\ s \to 0. \tag{10}$$

where $\hat{\gamma}^K_{n,k}$ is the kernel class of Hill estimator of the tail index γ.

Also, we define an estimator for $TVaR^*_F$ as the following:

$$\widetilde{TVaR}^*_{n,k} = \frac{1}{t}\left(\widetilde{TVaR}_{n,k}(0) + (1-t)\widetilde{TVaR}_{n,k}(t)\right).$$

Thus, the estimator (7) is already proposed by Greselin et al., 2014 [31] in the particular case where $K = \underline{K} := 1_{(0,1)}$. Asymptotic normality for $\widetilde{Z}^K_{n,k}$ is obviously related to the one $\hat{\gamma}^K_{n,k}$ estimator. As usual in the extreme value framework, to prove such type of results, we need a second-order condition on the tail quantile function $\mathbb{U}$, defined as

$$\mathbb{U}(z) = \inf\{y : F(y) \geq 1 - 1/z\},\ z > 1. \tag{11}$$

We say that the function $\mathbb{U}$ satisfies the second-order regular variation condition with second-order parameter $\rho \leq 0$ if there exists a function $A(t)$ which does not change its sign in a neighbourhood of infinity and such that, for every $x > 0$,

$$\lim_{t\to\infty} \frac{\log \mathbb{U}(tx) - \log \mathbb{U}(t) - \gamma \log(x)}{A(t)} = \frac{x^\rho - 1}{\rho}, \tag{12}$$

when $\rho = 0$, then the ratio on the right-hand side of equation (12) should be interpreted as $\log x$. As an example of heavy-tailed distributions satisfying the second-order condition, we have the so called and frequently used Hall's model which is a class of cdf's, such that

$$\mathbb{U}(t) = ct^\gamma(1 + dA(t)/\rho + o(t^\rho)) \text{ as } t \to \infty. \tag{13}$$

where $\gamma > 0$, $\rho \leq 0$, $c > 0$, and $d \in \mathbb{R}^*$.
This sub-class of heavy-tailed distributions contains the Pareto, Burr, Fréchet and $t-$Student, these cdf's are usually used in economic and insurance mathematics, as models for dangerous risks. For statistical inference concerning the second-order parameter ρ we refer, for example, to Peng and Qi (2004) [38], Gomes *et al.,* (2005) [24], Gomes and Pestana (2007) [25].
The study of the asymptotic distributions of the estimator given by equation (7) for the uniform kernel $K = \underline{K} := 1_{(0,1)}$, is established by Greselin et al., (2014) [31] under the second-order framework (12), which is asymptotically normal with null mean value whenever $\sqrt{k}A(n/k) \to 0$, but there appears a non-null asymptotic bias, whenever $\sqrt{k}A(n/k) \to \lambda$is finite. In this paper we are going to base on reduced bias estimators of the Zenga index, even when $\sqrt{k}A(n/k) \to \lambda$ is finite, non-necessarily Null. Our procedure is based on the exponential regression model.

3 Main results

To study the asymptotic normality of the estimator of $\widetilde{Z}^K_{n,k}$, we need some results and classical assumptions about the kernel:

Condition ($\mathscr{K}$): Let K be a function defined on $(0,1]$

CK1.$K(s) \geq 0$ whenever $0 < s \leq 1$ and $K(1) = K'(1) = 0$;
CK2.$K(\cdot)$ is differentiable, nonincreasing and right continuous on $(0,1]$;
CK3.K and K' are bounded (K' is the derivative function of K);
CK4.$\int_0^1 K(u)du = 1$;
CK5.$\int_0^1 u^{-1/2}K(u)du < \infty$.

3.1 Asymptotic result for the $\widetilde{Z}^K_{n,k}$ estimator

Theorem 1.*Assume that F satisfies the condition (12) for $\gamma \in (1/2,1)$. If further ($\mathscr{K}$) holds and the sequence k satisfies $k \to \infty$ and $k/n \to 0$ as $n \to \infty$, we have*

$$\frac{\sqrt{k}}{(k/n)\mathbb{U}(n/k)}\left(\widetilde{Z}^K_{n,k} - Z_F\right) \overset{d}{=} \sqrt{k}A\left(\frac{n}{k}\right)\mathscr{AB}_K(\gamma,\rho) + \mathbb{W}_{1,n} + \mathbb{W}_{2,n}(K) + \mathbb{W}_{3,n} + o_{\mathbb{P}}(1),$$

where

$$\mathscr{A}\mathscr{B}_K(\gamma,\rho) = \left(\frac{v(1)}{(1-\gamma)(\gamma+\rho-1)} + \frac{v(1)}{(1-\gamma)^2}\int_0^1 s^{-\rho}K(s)ds\right);$$
$$\mathbb{W}_{1,n} = \frac{\gamma v(1-k/n)}{(1-\gamma)}\sqrt{\frac{n}{k}}\mathbf{B}_n\left(1-\frac{k}{n}\right);$$
$$\mathbb{W}_{2,n}(K) = -\frac{\gamma v(1-k/n)}{(1-\gamma)^2}\sqrt{\frac{n}{k}}\left\{\int_0^1 \frac{1}{s}\mathbf{B}_n\left(1-\frac{sk}{n}\right)d(sK(s))\right\};$$
$$\mathbb{W}_{3,n} = -\frac{\int_0^{1-k/n} v(s)\mathbf{B}_n(s)}{(k/n)^{1/2}\mathbb{Q}(1-k/n)}d\mathbb{Q}(s) + o_{\mathbf{P}}(1),$$

and

$$v(s) = \int_0^s \frac{TVaR_F^*(t)}{(1-t)TVaR_F^2(t)}dt.$$

Now, by computing the asymptotic variances of the different processes appearing in Theorem 1, we deduce the following corollary.

Corollary 1. *Under the same assumptions of Theorem 1, if $\sqrt{k}A(n/k) \to \lambda \in \mathbb{R}$, we have*

$$\frac{\sqrt{k}}{(k/n)\mathbb{U}(n/k)}\left(\widetilde{Z}_{n,k}^K - Z_F\right) \xrightarrow{d} \mathscr{N}\left(\lambda\mathscr{A}\mathscr{B}_K(\gamma,\rho), \mathscr{A}\mathscr{C}_K(\gamma,\rho)\right),$$

where

$$\mathscr{A}\mathscr{C}_K(\gamma,\rho) = \frac{v^2(1)\gamma^2}{(1-\gamma)^2(2\gamma-1)} + \frac{v^2(1)\gamma^2}{(1-\gamma)^4}\int_0^1 K^2(s)ds.$$

The Corollary 1 generalizes the result of the Theorem 1 in Greselin, F. et al. (2014)[31] in case $\lambda \neq 0$ and when we use a general kernel instead of K.
In view of these results, $\widetilde{Z}_{n,k}^K$ is an estimator of Z_F with an asymptotic bias given by

$$(k/n)\mathbb{U}(n/k)A(n/k)\mathscr{A}\mathscr{B}_K(\gamma,\rho).$$

For a specific kernel, the asymptotic bias and variance can be computed. For instance, we have the following corollary 2 if $K = \underline{K}$.

Corollary 2. *Under the assumptions of Corollary 1 and in the special case where $K = \underline{K}$, we have*

$$\frac{\sqrt{k}(\widetilde{Z}_{n,k}^K - Z_F)}{(k/n)\mathbb{U}(n/k)} \xrightarrow{d} \mathscr{N}\left(\lambda\frac{\gamma\rho v(1)}{(1-\rho)(\gamma+\rho-1)(1-\gamma)^2}, \sigma_\gamma^2\right) \tag{14}$$

for any fixed $t \in (0,1)$, where the asymptotic variance σ_γ^2 is given by the formula

$$\sigma_\gamma^2 = \frac{\gamma^4 v^2(1)}{(1-\gamma)^4(2\gamma-1)}.$$

The next step is to propose a reduced-bias estimator of Z_F.

3.2 Bias-correction for the $\widetilde{Z}_{n,k}^K$ estimator

The problem of reduced bias of inequality measures estimation is well known in the literature, and it has been addressed recently by several authors, among whom we mention Fichtenbaum and Shahidi, 1988 [17], Breunig, 2002 [4], Deltas 2003 [13], all these researchers consider the possibility of dealing with the bias term in an appropriate way, building

different new reduced bias estimators.
For the kernel-type estimator $\widetilde{Z}^{K}_{n,k}$, we recall that, from Theorem 1,

$$\widetilde{Z}^{K}_{n,k} - (k/n)\,\mathbb{U}(n/k)A(n/k)\mathscr{A}\mathscr{B}_K(\gamma,\rho),$$

is an asymptotically unbiased estimator for Z_F. Note that γ,ρ, $\mathbb{U}(n/k)$ and $A(n/k)$ are unknown quantities that we have to estimate. Under the condition (12), Feuerverger and Hall (1999) [16] and Beirlant et al. (1999, 2002) [2, **?**] propose the following exponential regression model for the log-spacings of order statistics:

$$Y_i \sim \left(\gamma + A\left(\frac{n}{k}\right)\left(\frac{i}{k+1}\right)^{-\rho}\right) + \varepsilon_{i,k},\ 1 \le i \le k, \tag{15}$$

where $\varepsilon_{i,k}$ are zero-centered error terms. If we ignore the term $A(n/k)$ in equation (15), we obtain the Hill estimator $\widehat{\gamma}^{H}_{n,k}$ by taking the mean of the left-hand side of (15). By using a least-square approach, (15) can be further exploited to propose a reduced-bias estimator for γ in which ρ is substituted by a consistent estimator $\widehat{\rho} = \widehat{\rho}(n,k)$ (see for instance Beirlant et al., 2002 [3] and Fraga Alves et al., 2003) [19] or by a canonical choice, such as $\rho = -1$(see e.g. Feuerverger and Hall (1999) [16] or Beirlant et al., (1999)[2]). The least squares estimators for γ and $A(n/k)$ are then given by

$$\widehat{\gamma}^{L.S}_{n,k}(\widehat{\rho}) = \frac{1}{k}\sum_{i=1}^{k} Y_i - \frac{\widehat{A}^{L.S}_{n,k}(\widehat{\rho})}{1-\widehat{\rho}};$$

$$\widehat{A}^{L.S}_{n,k}(\widehat{\rho}) = \frac{(1-2\widehat{\rho})(1-\widehat{\rho})^2}{\widehat{\rho}^2}\frac{1}{k}\sum_{i=1}^{k}\left(\left(\frac{i}{k+1}\right)^{-\widehat{\rho}} - \frac{1}{1-\widehat{\rho}}\right)Y_i.$$

Note that $\widehat{\gamma}^{L.S}_{n,k}(\rho)$ can be viewed as the kernel estimator $\widehat{\gamma}^{K_\rho}_{n,k}$, where for $0 < u \le 1$:

$$K_\rho(u) := \frac{1-\rho}{\rho}\underline{K}(u) + \left(1 - \frac{1-\rho}{\rho}\right)\underline{K}_\rho(u) \tag{16}$$

with $\underline{K}(u) = 1_{\{0<u<1\}}$ and $\underline{K}_\rho(u) = \left(\frac{1-\rho}{\rho}\right)(u^{-\rho}-1)\,1_{\{0<u<1\}}$, both kernels satisfy condition $(\mathscr{K})$. On the contrary $\underline{K}_\rho$ does not satisfy statement (CK1) in $(\mathscr{K})$. We refer to Gomes and Martins (2004) [22] and Gomes et al., (2007) [26] for other techniques of bias reduction based on the estimation of the second-order parameter.

We are now able to obtain a reduced-bias estimator for the Zenga index Z_F from condition (12) and using the above estimators for the different unknown quantities:

$$\widetilde{Z}^{K,\widehat{\rho}}_{n,k} = \widetilde{Z}^{K}_{n,k} - \left(\frac{k}{n}\right)X_{n-k,n}\widehat{A}^{L.S}_{n,k}(\widehat{\rho})\mathscr{A}\mathscr{B}_K\left(\widehat{\gamma}^{L.S}_{n,k}(\widehat{\rho}),\widehat{\rho}\right). \tag{17}$$

The asymptotic normality of $\widetilde{Z}^{K,\widehat{\rho}}_{n,k}$ is established in the theorem 2.

Theorem 2._Under the assumptions of Theorem 1, if $\widehat{\rho}$ is a consistent estimator for ρ, then we have_

$$\frac{\sqrt{k}(\widetilde{Z}^{K,\widehat{\rho}}_{n,k} - Z_F)}{(k/n)\mathbb{U}(n/k)} \xrightarrow{d} \mathscr{N}\left(0, \widetilde{\mathscr{A}\mathscr{C}}_K(\gamma,\rho)\right),$$

where

$$\widetilde{\mathscr{A}\mathscr{C}}_K(\gamma,\rho) = \mathscr{A}\mathscr{C}_K(\gamma,\rho) + \frac{\gamma^2}{\rho^2}(1-2\rho)(1-\rho)^2\mathscr{A}\mathscr{B}^2_K(\gamma,\rho)$$
$$+\frac{2\gamma^2(1-2\rho)(1-\rho)}{\rho^2(1-\gamma)^2}\left(1-(1-\rho)\int_0^1\frac{K(s)}{s^\rho}ds\right)\mathscr{A}\mathscr{B}_K(\gamma,\rho).$$

Let us observe that $\widetilde{Z}^{K,\widehat{\rho}}_{n,k}$ has a null asymptotic bias, which is not the case for $\widetilde{Z}^{K}_{n,k}$ (Corollary 1).

Corollary 3. *Under the same assumptions as in Theorem 2 and in the special case where $K = \underline{K}$, we have*

$$\frac{\sqrt{k}\left(\widetilde{Z}_{n,k}^{K,\widehat{\rho}} - Z_F\right)}{(k/n)\mathbb{U}(n/k)} \xrightarrow{d} \mathcal{N}\left(0, \frac{\gamma^4(\gamma-\rho)^2 v^2(1)}{(2\gamma-1)(\gamma+\rho-1)^2(1-\gamma)^4}\right).$$

Now, in the special case where $K = K_\rho$, as already mentioned, the estimator $\widehat{\gamma}_{n,k}^{K_\rho}$ coincides with $\widehat{\gamma}_{n,k}^{L.S}(\rho)$. The aim of the next corollary is to establish the asymptotic normality of the resulting Zenga estimator $\widehat{Z}_{n,k}^{K_\rho,\widehat{\rho}}$, denoted by $\widehat{Z}_{n,k}^{L.S,\widehat{\rho}}$, when the least squares approach is adopted.

Corollary 4. *Under the same assumptions as in Theorem 2 and in the special case where $K = K_\rho$, we have*

$$\frac{\sqrt{k}\left(\widehat{Z}_{n,k}^{L.S,\widehat{\rho}} - Z_F\right)}{(k/n)\mathbb{U}(n/k)} \xrightarrow{d} \mathcal{N}\left(0, \widetilde{\mathcal{AC}}_{K_\rho}(\gamma,\rho)\right),$$

where

$$\widetilde{\mathcal{AC}}_{K_\rho}(\gamma,\rho) = \frac{\gamma^2(1-\rho)^2}{\rho^2(1-\gamma)^4}v^2(1) + \frac{v^2(1)\gamma^2}{(2\gamma-1)(1-\gamma)^2} + \frac{\gamma^2(1-2\rho)(1-\rho)(\gamma\rho+2\rho+\gamma-\rho-1)}{\rho^2(1-\gamma)^3(\gamma+\rho-1)^2}v^2(1).$$

4 Simulation study

In this section we examine via Monte Carlo simulations a finite-sample performance of the proposed (asymptotic) Zenga index measure, with particular emphasis on comparisons and illustrations of the performance of the biased estimator $\widetilde{Z}_{n,k}^{K}$ and the reduced-bias estimator $\widetilde{Z}_{n,k}^{L.S,\widehat{\rho}}$, we compare the two estimators in terms of the bias and Root Mean Square Error (RMSE), we note that bias1 and RMSE1 are for the estimator $\widetilde{Z}_{n,k}^{K}$, and bias2 and RMSE2 are for the estimator $\widetilde{Z}_{n,k}^{L.S,\widehat{\rho}}$, also we compare between the coverage probability (cov prob) of the bias and the RMSE for a significance level $\zeta = 0.95$, through its application to sets of samples taken from Pareto distribution with tail of distribution $\overline{F}(x) = 1 - x^{-1/\gamma}$, $x > 1$, and Fréchet distributions with tail of distribution $\overline{F}(x) = 1 - \exp\left(-x^{-1/\gamma}\right)$, $x > 0$, in this study, we consider two values of tail index ($\gamma = 2/3$ and $\gamma = 3/4$), and the second-order parameter $\rho = -1$, we generate 1000 independent replicates of samples sizes 1000, 2000 and 4000 from the selected parent distribution. With the optimal values of k, we estimate the following minimal Asymptotic Mean Squared Error (AMSE) values of $\widehat{\gamma}_{n,k_{opt}}^{K}$, several procedures have been suggested in the literature, and we refer to, e.g., (Dekkers and de Haan, (1993) [12], Drees and Kaufmann, (1998) [15], Danielsson et al., (2001) [10], Cheng and Peng, (2001) [6], Neves and Fraga Alves, (2004) [37]) and references therein.
In our current study we employ the method of Cheng and Peng (2001) [6] for deciding on an appropriate value k_{opt} of k. For each simulated sample, we obtain an approximation of the estimators of Zenga index Z_F. The overall estimated Z_F is then taken as the empirical mean of the values in the 1000 repetitions. To this end, we summarize the results in Table 1 and Table 2

5 Proofs

Let $Y_1,...,Y_n$ be i.i.d. r.v.'s from the unit Pareto distribution G, defined as $G(y) = 1 - 1/y, y > 1$. 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 results from Csörgö et al., (1986) [8], a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ is constructed carrying a sequence $1, 2, ...$ of independent random variables uniformly distributed on $(0,1)$ and a sequence of Brownian bridges $\{\mathbf{B}_n(s), 0 \leq s \leq 1, n = 1, 2....\}$. The resulting empirical quantile is denoted by

$$\beta_n(r) = \sqrt{n}(r - \mathbb{V}(r)),$$

where

$$\mathbb{V}(r) = \zeta_{j,n}, \frac{j-1}{n} \leq r \leq \frac{j}{n}, j = 1, 2, ..., n \quad and \quad \mathbb{V}(0) = 0.$$

Table 1: Simulations results based on Pareto distribution with $\gamma = 3/4$ and $\gamma = 2/3$, the corresponding Zenga index values 0.8247 and 0.7590, respectively

γ	2/3				3/4			
	bias1	RMSE1	bias2	RMSE2	bias1	RMSE1	bias2	RMSE2
n=1000	-0.0024	0.0075	-0.0012	0.0031	-0.0026	0.0069	-0.0016	0.0038
n=2000	-0.0018	0.0049	-0.0013	0.0023	-0.0019	0.0055	-0.0012	0.0029
n=4000	-0.0013	0.0036	-0.0010	0.0019	-0.0013	0.0038	-0.0010	0.0015
cov prob								
n=1000	0.627	0.665	0.714	0.755	0.605	0.685	0.669	0.775
n=2000	0.635	0.676	0.726	0.762	0.619	0.697	0.682	0.782
n=4000	0.639	0.681	0.744	0.785	0.632	0.701	0.705	0.796

Table 2: Simulations results based on Fréchet distribution with $\gamma = 3/4$ and $\gamma = 2/3$, the corresponding Zenga index values 0.8652 and 0.8229, respectively

γ	2/3				3/4			
	bias1	RMSE1	bias2	RMSE2	bias1	RMSE1	bias2	RMSE2
n=1000	0.0019	0.0073	0.0013	0.0055	0.0021	0.0055	0.0012	0.0024
n=2000	0.0010	0.0048	0.0009	0.0026	0.0015	0.0034	0.0011	0.0019
n=4000	0.0008	0.0022	0.0006	0.0012	0.0012	0.0023	0.0009	0.0010
cov prob								
n=1000	0.625	0.701	0.734	0.755	0.596	0.659	0.682	0.748
n=2000	0.633	0.713	0.747	0.774	0.613	0.665	0.699	0.768
n=4000	0.639	0.721	0.762	0.781	0.629	0.671	0.718	0.782

Proof(Proof of Theorem 1). Theorem 1 follows from the asymptotic expansion

$$\widetilde{Z}^K_{n,k} = 1 - \int_0^1 \left(\frac{\widetilde{TVaR}^{*K}_{n,k}(t)}{\widetilde{TVaR}^{K}_{n,k}(t)} \right) dt,$$

we have

$$\begin{aligned}
\frac{\sqrt{k}}{\sqrt{k/n}\mathbb{U}(n/k)} \left(\widetilde{Z_{n,k}}^K - Z_F \right) &= -\int_0^1 \frac{1}{TVaR_F(t)} \sqrt{n} \frac{\widetilde{TVaR}^{*K}_{n,k}(t) - TVaR^*_F(t)}{\sqrt{k/n}\mathbb{U}(n/k)} dt \\
&+ \int_0^1 \frac{TVaR^*_F(t)}{TVaR^2_F(t)} \sqrt{n} \frac{\widetilde{TVaR}^{K}_{n,k}(t) - TVaR_F(t)}{\sqrt{k/n}\mathbb{U}(n/k)} dt \\
&+ \int_0^1 \left(\frac{1}{TVaR^K_{n,k}(t)} - \frac{1}{TVaR_F(t)} \right) \\
&\times \sqrt{n} \frac{\widetilde{TVaR}^{*K}_{n,k}(t) - TVaR^*_F(t)}{\sqrt{k/n}\mathbb{U}(n/k)} dt \\
&+ \int_0^1 \frac{TVaR^*_F(t)}{TVaR_F(t)} \left(\frac{1}{\widetilde{TVaR}^{K}_{n,k}(t)} - \frac{1}{TVaR_F(t)} \right) \\
&\times \sqrt{n} \frac{\widetilde{TVaR}^{K}_{n,k}(t) - TVaR_F(t)}{\sqrt{k/n}\mathbb{U}(n/k)} dt \\
&= I_1 + I_2 + I_3 + I_4.
\end{aligned}$$

Greselin et al. (2014) [31] showed that $I_3 = o_P(1)$ and $I_4 = o_P(1)$ when $n \to \infty$.
On the other hand, we have

$$(1-t) \left(\widetilde{TVaR}^{K}_{n,k}(t) - TVaR_F(t) \right) = A_{n,1} + A_{n,2}(t),$$

and

$$t\left(\widetilde{TVaR}_{n,k}^{K*}(t)-TVaR_F^*(t)\right)=A_{n,2}(0)-A_{n,2}(t),$$

for all t such that $0 \le t \le 1-k/n$, where

$$A_{n,1}=\left(\frac{k}{n}\right)\frac{X_{n,k}}{1-\widehat{\gamma}_{n,k}^{K}}-\int_{1-k/n}^{1}Q(s)ds$$

$$A_{n,2}(t)=\int_{t}^{1-k/n}(Q_n(s)-Q(s))ds$$

We have, from Rassoul, (2013) [39], that

$$\frac{\sqrt{k}A_{n,1}(t)}{(k/n)\mathbb{Q}(1-k/n)}=-\frac{\int_0^{1-k/n}\mathbf{B}_n(s)d\mathbb{Q}(s)}{(k/n)^{1/2}\mathbb{Q}(1-k/n)}+o_{\mathbf{P}}(1)=\mathbb{W}_{n,3} \tag{18}$$

and

$$\frac{\sqrt{k}A_{n,2}(t)}{(k/n)\mathbb{Q}(1-k/n)}=\sqrt{k}A(n/k)\mathscr{A}\mathscr{B}_K(\gamma,\rho)+\mathbb{W}_{1,n}+\mathbb{W}_{2,n}. \tag{19}$$

The proof of statement (18) is similar to that of Theorem 2 in Necir *et al.,* (2010) [36], though some adjustments are needed since we are now concerned with the Zenga index. Therefore, we present the main blocks of the proof together with pinpointed references to Necir *et al.,* (2010) [36] for specific technical details.
Next, from theorem 1 in Rassoul (2013) [39], we have

$$\frac{(1-t)\sqrt{k}}{(k/n)\mathbb{U}(n/k)}\left(\widetilde{TVaR}_{n,k}^{K}(t)-TVaR_F(t)\right)$$
$$\stackrel{d}{=}\sqrt{k}A\left(\frac{n}{k}\right)\mathscr{A}\mathscr{B}_K(\gamma,\rho)+\mathbb{W}_{1,n}+\mathbb{W}_{2,n}(K)+\mathbb{W}_{3,n}+o_{\mathbb{P}}(1),$$

where

$$\mathscr{A}\mathscr{B}_K(\gamma,\rho)=\left(\frac{v(1)}{(1-\gamma)(\gamma+\rho-1)}+\frac{v(1)}{(1-\gamma)^2}\int_0^1 s^{-\rho}K(s)ds\right);$$

$$\mathbb{W}_{1,n}=\frac{\gamma v(1-k/n)}{(1-\gamma)}\sqrt{\frac{n}{k}}\mathbf{B}_n\left(1-\frac{k}{n}\right)(1+o_{\mathbb{P}}(1));$$

$$\mathbb{W}_{2,n}(K)=-\frac{\gamma v(1-k/n)}{(1-\gamma)^2}\sqrt{\frac{n}{k}}\left\{\int_0^1\frac{1}{s}\mathbf{B}_n\left(1-\frac{sk}{n}\right)d(sK(s))\right\};$$

$$\mathbb{W}_{3,n}=-\frac{\int_0^{1-k/n}v(s)\mathbf{B}_n(s)}{(k/n)^{1/2}\mathbb{Q}(1-k/n)}d\mathbb{Q}(s)+o_{\mathbf{P}}(1).$$

Finally

$$\frac{\sqrt{k}}{\sqrt{k/n}\mathbb{U}(n/k)}\left(\widetilde{Z}_{n,k}^{K}-Z_F\right)\stackrel{d}{=}\sqrt{k}A\left(\frac{n}{k}\right)\mathscr{A}\mathscr{B}_K(\gamma,\rho)+\mathbb{W}_{1,n}+\mathbb{W}_{2,n}(K)+\mathbb{W}_{3,n}+o_{\mathbb{P}}(1).$$

Proof(Proof of corollary 1). From Theorem 1, the sum $\mathbb{W}_{1,n}+\mathbb{W}_{2,n}(K)+\mathbb{W}_{3,n}$ is a centered Gaussian random variable. To calculate its asymptotic variance, the computations are tedious but quite direct. The classical Sultsky's lemma completes the proof of corollary 1.

Proof(Proof of corollary 2). The proof of the corollary 2 is a direct result of the corollary 1 with the kernel $K=\underline{K}=1_{(0,1)}$.

Proof(Proof of Theorem 2).
We have

$$\frac{\sqrt{k}(\widetilde{Z}_{n,k}^{K,\widehat{\rho}}-Z_F)}{(k/n)\mathbb{U}(n/k)}\stackrel{d}{=}\mathbb{W}_{1,n}+\mathbb{W}_{2,n}(K)+\mathbb{W}_{3,n}+\mathbb{W}_{4,n}+o_{\mathbb{P}}(1),$$

where

$$\begin{aligned}
\mathbb{W}_{4,n} &= \sqrt{k}\left(A(n/k)\mathscr{A}\mathscr{B}_K(\gamma,\rho) - \widehat{A}_{n,k}^{L.S}(\widehat{\rho})\mathscr{A}\mathscr{B}_K\left(\widehat{\gamma}_{n,k}^{L.S}(\widehat{\rho}),\widehat{\rho}\right)\frac{X_{n-k,n}}{\mathbb{U}(n/k)}\right) \\
&= -\mathscr{A}\mathscr{B}_K(\gamma,\rho)\sqrt{k}\left(\widehat{A}_{n,k}^{L.S}(\widehat{\rho}) - A(n/k)\right) \\
&\quad -\sqrt{k}\widehat{A}_{n,k}^{L.S}(\widehat{\rho})\left(\mathscr{A}\mathscr{B}_K\left(\widehat{\gamma}_{n,k}^{L.S}(\widehat{\rho}),\widehat{\rho}\right) - \mathscr{A}\mathscr{B}_K(\gamma,\rho)\right) \\
&\quad -\sqrt{k}\widehat{A}_{n,k}^{L.S}(\widehat{\rho})\mathscr{A}\mathscr{B}_K\left(\widehat{\gamma}_{n,k}^{L.S}(\widehat{\rho}),\widehat{\rho}\right)\left(\frac{X_{n-k,n}}{\mathbb{U}(n/k)} - 1\right) \\
&\overset{d}{=} -\mathscr{A}\mathscr{B}_K(\gamma,\rho)\gamma(1-\rho)\sqrt{\frac{n}{k}}\left\{\int_0^1 \frac{1}{s}\mathbf{B}_n\left(1-\frac{sk}{n}\right)d\left(s\left(\underline{K}(s) - K_\rho(s)\right)\right)\right\} \\
&\quad + o_{\mathbb{P}}(1).
\end{aligned}$$

By the result of Lemma 5 of Girard and Guillou, (2013) [21], for any consistent estimator $\widehat{\rho}$ of ρ, we have

$$\sqrt{k}\left(\widehat{\gamma}_{n,k}^{L.S}(\widehat{\rho}) - \gamma\right) \overset{d}{=} \gamma\sqrt{\frac{n}{k}}\int_0^1 \frac{1}{s}\mathbf{B}_n\left(1-\frac{sk}{n}\right)d\left(sK_\rho(s)\right) + o_{\mathbb{P}}(1), \tag{20}$$

and

$$\begin{aligned}
&\sqrt{k}\left(\widehat{A}_{n,k}^{L.S}(\widehat{\rho}) - A\left(\frac{n}{k}\right)\right) \\
&\overset{d}{=} \gamma(1-\rho)\sqrt{\frac{n}{k}}\int_0^1 \frac{1}{s}\mathbf{B}_n\left(1-\frac{sk}{n}\right)d\left(s\left(\underline{K}(s) - K_\rho(s)\right)\right) + o_{\mathbb{P}}(1),
\end{aligned} \tag{21}$$

and by using the consistency and the inequality $|\frac{e^x-1}{x} - 1| \leq e^{|x|} - 1$ for all $x \in \mathbb{R}$. Moreover, direct computations lead to the desired asymptotic variance which ends the proof of Theorem 2.

Proof(Proof of corollary 3). The proof of the corollary 3 is a direct result of the Theorem 2 with the kernel $K = \underline{K} = 1_{(0,1)}$.

Proof(Prooof of corollary 4). Recall that K_ρ does not satisfy condition $(\mathscr{K})$ but it can be rewritten as (16) with both K and K_ρ satisfying $(\mathscr{K})$. So, following the lines of the proof of Theorem 2, Corollary 4 follows.

6 Conclusion

This paper develops a kernel-type estimator of the Zenga index for heavy-tailed model. Our estimator is consistent and asymptotically normal, and performs excellently in finite samples as shown by Monte Carlo simulations.

Acknowledgements

The authors are very grateful to the editor and the two referees for their careful reading of the paper, and for their comments and suggestions on an earlier draft of this paper, which led to significant improvements of the initial draft. Any errors are our responsibility.

References

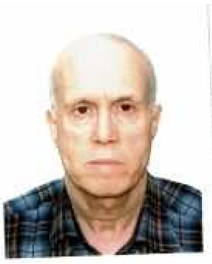
[1] A. B. Abel, Optimal capital income taxation. NBER Working Paper No. 13354. http://www.nber.org/papers/w13354,(2007).
[2] J. Beirlant, G. Dierckx, M.Goegebeur, and G. Matthys, Tail index estimation and an exponential regression model, *Extremes*, **2**, 177-200, (1999).
[3] J. Beirlant, G. Dierckx, A. Guillou, and C. Starica, On exponential representations of log-spacings of extreme order statistics, *Extremes*, **5**, 157-180, (2002).
[4] R. Breunig, Bias correction for inequality measures: an application to China and Kenya, *Applied Economics Letters*, **9(12)**, 783-786, (2002).
[5] C. Chamley, Optimal taxation of capital income in general equilibrium with infinite lives. *Econometrica* **54**, 607-622, (1986).
[6] S. Cheng, and L. Peng, Confidence intervals for the tail index, *Bernoulli*, **7(5)**, 751-760,(2001).

[7] M. Csörgo, and D. M., Mason, Central limit theorems for sums of extreme values. *Math. Proc. Camb. Phil. Soc.*, **98**, 547-558, (1985)..
[8] M. Csörgo, S. Csörgo, L. Horvärth, and D. M. Mason, Weighted empirical and quantile processes, *Ann. Prob.*, **14**, 31-85, (1986).
[9] S. Csörgo, P. Deheuvels, and D. M. Mason, Kernel estimates of the tail index of a distribution, *Annals of Statistics*, **13**, 1050- 1077, (1985).
[10] J. Danielsson, L. de Haan, L. Peng, and C. G. de Vries, Using a bootstrap method to choose the sample fraction in tail index estimation. *Journal of Multivariate analysis*, **76(2)**, 226-248, (2001).
[11] L. de Haan, and L. Peng, Comparison of tail index estimators, *Statistica Neerlandica*, **52**, 60-70, (1998).
[12] A. L. Dekkers, and L. de Haan, Optimal choice of sample fraction in extreme value estimation. *Journal of Multivariate Analysis*, **47(2)**, 173-195, (1993).
[13] G. Deltas, The Small-Sample Bias of the Gini Coefficient: Results and Implications for Empirical Research, *The Review of Economics and Statistics*, **85(1)**, 226-234, (2003).
[14] M. Denuit, J. Dhaene, M. J. Goovaerts, and R. Kaas, *Actuarial Theory for Dependent Risk:Measures, Orders and Models,* Wiley, New York, (2005).
[15] H. Drees, and E. Kaufmann, Selecting the optimal sample fraction in univariate extreme value estimation, *Stochastic Processes and their Applications*, **75(2)**, 149-172, (1998).
[16] A. Feuerverger, and P. Hall, Estimating a tail exponent by modelling departure from a Pareto distribution, *Annals of Statistics*, **27**, 760-781, (1999).
[17] R. Fichtenbaum, and H. Shahidi, Truncation Bias and the Measurement of Income Inequality, Journal of Business & Economic Statistics, **6(3)**, 335-337, (1988).
[18] A. Frädorf, M. M. Grabka, and J. Schwarze, The impact of household capital income on income inequality: a factor decomposition analysis for the UK, Germany and the USA. *J. Econ. Inequal.* **9**, 35-56,(2011).
[19] M. I. Fraga Alves, L. de Haan, and T. Lin, Estimation of the parameter controlling the speed of convergence in extreme value theory. *Math. Methods Statist.*, **12**, 155-176, (2003).
[20] J. L. Gastwirth, The estimation of the Lorenz curve and Gini index. *The Review of Economics and Statistics*, 306-316, (1972).
[21] S. Girard, and A. Guillou, Reduced-bias estimator of the Proportional Hazard Premium for heavy-tailed distributions. *Insurance: Mathematics and Economics*, **52(3)**, 550-559, (2013).
[22] M. I. Gomes, and M.J. Martins, Bias reduction and explicit semi-parametric estimation of the tail index, *Journal of Statistical Planning and Inference*, **124**, 361-378, (2004).
[23] M. J. Goovaerts, F. de Vlyder, and J. Haezendonck, *Insurance premiums, theory and applications*, North Holland, Amsterdam, (1984).
[24] M.I. Gomes, F. Figueiredo, and S. Mendona, Asymptotically best linear unbiased tail estimators under a second-order regular variation condition. *J. Stat. Plann. Inference*, textbf134(2), 409-433, (2005).
[25] M.I. Gomes, and D. Pestana, A sturdy reduced-bias extreme quantile (VaR) estimator, *Journal of the American Statistical Association*, **102(477)**, 280-292, (2007).
[26] M. I. Gomes, M. I. Martins, and M. Neves, Improving second order reduced bias extreme value index estimator, *REVSTAT-Statistical* Journal, **5(2)**, 177-207, (2007).
[27] M. Golosov, N. Kocherlakota, and A. Tsyvinski, Optimal indirect and capital taxation, *Rev. Econ. Studies*, **70**, 569-587, (2003).
[28] F. Greselin, and L. Pasquazzi, Asymptotic confidence intervals for a new inequality measure, *Comm. Stat. Comput. Simul.*, **38**, 17-42, (2009).
[29] F. Greselin, L. Pasquazzi, and R. Zitikis, Zenga's new index of economic inequality, its estimation, and an analysis of incomes in Italy. J. Prob. Stat. (Spec. Issue on Actuarial and Financial Risks: Models, Statistical Inference, and Case Studies). Article ID 718905, p. 26, (2010).
[30] F. Greselin, L. Pasquazzi, and R. Zitikis, Contrasting the Gini and Zenga indices of economic inequality, *J. Appl. Stat.*, **40**, 282-297, (2013).
[31] F. Greselin, L. Pasquazzi, and R. Zitikis, Heavy tailed capital incomes: Zenga index, statistical inference, and ECHP data analysis, *Extremes*, **17**,127-155, (2014).
[32] P. Groeneboom, H. P. LopuhaA, and P. de Wolf, Kernel-type estimators for the extreme value index, *Annals of Statistics*, **31**, 1956-1995, (2003).
[33] B. M. Hill, A simple approach to inference about the tail of a distribution, *Annals of Statistics*, **3**, 1136-1174, (1975).
[34] K.L. Judd, Capital-income taxation with imperfect competition, *Am. Econ. Rev.*, **92**, 417-421, (2002).
[35] R. I. Lerman, and S. Yitzhaki, Income inequality effects by income source: a new approach and applications to the United States, *Rev. Econ. Stat.*, **67**, 151-156, (1985).
[36] A. Necir, A. Rassoul, and R. Zitikis, Estimating the conditional tail expectation in the case of heavy-tailed losses, *JPS*, doi:10.1155/596839, (2010).
[37] C. Neves, and M. F. Alves, Reiss and Thomas automatic selection of the number of extremes, *Computational statistics data analysis*, **47(4)**, 689-704, (2004).
[38] L. Peng, and Y. Qi, Estimating the first- and second-order parameters of a heavy-tailed distribution, *Aust. N. Z. J. Stat.*, **46**, 305-312, (2004).
[39] A. Rassoul, Kernel-type estimator of the conditional tail expectation for a heavy-tailed distribution, *Insurance: Mathematics and Economics*, **53(3)**, 698-703, (2013).

[40] P.B. Sörensen, Can capital income taxes survive? And should they? *CESifo Econ. Stud.*, **53**, 172-228, (2007).
[41] E. Saez, Top incomes in the United States and Canada over the twentieth century. *J. Eur. Econ. Assoc.*, **3**, 402-411, (2005).
[42] I. Weissman, Estimation of parameters and larges quantiles based on the k largest observations, *Journal of American Statistical Association*, **73**, 812-815, (1978).
[43] M. Zenga, Inequality curve and inequality index based on the ratios between lower and upper arithmetic means, *Stat. Appl.*, **5**, 3-27, (2007).

Tami Omar The Head of Department of Mathematics of the Faculty of Sciences of the University of Blida 1 in Algeria, He is assistant lecturer and spent decades teaching basic and applied mathematics, also teaching differents program of probability theory and inference statistics. He is member in LAMDA-RO laboratory. His research interests are in the areas of applied mathematics and statistics, Extremes values theory, pParametric and non parametric statistics, mathematical economics, measure of risk and applications in actuariat and hydrology, applied economics.

Rassoul Abdelaziz is Associate Professor of Mathematics at Higher National School of Hydraulics of Blida in Algeria Research Laboratory of GEE of ENSH. He received the PhD degree in “Mathematical Statistics” at University of Biskra (Algeria). He is referee of several international journals in the frame of pure and applied mathematics, applied economics. His main research interests are: Application of Extremes values theory in many fields as actuariat, hydrology and meteorology, dangerous models, times series, max stables processes, mathematical economics.



Ould Rouis Hamid is Associate Professor of Mathematics at Department of Mathematics of the Faculty of Sciences of the University of Blida 1 in Algeria. Research Laboratory of LAMDA-RO of university of Blida 1. He received the PhD degree in “Mathematical Probability and Statistics” at University of Boston (USA). His main research interests are: theory of probability, stochastic process and their applications, limites theoremes, asymptotics theory, order statistics.