# Estimation of Zenga's new index of economic inequality in heavy tailed populations

Greselin, Francesca and Pasquazzi, Leo

Department of Quantitative Methods for Economics and Business Sciences, University of Milano Bicocca

2 June 2011

# ESTIMATION OF ZENGA'S NEW INDEX OF ECONOMIC INEQUALITY IN HEAVY TAILED POPULATIONS.

Francesca Greselin
University of Milano-Bicocca
francesca.greselin@unimib.it

Leo Pasquazzi
University of Milano-Bicocca
leo.pasquazzi@unimib.it

Abstract

In this work we propose a new estimator for Zenga's inequality measure in heavy tailed populations. The new estimator is based on the Weissman estimator for high quantiles. We will show that, under fairly general conditions, it has asymptotic normal distribution. Further we present the results of a simulation study where we compare confidence intervals based on the new estimator with those based on the plug-in estimator.

## 1. Introduction

In this work we propose a new estimator for Zenga's (2007) inequality measure in heavy tailed populations. Zenga's index is a recently introduced risk and inequality measure, which is the based on the ratio between the lower and upper conditional tail expectations. Let $F(x)$ be the cdf of a non-negative random variable $X$, which describes income, wealth, an actuarial risk or a financial loss. Throughout this paper we shall assume that $F$ is continuous. Let $Q(s)$ be the quantile function and denote the upper and lower conditional tail expectations by

$$CTE_F(t) = \frac{1}{1-t}\int_t^1 Q(s)ds, \quad 0 \le t < 1,$$

and

$$CTE_F^*(t) = \frac{1}{t}\int_0^t Q(s)ds, \quad 0 < t \le 1,$$

respectively. We may measure inequality between the "poorest" $t$ percent of the population and the remaining "richer" part of it by

$$Z(t) = 1 - \frac{CTE_F^*(t)}{CTE_F(t)}, \qquad 0 < t < 1,$$

and averaging Zenga's inequality curve $Z(t)$ over $t$ we get Zenga's inequality index

$$Z = \int_0^1 Z(t)dt,$$

which measures the overall inequality in the population.

Greselin *et al.* (2010) have recently derived the asymptotic normality of the plug-in estimator for $Z$, assuming that $E(X^{2+\epsilon}) < \infty$ for some $\epsilon > 0$. In this work we shall be concerned with heavy tailed populations. More precisely, we will deal with the case where $F$ is regularly varying at infinity with tail index $\gamma$ between 0.5 and 1. Formally, we shall assume that

$$\lim_{t\to\infty} \frac{1 - F(tx)}{1 - F(t)} = x^{-1/\gamma}$$

for some $\gamma \in (0.5, 1)$. Notice that it does not make sense to consider larger values of $\gamma$ since this would imply that $F$ has infinite mean, in which case Zenga's index is not even defined.

The new estimator for $Z$ we are going to introduce is based on the Weissman (1978) estimator

$$\widetilde{q}_s := \left(\frac{k}{n}\right)^{\widehat{\gamma}_n} \frac{X_{n-k:n}}{(1-s)^{\widehat{\gamma}_n}}, \qquad 1 - \frac{k}{n} < s < 1,$$

for large quantiles. In the definition of $\widetilde{q}_s$ we indicate by $X_{1:n} < X_{2:n} < \cdots < X_{n:n}$ the order statistics associated with the i.i.d. sample random variables $X_1, X_2, ..., X_n \sim F$ and by $k$ the sample fraction in the Hill estimator

$$\widehat{\gamma}_n = \frac{1}{k} \sum_{i=1}^{k} \ln X_{n-i+1:n} - \ln X_{n-k:n}$$

for the tail index $\gamma$.

If we estimate $CTE_F(t)$ and $CTE_F^*(t)$ by

$$CTE_n(t) = \begin{cases} \frac{1}{1-t}\left(\int_t^{1-k/n} Q_n(s)ds + \frac{kX_{(n-k)}}{n(1-\widehat{\gamma})}\right) & 0 \le t < 1 - k/n \\ \left(\frac{k}{n}\right)^{\widehat{\gamma}} \frac{X_{(n-k)}}{(1-t)^{\widehat{\gamma}}(1-\widehat{\gamma})} & 1 - k/n \le t < 1 \end{cases}$$

and by

$$CTE_n^*(t) = \frac{1}{t}\left[CTE_n(0) - (1-t)CTE_n(t)\right], \qquad 0 < t \le 1$$

respectively, we get the following estimator for $Z$:

$$Z_n = 1 - \int_0^1 \frac{CTE_n^*(t)}{CTE_n(t)} dt.$$

In the above definition of $CTE_n$ and what follows we indicate by $Q(t)$ and $Q_n(t)$ the quantile function and its empirical counterpart (both functions are left-continuous).

Necir *et al.* (2010) have recently derived asymptotic normality of $CTE_n(t)$ (to be precise of the expression in the first line of the definition) assuming that $F$ satisfies the generalized second order regular variation condition with second order parameter $\rho \leq 0$, i.e. assuming that there exists a function $\alpha(t) \to 0$ as $t \to \infty$, which does not change sign in a neighborhood of infinity, such that for every $x > 0$,

$$\lim_{t\to\infty} \frac{1}{\alpha(t)} \left( \frac{1 - F(tx)}{1 - F(t)} - x^{-1/\gamma} \right) = x^{-1/\gamma} \frac{x^{\rho/\gamma} - 1}{\rho/\gamma}, \tag{1.1}$$

and that the sample fraction $k = k_n \to \infty$ as $n \to \infty$ in such a way that $k/n \to 0$ and $\sqrt{k}\, \alpha(Q(1 - k/n)) \to 0$. In this paper we will see that the same set of assumptions ensures the asymptotic normality of the new estimator for Zenga's index as well. As in Necir *et al.* (2010), the proofs are crucially based on the extreme value theory.

The rest of the paper is organized as follows. In section 2 we state the main theorem and discuss its practical implementation. The proof of the theorem is deferred to the appendix. In section 3 we present the results of a simulation study, where we compare confidence intervals based on the new estimator with those based on the plug-in estimator in a variety of settings. Conclusions and final remarks end the paper in section 4.

## 2. Main theorem and Practical Implementation

**Theorem 2.1.** *Assume that $F$ satisfies condition (1.1) with $0 < \gamma < 1/2$, $\rho \leq 0$ and let $k = k_n \to \infty$ as $n \to \infty$ in such a way that $k/n \to 0$ and $\sqrt{k}\, \alpha(Q(1 - k/n)) \to 0$. Then,*

$$\begin{aligned} \frac{\sqrt{n}\,(Z_n - Z)}{\sqrt{k/n} X_{n-k:n}} &= -\int_0^{1-k/n} \frac{\mathcal{B}_n(s) v(s)}{\sqrt{k/n} Q(1 - k/n)} dQ(s) \\ &\quad + \frac{\gamma^2\, v(1 - k/n)}{(1 - \gamma)^2} \sqrt{\frac{n}{k}} \mathcal{B}_n \left( 1 - \frac{k}{n} \right) \\ &\quad - \frac{\gamma\, v(1 - k/n)}{(1 - \gamma)^2} \sqrt{\frac{n}{k}} \int_{1-k/n}^1 \frac{\mathcal{B}_n(s)}{1 - s} ds + o_{\mathbf{P}}(1) \end{aligned}$$

*and hence we have*

$$\frac{\sqrt{n}\,(Z_n - Z)}{\sqrt{k/n} X_{n-k:n}} \xrightarrow{d} N(0, \sigma^2),$$

*where*

$$\sigma^2 = \frac{\gamma^4}{(1 - \gamma)^4 (2\gamma - 1)} \int_0^1 \frac{CTE_F^*(t)}{(1 - t) CTE_F(t)^2} dt.$$

The proof of the theorem is deferred to the appendix.

Notice that both the new estimator for Zenga's index as well as the plug-in estimator for its variance depend on the sample fraction $k$. In fact, in the former this dependence is direct as well as mediated through the Hill estimator, while the latter depends on $k$ only through the Hill estimator. It is well known that the Hill estimator has large variance for small values of $k$ and large bias if $k$ is large. We therefore need to balance between these two shortcomings. Several adaptive procedures for an optimal choice of $k$ have been proposed in the literature. In the simulation study in the next section we will employ the method of Cheng and Peng (2001). According to the simulation results it works reasonably well.

## 3. Simulations

In this section we present the results of a simulation study to assess the performance of the asymptotic normal confidence intervals based on the normal limit distribution of the plug-in estimator and the new estimator. As parent distribution we used Zenga's model for economic size distributions (Zenga, 2010), a very flexible three parameter family with paretian right tail. It depends on a scale parameter that coincides with the model mean and two shape parameters $\theta$ and $\alpha$, affecting the center and the tails of the distribution, respectively. The interested reader may find more information about Zenga's model in Zenga *et al.* (2010a), (2010b).

In our simulations we set the parameter values to ML-estimates obtained on capital income data from the 2001 wave of the ECHP survey. We considered three parent distributions with low, intermediate and high tail index. Our simulation results (table 4) confirm that the confidence intervals based on the plug-in estimator suffer from undercoverage. The new ones seem to solve this issue. They are slightly larger (cfr. the quantiles of the estimated standard errors) but their coverage accuracy is almost exact.

## Appendix A

In this appendix we prove, under the assumptions of theorem 2.1, the following asymptotic expansion

$$\frac{\sqrt{n}\,(Z_n - Z)}{\sqrt{k/n}Q(1-k/n)} = \sum_{i=1}^{3} T_{n,i} + o_{\mathbf{P}}(1). \tag{3.1}$$

The three main terms in the expansion are given by

$$T_{n,1} = -\int_0^{1-k/n} \frac{\mathcal{B}_n(s)v(s)}{\sqrt{k/n}Q(1-k/n)}dQ(s),$$

$$T_{n,2} = \frac{\gamma^2\ v(1-k/n)}{(1-\gamma)^2}\sqrt{\frac{n}{k}}\mathcal{B}_n\left(1-\frac{k}{n}\right),$$

$$T_{n,3} = -\frac{\gamma\ v(1-k/n)}{(1-\gamma)^2}\sqrt{\frac{n}{k}}\int_{1-k/n}^1 \frac{\mathcal{B}_n(s)}{1-s}ds,$$

where, for each $n$, $\mathcal{B}_n$ is a Brownian bridge that approximates the empirical process induced by the cdf transformations of the sample random variables and

$$v(s) \;=\; \int_0^s \frac{CTE_F^*(t)}{CTE_F(t)^2(1-t)}dt, \qquad 0 < s \leq 1.$$

It is easily verified that

$$\begin{array}{lll} E(T_{n,1}^2)\longrightarrow\frac{2\gamma}{2\gamma-1}v(1)^2 & E(T_{n,1}T_{n,2})\longrightarrow-\frac{\gamma^2}{(1-\gamma)^2}v(1)^2 & E(T_{n,1}T_{n,3})\longrightarrow\frac{\gamma}{(1-\gamma)^2}v(1)^2 \\ E(T_{n,2}^2)\longrightarrow\frac{\gamma^4}{(1-\gamma)^4}v(1)^2 & E(T_{n,2}T_{n,3})\longrightarrow\frac{\gamma^3}{(1-\gamma)^4}v(1)^2 & E(T_{n,3}^2)\longrightarrow\frac{2\gamma^2}{(1-\gamma)^4}v(1)^2, \end{array}$$

so that

$$\frac{\sqrt{n}\,(Z_n - Z)}{\sqrt{k/n}Q(1-k/n)} \xrightarrow{d} N(0,\sigma^2), \tag{3.2}$$

where

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

By theorem 2.4.8 in de Haan and Ferreira (2006), the second order condition in (1.1) and $\sqrt{k}\alpha(Q(1-k/n)) \to 0$ ensure that

$$\sqrt{k}\left(\frac{X_{n-k:n}}{Q(1-k/n)} - 1\right) = O_{\mathbf{P}}(1), \tag{3.3}$$

so that we may substitute $Q(1-k/n)$ with $X_{n-k}$ in the limit relation in (3.2).

To prove (3.1), we first linearize $Z_n - Z_F$ with respect to the $CTE$'s. This yields

$$\begin{aligned} \frac{\sqrt{n}\,(Z_n - Z)}{\sqrt{k/n}Q(1-k/n)} &= -\int_0^1 \frac{1}{CTE_F(t)}\frac{\sqrt{n}\,(CTE_n^*(t) - CTE_F^*(t))}{\sqrt{k/n}Q(1-k/n)}dt+ \\ &\quad + \int_0^1 \frac{CTE_F^*(t)}{CTE_F(t)^2}\frac{\sqrt{n}\,(CTE_n(t) - CTE_F(t))}{\sqrt{k/n}Q(1-k/n)}dt+ \\ &\quad + r_{n,1} + r_{n,2}, \end{aligned} \tag{3.4}$$

where

$$r_{n,1} \;=\; -\int_0^1 \left(\frac{1}{CTE_n(t)} - \frac{1}{CTE_F(t)}\right)\frac{\sqrt{n}\,(CTE_n^*(t) - CTE_F^*(t))}{\sqrt{k/n}Q(1-k/n)}dt$$

$$r_{n,2} \;=\; \int_0^1 \frac{CTE_F^*(t)}{CTE_F(t)}\left(\frac{1}{CTE_n(t)} - \frac{1}{CTE_F(t)}\right)\frac{\sqrt{n}\,(CTE_n(t) - CTE_F(t))}{\sqrt{k/n}Q(1-k/n)}dt$$

We shall show below (in Appendix B) that

$$r_{n,1} = o_{\mathbf{P}}(1), \tag{3.5}$$

$$r_{n,2} = o_{\mathbf{P}}(1). \tag{3.6}$$

$$\int_{1-k/n}^{1} \frac{1}{CTE_F(t)} \frac{\sqrt{n}\left(CTE_n^*(t) - CTE_F^*(t)\right)}{\sqrt{k/n}Q(1-k/n)} dt = o_{\mathbf{P}}(1), \tag{3.7}$$

and finally that

$$\int_{1-k/n}^{1} \frac{CTE_F^*(t)}{CTE_F(t)^2} \frac{\sqrt{n}\left(CTE_n(t) - CTE_F(t)\right)}{\sqrt{k/n}Q(1-k/n)} dt = o_{\mathbf{P}}(1). \tag{3.8}$$

Thus we have

$$\begin{aligned} \frac{\sqrt{n}\left(Z_n - Z\right)}{\sqrt{k/n}Q(1-k/n)} = & -\int_{0}^{1-k/n} \frac{1}{CTE_F(t)} \frac{\sqrt{n}\left(CTE_n^*(t) - CTE_F^*(t)\right)}{\sqrt{k/n}Q(1-k/n)} dt + \\ & + \int_{0}^{1-k/n} \frac{CTE_F^*(t)}{CTE_F(t)^2} \frac{\sqrt{n}\left(CTE_n(t) - CTE_F(t)\right)}{\sqrt{k/n}Q(1-k/n)} dt + o_{\mathbf{P}}(1). \end{aligned} \tag{3.9}$$

Like in Necir *et al.* (2010), we now write

$$CTE_n(t) - CTE_F(t) = \frac{1}{1-t}\left(A_{n,1}(t) + A_{n,2}\right), \qquad 0 \le t \le 1-k/n, \tag{3.10}$$

and

$$CTE_n^*(t) - CTE_F^*(t) = \frac{1}{t}\left(A_{n,1}(0) - A_{n,1}(t)\right), \qquad 0 < t \le 1-k/n, \tag{3.11}$$

where

$$\begin{aligned} A_{n,1}(t) &= \int_{t}^{1-k/n} \left(Q_n(s) - Q(s)\right) ds, \qquad 0 \le t \le 1-k/n, \\ A_{n,2} &= \frac{k/n}{1-\widehat{\gamma}} X_{n-k:n} - \int_{1-k/n}^{1} Q(s) ds. \end{aligned}$$

Notice that

$$A_{n,1}(t) = -\int_{Q(t)}^{Q(1-k/n)} (F_n(x) - F(x))dx + V_n\left(1 - \frac{k}{n}\right) - V_n(t). \tag{3.12}$$

where

$$V_n(t) = \int_0^t (Q_n(s) - Q(s))ds + \int_{-\infty}^{Q(t)} (F_n(x) - F(x))dx,$$

is the Vervaat process. In the appendix we list some properties of the Vervaat process, which will be needed in this proof.

Substituting (3.12) in (3.10) and (3.11), respectively, and changing the variables of integration, we get

$$CTE_n(t) - CTE_F(t) = \frac{1}{1-t}\left[-\int_t^{1-k/n} \frac{e_n(s)}{\sqrt{n}} dQ(s) + V_n\left(1-\frac{k}{n}\right) - V_n(t) + A_{n,2}\right] \tag{3.13}$$

and (since $V_n(0) = 0$)

$$CTE_n^*(t) - CTE_F^*(t) = \frac{1}{t}\left[-\int_0^t \frac{e_n(s)}{\sqrt{n}} dQ(s) + V_n(t)\right], \tag{3.14}$$

where $e_n(t) = \sqrt{n}(F_n(F^{-1}(t)) - t)$, is the uniform on $[0,1]$ empirical process. Substituting now (3.13) and (3.14) in (3.9), yields

$$\begin{aligned}\frac{\sqrt{n}\,(Z_n - Z)}{\sqrt{k/n}Q(1-k/n)} = &-\int_0^{1-k/n} \frac{1}{CTE_F(t)\; t} \frac{\int_0^t e_n(s)dQ(s)}{\sqrt{k/n}Q(1-k/n)} dt \\ &-\int_0^{1-k/n} \frac{CTE_F^*(t)}{CTE_F(t)^2(1-t)} \frac{\int_t^{1-k/n} e_n(s)dQ(s)}{\sqrt{k/n}Q(1-k/n)} dt \\ &+\frac{\sqrt{n}A_{n,2}}{\sqrt{k/n}Q(1-k/n)} \int_0^{1-k/n} \frac{CTE_F^*(t)}{CTE_F(t)^2(1-t)} dt \\ &+ r_{n,3} + r_{n,4},\end{aligned} \tag{3.15}$$

where

$$\begin{aligned} r_{n,3} &= \frac{\sqrt{n}}{\sqrt{k/n}Q(1-k/n)} \int_0^{1-k/n} \frac{1}{CTE_F(t)\; t} V_n(t)dt \\ r_{n,4} &= -\frac{\sqrt{n}}{\sqrt{k/n}Q(1-k/n)} \int_0^{1-k/n} \frac{CTE_F^*(t)}{CTE_F(t)^2} \frac{1}{1-t}\left(V_n\left(1-\frac{k}{n}\right) - V_n(t)\right) dt.\end{aligned}$$

Below we shall show that

$$r_{n,3} = o_{\mathbf{P}}(1), \tag{3.16}$$

and finally that

$$r_{n,4} = o_{\mathbf{P}}(1), \tag{3.17}$$

as well.

Applying Fubini's theorem in the first two terms in (3.15) yields

$$\begin{aligned}\frac{\sqrt{n}\,(Z_n - Z)}{\sqrt{k/n}Q(1-k/n)} = &-\int_0^{1-k/n} \frac{e_n(s)w_n(s)}{\sqrt{k/n}Q(1-k/n)} dQ(s) \\ &-\int_0^{1-k/n} \frac{e_n(s)v(s)}{\sqrt{k/n}Q(1-k/n)} dQ(s) \\ &+\frac{\sqrt{n}A_{n,2}\; v(1-k/n)}{\sqrt{k/n}Q(1-k/n)} + o_{\mathbf{P}}(1),\end{aligned} \tag{3.18}$$

where

$$\begin{aligned} w_n(s) &= \int_s^{1-(k/n)} \frac{1}{CTE_F(t)\ t} dt, \qquad 0 < s \le 1-(k/n) \\ v(s) &= \int_0^s \frac{CTE_F^*(t)}{CTE_F(t)^2(1-t)} dt, \qquad 0 < s \le 1. \end{aligned}$$

Below we shall show that

$$\int_0^{1-k/n} \frac{e_n(s)w_n(s)}{\sqrt{k/n}Q(1-k/n)} dQ(s) = o_{\mathbf{P}}(1). \tag{3.19}$$

Finally we choose a sequence of Brownian bridges $\mathcal{B}_n$ as in Result 1 that replaces the emprical process $e_n(s)$ in (3.18). In view of (3.19), this yields

$$\begin{aligned} \frac{\sqrt{n}\,(Z_n - Z)}{\sqrt{k/n}Q(1-k/n)} &= -\int_0^{1-k/n} \frac{\mathcal{B}_n(s)v(s)}{\sqrt{k/n}Q(1-k/n)} dQ(s) \\ &\quad + \frac{\sqrt{n}A_{n,2}\ v(1-k/n)}{\sqrt{k/n}Q(1-k/n)} + r_{n,5} + o_{\mathbf{P}}(1), \end{aligned} \tag{3.20}$$

with

$$r_{n,5} = O_{\mathbf{P}}(1)\frac{1}{n^{\nu_1}} \int_0^{1-k/n} (1-s)^{(1/2)-\nu_1} v(s) ds,$$

for any $0 \le \nu_1 < \frac{1}{4}$. Since

$$\sup_{0 \le s \le 1} v(s) < \infty, \tag{3.21}$$

the integral in $r_{n,5}$ remains bounded as $n$ goes to infinity. Thus, it follows that

$$r_{n,5} = o_{\mathbf{P}}(1). \tag{3.22}$$

Now, using the same sequence of Brownian bridges $\mathcal{B}_n$ as in (3.20), it may be shown (see Necir *et al.*, 2010 and references therein) that

$$\begin{aligned} \frac{\sqrt{n}A_{n,2}}{\sqrt{k/n}Q(1-k/n)} &= \frac{\gamma^2}{(1-\gamma)^2}\sqrt{\frac{n}{k}}\mathcal{B}_n\left(1-\frac{k}{n}\right) \\ &\quad - \frac{\gamma}{(1-\gamma)^2}\sqrt{\frac{n}{k}}\int_{1-(k/n)}^1 \frac{\mathcal{B}_n(s)}{1-s} ds + o_{\mathbf{P}}(1) \end{aligned} \tag{3.23}$$

which finally, in view of (3.20), (3.21) and (3.22), yields the asymptotic expansion in (3.1).

## 4. Appendix B

In this section we will prove negligibility of the remainder terms in appendix A. In our proofs we will need the following lemmas.

**Lemma 4.1.** *For each $\epsilon > 0$, as small as desired, we have*

$$\sup_{0\le t<1-k/n}(1-t)\left|CTE_n(t)-CTE_F(t)\right| = o_{\mathbf{P}}(1)\frac{1}{\sqrt{k}}\left(\frac{k}{n}\right)^{1-\gamma-\epsilon} \tag{4.1}$$

$$\sup_{0<t\le 1-k/n} t\left|CTE_n^*(t)-CTE_F^*(t)\right| = o_{\mathbf{P}}(1)\frac{1}{\sqrt{k}}\left(\frac{k}{n}\right)^{1-\gamma-\epsilon} \tag{4.2}$$

*Proof.* By (3.13), (3.14) and (3.23), the proof reduces to showing that

$$\frac{1}{\sqrt{n}}\int_0^{1-k/n}|e_n(s)|dQ(s) \le o_{\mathbf{P}}(1)\frac{1}{\sqrt{k}}\left(\frac{k}{n}\right)^{1-\gamma-\epsilon} \tag{4.3}$$

and that

$$\sup_{0\le t\le 1-k/n} V_n(t) \le o_{\mathbf{P}}(1)\frac{1}{\sqrt{k}}\left(\frac{k}{n}\right)^{1-\gamma-\epsilon}. \tag{4.4}$$

By result 3, we immediately get

$$\frac{1}{\sqrt{n}}\int_0^{1-k/n}|e_n(s)|dQ(s) \le O_{\mathbf{P}}(1)\frac{1}{\sqrt{n}}\int_0^{1-k/n}(1-s)^{1/2-\epsilon}dQ(s)$$

Changing the integration variable on the RHS yields

$$\begin{aligned}\frac{1}{\sqrt{n}}\int_0^{1-k/n}(1-s)^{1/2-\epsilon}dQ(s) &= \frac{1}{\sqrt{n}}\int_{Q(0)}^{Q(1-k/n)}(1-F(x))^{1/2-\epsilon}dx \\ &\sim \frac{2\gamma}{1-2\epsilon}\frac{1}{\sqrt{k}}\left(\frac{k}{n}\right)^{1-\epsilon}Q(1-k/n).\end{aligned}$$

Since $Q(1-k/n) = o((n/k)^{\gamma+\epsilon})$, the proof of (4.3) is complete.

To get a bound for the LHS in (4.4), we use first property (c) of the Vervaat process and then results 3 and 4. This yields

$$V_n(t) \le o_{\mathbf{P}}(1)\frac{1}{\sqrt{n}}(1-t)^{1/2-\gamma-2\epsilon}$$

But by hypothesis we have $1/2-\gamma-2\epsilon < 0$. For $0 \le t \le 1-k/n$ we thus have the following bounded for the RHS of the previous inequality:

$$o_{\mathbf{P}}(1)\frac{1}{\sqrt{k}}\left(\frac{k}{n}\right)^{1-\gamma-2\epsilon}.$$

This completes the proof of (4.4). □

**Lemma 4.2.** *For each $\epsilon > 0$, as small as desired, we have*

$$\sup_{s\geq 1} \frac{\sqrt{k}}{Q(1-k/n)} \frac{CTE_n\left(1-\frac{k}{ns}\right) - CTE_F\left(1-\frac{k}{ns}\right)}{s^{\gamma+\epsilon}} = O_{\mathbf{P}}(1) \tag{4.5}$$

*and*

$$\sup_{s\geq 1} \frac{\sqrt{n}}{\sqrt{k/n}Q(1-k/n)} \frac{CTE_n^*\left(1-\frac{k}{ns}\right) - CTE_F^*\left(1-\frac{k}{ns}\right)}{s^{\gamma+\epsilon}} = O_{\mathbf{P}}(1) \tag{4.6}$$

*Proof.* Suppose that (4.5) is true. Then, by lemma (4.1), it is easily seen that (4.6) is also true.

We shall now prove (4.5). Note that

$$\begin{aligned}\frac{CTE_n\left(1-\frac{k}{ns}\right) - CTE_F\left(1-\frac{k}{ns}\right)}{Q(1-k/n)} &= \frac{X_{n-k:n}}{Q(1-k/n)}\frac{s^{\widehat{\gamma}}}{1-\widehat{\gamma}} - \frac{s^{\gamma}}{1-\gamma} \\ &\quad + \frac{s^{\gamma}}{1-\gamma} - \frac{CTE_F\left(1-\frac{k}{ns}\right)}{Q(1-k/n)}\end{aligned}$$

We will prove the lemma by showing that

$$\sup_{s\geq 1} \frac{\sqrt{k}}{s^{\gamma+\epsilon}} \left| \frac{X_{n-k:n}}{Q(1-k/n)}\frac{s^{\widehat{\gamma}}}{1-\widehat{\gamma}} - \frac{s^{\gamma}}{1-\gamma}\right| = O_{\mathbf{P}}(1) \tag{4.7}$$

and that

$$\sup_{s\geq 1} \frac{\sqrt{k}}{s^{\gamma+\epsilon}} \left| \frac{s^{\gamma}}{1-\gamma} - \frac{CTE_F\left(1-\frac{k}{ns}\right)}{Q(1-k/n)}\right| = o(1). \tag{4.8}$$

Consider first the assertion in (4.8). Notice that

$$\frac{s^{\gamma}}{1-\gamma} - \frac{CTE_F\left(1-\frac{k}{ns}\right)}{Q(1-k/n)} = s\int_s^{\infty}\left(t^{\gamma} - \frac{Q\left(1-\frac{k}{nt}\right)}{Q\left(1-\frac{k}{n}\right)}\right)\frac{1}{t^2}dt \tag{4.9}$$

We will use the second order regular variation condition (1.1) to show that the integral converges to 0 faster than $k^{-1/2}$. By theorem 2.3.9 condition in de Haan and Ferreira (2006), condition (1.1) is equivalent to

$$\lim_{t\to\infty} \frac{\frac{Q(1-1/(tx)}{Q(1-1/t)} - x^{\gamma}}{\alpha(Q(1-1/t)} = x^{\gamma}\frac{x^{\rho}-1}{\rho}, \qquad x>0$$

and it implies that, for any positive $\epsilon$ and $\delta$, there exists $t_0 > 1$ ($t_0$ depends on $\epsilon$ and $\delta$), such that for all $t, tx \geq t_0$,

$$\left|\frac{1}{A_0(t)}\left(\frac{Q(1-1/(tx))}{Q(1-1/t)} - x^{\gamma}\right) - x^{\gamma}\frac{x^{\rho}-1}{\rho}\right| \leq \epsilon x^{\gamma+\rho}\max(x^{\delta}, x^{-\delta})$$

for some function $A_0(t) \sim \alpha(Q(1-1/t))$. Applying this inequality with $0<\delta<|\rho|$ in (4.8), yields, for large enough $n$,

$$\frac{\sqrt{k}}{s^{\gamma+\epsilon}}\left|\frac{s^{\gamma}}{1-\gamma} - \frac{CTE_F\left(1-\frac{k}{ns}\right)}{Q(1-k/n)}\right| \leq \sqrt{k}A_0(n/k)\left(\frac{1}{|\rho|}+\epsilon\right) s^{1-\gamma-\epsilon}\int_s^{\infty} t^{\gamma-2}dt \tag{4.10}$$

Since

$$\lim_{s\to\infty} s^{1-\gamma-\epsilon}\int_s^\infty t^{\gamma-2}dt = 0,$$

it follows that

$$\sup_{s\geq 1} s^{1-\gamma-\epsilon}\int_s^\infty t^{\gamma-2}dt < \infty.$$

Moreover, by the hypothesis on the function $\alpha(Q(1-1/t))$, $\sqrt{k}A_0(n/k) \to 0$ as $n \to \infty$ and thus the RHS in (4.10) converges to zero uniformly for $s \geq 1$ as $n$ goes to infinity. This completes the proof of the assertion in (4.8).

Consider now (4.7) and note that

$$\begin{aligned}\frac{1}{s^{\gamma+\epsilon}}\left(\frac{X_{n-k:n}}{Q(1-k/n)}\frac{s^{\widehat{\gamma}}}{1-\widehat{\gamma}} - \frac{s^\gamma}{1-\gamma}\right) = {} & \frac{s^{\widehat{\gamma}-\gamma-\epsilon}}{1-\widehat{\gamma}}\left(\frac{X_{n-k:n}}{Q(1-k/n)} - 1\right) \\ & + \frac{s^{\widehat{\gamma}-\gamma-\epsilon} - s^{-\epsilon}}{1-\widehat{\gamma}} \\ & + \frac{s^{-\epsilon}(\widehat{\gamma}-\gamma)}{(1-\widehat{\gamma})(1-\gamma)}.\end{aligned} \tag{4.11}$$

We will show that each term on the RHS is uniformly bounded (for $s \geq 1$) by a random variable of order $O_{\mathbf{P}}(1)$. To this aim we first observe that for $s \geq 1$,

$$|s^{\widehat{\gamma}-\gamma-\epsilon} - s^{-\epsilon}| \leq \begin{cases} \left(1 - \frac{\widehat{\gamma}-\gamma}{\epsilon}\right)^{\epsilon/(\widehat{\gamma}-\gamma)}\left(\frac{\epsilon}{\epsilon-(\widehat{\gamma}-\gamma)} - 1\right), & \text{if } |\widehat{\gamma}-\gamma| < \epsilon \text{ and } \widehat{\gamma} \neq \gamma, \\ 0, & \text{if } \widehat{\gamma} = \gamma \end{cases}$$

and that

$$\left(1 - \frac{\widehat{\gamma}-\gamma}{\epsilon}\right)^{\epsilon/(\widehat{\gamma}-\gamma)}\left(\frac{\epsilon}{\epsilon-(\widehat{\gamma}-\gamma)} - 1\right) \to 0 \quad \text{if} \quad \widehat{\gamma} \to \gamma.$$

Since (Mason, 1982)

$$\widehat{\gamma} \xrightarrow{\mathbf{P}} \gamma, \tag{4.12}$$

it follows that

$$\sup_{s\geq 1} |s^{\widehat{\gamma}-\gamma-\epsilon} - s^{-\epsilon}| = o_{\mathbf{P}}(1), \tag{4.13}$$

which, along with (3.3), implies that

$$\sup_{s\geq 1} \frac{\sqrt{k}s^{\widehat{\gamma}-\gamma-\epsilon}}{1-\widehat{\gamma}}\left(\frac{X_{n-k:n}}{Q(1-k/n)} - 1\right) = O_{\mathbf{P}}(1). \tag{4.14}$$

In order to deal with the second and third term on the RHS of (4.11), we need to know about the asymptotic behaviour of the Hill estimator. Under the second order condition in (1.1) and the assumption on the asymptotic behaviour of the function $\alpha(Q(1-1/t))$, we have (see theorem 3.2.5 in de Haan and Ferreira, 2006)

$$\sqrt{k}(\widehat{\gamma}-\gamma) = O_{\mathbf{P}}(1) \tag{4.15}$$

By the mean value theorem, we may write the second term on the RHS in (4.11) in the following way:

$$\frac{s^{\widehat{\gamma}-\gamma-\epsilon}-s^{-\epsilon}}{1-\widehat{\gamma}}=\frac{s^{\gamma^{*}-\gamma-\epsilon/4}\ln s}{s^{\epsilon 3/4}(1-\widehat{\gamma})}(\widehat{\gamma}-\gamma), \tag{4.16}$$

where $\gamma<\gamma^{*}<\widehat{\gamma}$, and thus, by (4.12), $\gamma^{*}\xrightarrow{\mathbf{P}}\gamma$. Using now (4.13) with $\gamma^{*}$ in the place of $\widehat{\gamma}$ and $\epsilon/4$ instead of $\epsilon$, the RHS in (4.16) may be written as

$$\frac{(s^{\epsilon/4}+o_{\mathbf{P}}(1))\ln s}{s^{\epsilon 3/4}(1-\widehat{\gamma})}(\widehat{\gamma}-\gamma)$$

which, multiplied by $\sqrt{k}$, is bounded by

$$\left|\frac{\sqrt{k}(\widehat{\gamma}-\gamma)}{(1-\widehat{\gamma})}\right|$$

for all $s\geq 1$. By (4.15), it follows that

$$\sup_{s\geq 1}\sqrt{k}\frac{s^{\widehat{\gamma}-\gamma-\epsilon}-s^{-\epsilon}}{1-\widehat{\gamma}}=O_{\mathbf{P}}(1). \tag{4.17}$$

For the last term in (4.11), we immediately notice that (4.15) implies

$$\sup_{s\geq 1}\frac{s^{-\epsilon}(\widehat{\gamma}-\gamma)}{(1-\widehat{\gamma})(1-\gamma)}=O_{\mathbf{P}}(1). \tag{4.18}$$

Finally we notice that (4.11), (4.14), (4.17) and (4.18) imply that the assertion in (4.7) is true. The proof of the lemma is thus complete. □

It is worth noting that the function $A_0(t)$, and thus also the function $\alpha(Q(1-1/t)$, are regularly varying at infinity with tail index $\rho$.

**Proof of (3.5) and (3.6).** We will show that

$$\frac{\sqrt{n}}{\sqrt{k/n}Q(1-k/n)}\int_0^1\left|\frac{1}{CTE_n(t)}-\frac{1}{CTE_F(t)}\right||CTE_n^{*}(t)-CTE_F^{*}(t)|\,dt=o_{\mathbf{P}}(1). \tag{4.19}$$

Notice that this implies both (3.5) and (3.6).

We proceed by splitting the integral into three parts

$$\int_0^1\cdots=\int_0^{\delta}\cdots+\int_{\delta}^{1-k/n}\cdots+\int_{1-k/n}^{1}\cdots$$

and showing that each of them is of order $o_{\mathbf{P}}(1)$ when multiplied by $\frac{\sqrt{n}}{\sqrt{k/nQ(1-k/n)}}$.

Consider first the term with $\int_0^{\delta}\cdots$. By (4.1) we see that it is bounded by

$$\frac{\sqrt{n}}{\sqrt{k/n}Q(1-k/n)}\frac{1}{CTE_F^2(0)+o_{\mathbf{P}}(1)}\int_0^{\delta}\frac{1-t}{1-\delta}|CTE_n(t)-CTE_F(t)||CTE_n^{*}(t)-CTE_F^{*}(t)|dt$$

and using (4.1) again, we may bound the latter expression by

$$\frac{o_{\mathbf{P}}(1)}{(k/n)^{\gamma+\epsilon}Q(1-k/n)}\frac{1}{CTE_F^2(0)+o_{\mathbf{P}}(1)}\frac{1}{1-\delta}\int_0^\delta |CTE_n^*(t)-CTE_F^*(t)|dt.$$

Since, for small enough $\epsilon>0$, $\sqrt{n}(k/n)^{\gamma+\epsilon}Q(1-k/n)\to\infty$ as $n\to\infty$, we need to show that

$$\int_0^\delta |CTE_n^*(t)-CTE_F^*(t)|dt = O_p(n^{-1/2}).$$

By (3.14) this will be true if

$$\int_0^\delta \frac{1}{t}\int_0^t |e_n(s)|dQ(s)dt = O_{\mathbf{P}}(1) \tag{4.20}$$

and

$$\int_0^\delta \frac{\sqrt{n}V_n(t)}{t}dt = o_{\mathbf{P}}(1). \tag{4.21}$$

Now, apply Fubini's theorem in (4.20) to get

$$\int_0^\delta |e_n(s)|\ln(\delta/s)dQ(s),$$

which, by result 3, is bounded by

$$O_{\mathbf{P}}(1)\int_0^\delta s^{1/2-\epsilon}\ln(\delta/s)dQ(s) = O_{\mathbf{P}}(1).$$

This proves (4.20). To see that (4.21) is also true, we first use the bound in property (c) of the Vervaat process in result 2. This yields

$$\begin{aligned}\int_0^\delta \frac{\sqrt{n}V_n(t)}{t}dt &\le \int_0^\delta \frac{|e_n(t)||Q_n(t)-Q(t)|}{t}dt\\ &\le O_{\mathbf{P}}(1)\int_0^\delta t^{-1/2-\epsilon}|Q_n(t)-Q(t)|dt\\ &\le o_{\mathbf{P}}(1)\int_0^\delta t^{-1/2-\epsilon}(1-t)^{1/\gamma-\epsilon}dt\end{aligned}$$

In the above chain of inequalities we used result 3 in line two and result 4 in line three. Since the integral in line three is finite, this completes the proof of (4.21) and of the fact that $\frac{\sqrt{n}}{\sqrt{k/n}Q(1-k/n)}\int_0^\delta\cdots = o_{\mathbf{P}}(1)$.

We now turn our attention to the term with the $\int_\delta^{1-k/n}$ part of the integral in (4.19). Using lemma (4.1) we see that it is bounded by

$$\frac{o_{\mathbf{P}}(1)}{\sqrt{k}Q(1-k/n)}\left(\frac{k}{n}\right)^{1-2\gamma-2\epsilon}\frac{1}{\delta}\int_\delta^{1-k/n}\frac{1}{(1-t)CTE_F^2(t)+o_{\mathbf{P}}(1)}dt,$$

where

$$\frac{1}{\sqrt{k}Q(1-k/n)}\left(\frac{k}{n}\right)^{1-2\gamma-2\epsilon} = o(1),$$

since, for each $\epsilon > 0$, $(k/n)^{\gamma-\epsilon}Q(1-k/n) \to \infty$, and

$$\int_{\delta}^{1-k/n} \frac{1}{(1-t)CTE_F^2(t) + o_{\mathbf{P}}(1)} dt = O_{\mathbf{P}}(1),$$

since, for each $\epsilon > 0$, $CTE_F(t)(1-t)^{\gamma-\epsilon} \to \infty$ as $t$ approaches 1.

Finally, we shall deal with the term with the $\int_{1-k/n}^{1} \cdots$ part of the integral in (4.19). Changing the integration variable, this term may be written as

$$\frac{\sqrt{n}}{\sqrt{k/n}Q(1-k/n)} \int_1^\infty \frac{\left[CTE_n\left(1-\frac{k}{ns}\right) - CTE_F\left(1-\frac{k}{ns}\right)\right]\left[CTE_n\left(1-\frac{k}{ns}\right) - CTE_F\left(1-\frac{k}{ns}\right)\right]}{CTE_n\left(1-\frac{k}{ns}\right)CTE_F\left(1-\frac{k}{ns}\right)} \frac{k}{n}\frac{1}{s^2}ds$$

and, by lemma 4.2, it is bounded by

$$\frac{O_{\mathbf{P}}(1)}{\sqrt{k}}\frac{k}{n}Q\left(1-\frac{k}{n}\right)\int_1^\infty \frac{s^{2\gamma+2\epsilon-2}}{CTE_n\left(1-\frac{k}{ns}\right)CTE_F\left(1-\frac{k}{ns}\right)}ds$$

which in turn is bounded by

$$\frac{O_{\mathbf{P}}(1)}{\sqrt{k}}\frac{k}{n}\frac{Q\left(1-\frac{k}{n}\right)}{CTE_n\left(1-\frac{k}{n}\right)}\int_1^\infty \frac{s^{2\gamma+2\epsilon-2}}{CTE_F\left(1-\frac{k}{ns}\right)}ds.$$

Since, again by lemma 4.2, $CTE_n\left(1-k/n\right) \xrightarrow{\mathbf{P}} \infty$ and since, by Lebesgue's monotone convergence theorem,

$$\int_1^\infty \frac{s^{2\gamma+2\epsilon-2}}{CTE_F\left(1-\frac{k}{ns}\right)}ds = o(1),$$

it finally follows that the $\int_{1-k/n}^{1} \cdots$ part is of order $o_{\mathbf{P}}(1)$ as well.

**Proof of (3.7) and (3.8).** Changing the integration variable, the rest term in (3.7) may be written as

$$\int_1^\infty \frac{1}{CTE_F\left(1-\frac{k}{ns}\right)}\frac{\sqrt{n}\left[CTE_n^*\left(1-\frac{k}{ns}\right) - CTE_F^*\left(1-\frac{k}{ns}\right)\right]}{\sqrt{k/n}Q(1-k/n)}\frac{k}{n}\frac{1}{s^2}ds, \tag{4.22}$$

and the rest term in (3.8) as

$$\int_1^\infty \frac{CTE_F^*\left(1-\frac{k}{ns}\right)}{CTE_F\left(1-\frac{k}{ns}\right)}\frac{\sqrt{k}\left[CTE_n\left(1-\frac{k}{ns}\right) - CTE_F\left(1-\frac{k}{ns}\right)\right]}{Q(1-k/n)}\frac{1}{s^2}ds. \tag{4.23}$$

By lemma 4.2 we see that (4.22) is bounded by

$$O_{\mathbf{P}}(1)\frac{k}{n}\int_1^\infty \frac{s^{\gamma+\epsilon-2}}{CTE_F\left(1-\frac{k}{ns}\right)}ds$$

and that (4.23) is bounded by

$$O_{\mathbf{P}}(1)CTE_F^*(1)\int_1^\infty \frac{s^{\gamma+\epsilon-2}}{CTE_F\left(1-\frac{k}{ns}\right)}ds.$$

Notice that both bounds are of order $o_{\mathbf{P}}(1)$, since, by Lebesgue's monotone convergence theorem,

$$\int_1^\infty \frac{s^{\gamma+\epsilon-2}}{CTE_F\left(1-\frac{k}{ns}\right)}ds = o(1).$$

**Proof of (3.16) and (3.17).** Notice that

$$r_{n,3} = \frac{1}{\sqrt{k/n}Q(1-k/n)}\int_0^{1-k/n}\frac{1}{CTE_F(t)\ t}V_n(t)dt \geq 0$$

by property b) in result 2 about Vervaat processes and that

$$\begin{aligned} r_{n,3} &\leq \frac{1}{\sqrt{k/n}Q(1-k/n)}\int_0^{1-k/n}\frac{1}{CTE_F(t)\ t}|e_n(t)||Q_n(t)-Q(t)|dt \\ &\leq \frac{o_{\mathbf{P}}(1)}{\sqrt{k/n}Q(1-k/n)}\int_0^{1-k/n}\frac{(1-t)^{\gamma+\epsilon}}{CTE_F(t)\ t}|e_n(t)|dt \\ &\leq \frac{o_{\mathbf{P}}(1)}{\sqrt{k/n}Q(1-k/n)}\int_0^{1-k/n}\frac{(1-t)^{\gamma+\epsilon}}{CTE_F(t)\ t}t^{1/2-\epsilon}dt \end{aligned}$$

In the second line we used property c) in result 2 about Vervaat processes and in the third and fourth line we used results 4 and 3, respectively. Since the integral in the last line remains bounded as $n$ goes to infinity and since, by the hypothesis on the tail index, $\sqrt{k/n}Q(1-k/n)$ goes to infinity, the last bound is of order $o_{\mathbf{P}}(1)$ and the proof of (3.16) is complete.

In order to prove (3.17), we split the remainder term in two parts

$$r_{n,4}^{(1)} = -\frac{\sqrt{n}}{\sqrt{k/n}Q(1-k/n)}V_n\left(1-\frac{k}{n}\right)\int_0^{1-k/n}\frac{CTE_F^*(t)}{CTE_F(t)^2}\frac{1}{1-t}dt$$

and

$$r_{n,4}^{(2)} = \frac{\sqrt{n}}{\sqrt{k/n}Q(1-k/n)}\int_0^{1-k/n}\frac{CTE_F^*(t)}{CTE_F(t)^2}\frac{1}{1-t}V_n(t)dt,$$

and show that each of them is of order $o_{\mathbf{P}}(1)$. Indeed, the integral $r_{n,4}^{(1)}$ remains bounded as $n$ goes to infinity and, by properties b) and c) in result 2 about Vervaat processes, we have

$$0 \leq \frac{\sqrt{n}}{\sqrt{k/n}Q(1-k/n)}V_n\left(1-\frac{k}{n}\right) \leq \left|\sqrt{\frac{n}{k}}e_n\left(1-\frac{k}{n}\right)\right|\left|\frac{Q_n(1-k/n)}{Q(1-k/n)}-1\right|,$$

where

$$\sqrt{\frac{n}{k}}e_n\left(1-\frac{k}{n}\right) = O_{\mathbf{P}}(1)$$

and, by (3.3),

$$\frac{Q_n(1-k/n)}{Q(1-k/n)}-1 = \frac{X_{n-k:n}}{Q(1-k/n)}-1 = o_{\mathbf{P}}(1).$$

This proves that $r_{n,4}^{(1)} = o_{\mathbf{P}}(1)$. For $r_{n,4}^{(2)}$ we may again use properties b) and c) in result 2 about Vervaat processes along with results 3 and 4, to get

$$\begin{aligned} 0 \le r_{n,4}^{(2)} &\le \frac{1}{\sqrt{k/n}Q(1-k/n)} \int_0^{1-k/n} \frac{CTE_F^*(t)}{CTE_F(t)^2} \frac{1}{1-t} |e_n(t)||Q_n(t)-Q(t)|dt \\ &\le \frac{O_{\mathbf{P}}(1)}{\sqrt{k/n}Q(1-k/n)} \int_0^{1-k/n} \frac{CTE_F^*(t)}{CTE_F(t)^2} (1-t)^{-1/2-\epsilon} |Q_n(t)-Q(t)|dt \\ &\le \frac{o_{\mathbf{P}}(1)}{\sqrt{k/n}Q(1-k/n)} \int_0^{1-k/n} \frac{CTE_F^*(t)}{CTE_F(t)^2} (1-t)^{-1/2-\gamma-2\epsilon} dt \end{aligned}$$

Since the latter integral remains bounded as $n$ goes to infinity and since, by the hypothesis on the tail index $\sqrt{k/n}Q(1-k/n) \to \infty$, this implies that $r_{n,4}^{(2)} = o_{\mathbf{P}}(1)$.

**Proof of (3.19).** By result 3, we have

$$\frac{\int_0^{1-k/n} e_n(s) w_n(s)\, dQ(s)}{\sqrt{k/n}Q(1-k/n)} \le \frac{O_{\mathbf{P}}(1)}{\sqrt{k/n}Q(1-k/n)} \int_0^{1-k/n} s^{1/2-\epsilon} w_n(s) dQ(s)$$

Since $\sqrt{k/n}Q(1-k/n) \to \infty$, our task reduces to showing that the integral on the RHS remains bounded as $n$ goes to infinity. But this is certainly true since

$$\begin{aligned} \int_0^{1-k/n} s^{1/2-\epsilon} w_n(s) dQ(s) &= \int_0^{1-k/n} s^{1/2-\epsilon} \int_s^{1-k/n} \frac{1}{CTE_F(t)\ t} dt\ dQ(s) \\ &= \int_0^{1-k/n} \frac{1}{CTE_F(t)\ t} \int_0^t s^{1/2-\epsilon} dQ(s)\ dt \\ &\le \int_0^{1-k/n} \frac{Q(t)}{CTE_F(t)\ t^{1/2+\epsilon}} dt \end{aligned}$$

and

$$\lim_{t\to 1} \frac{Q(t)}{CTE_F(t)} = \lim_{s\to\infty} \frac{Q(1-1/s)}{s \int_s^\infty Q(1-1/s) s^{-2} ds} = 1-\gamma.$$

## Appendix C

This section contains some results that we used in our proofs.

**Result 1.** *There exists a probability space* $(\Omega, \mathcal{A}, P)$ *carrying a sequence* $U_1, U_2, \ldots$ *of independent random variables uniformly distributed on* $(0,1)$ *and a sequence of Brownian bridges* $\mathcal{B}_n$, $0 \le s \le 1$, $n = 1, 2, \ldots$ *such that for the uniform empirical process*

$$e_n(s) = \sqrt{n}\{G_n - s\}, \qquad 0 \le s \le 1,$$

*and the quantile process*

$$\beta_n(s) = \sqrt{n}\{s - U_n(s)\}, \qquad 0 \le s \le 1,$$

*where*

$$G_n(s) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(s \ge U_i), \qquad 0 \le s \le 1,$$

*and, with* $U_{1,n} \le U_{2,n} \le \cdots \le U_{n,n}$ *denoting the order statistics corresponding to* $U_1, U_2, \ldots, U_n$,

$$U_n(s) = \begin{cases} U_{k,n}, & \text{if } (k-1)/n < s \le k/n,\ k = 1, 2, \ldots, n \\ U_{1,n}, & \text{if } s = 0, \end{cases}$$

*we have*

$$\sup_{0 \le 1-(1/n)} n^{\nu_1} \frac{|e_n(s) - \mathcal{B}_n(s)|}{(1-s)^{\frac{1}{2}-\nu_1}} = O_{\mathbf{P}}(1) \tag{4.24}$$

*and*

$$\sup_{0 \le 1-(1/n)} n^{\nu_2} \frac{|\beta_n(s) - \mathcal{B}_n(s)|}{(1-s)^{\frac{1}{2}-\nu_2}} = O_{\mathbf{P}}(1), \tag{4.25}$$

*where* $\nu_1$ *and* $\nu_2$ *are any fixed numbers such that* $0 \le \nu_1 < \frac{1}{4}$ *and* $0 \le \nu_2 < \frac{1}{2}$. *The statement in (4.25) follows from theorem 2.1 in M. Csorgo et al. (1986), while the statement in (4.24) is contained in Corollary 2.1 of the above paper.*

**Result 2.** *The Vervaat process, defined by*

$$V_n(t) = \int_0^t (Q_n(s) - Q(s))ds + \int_{-\infty}^{Q(t)} (F_n(x) - F(x))dx,$$

*has the following properties:*

(a) $V_n(0) = 0$,

(b) $V_n(t) \ge 0$ *for all* $t \in [0,1]$,

(c) $\sqrt{n}V_n(t) \le |e_n(t)||Q_n(t) - Q(t)|$, *where* $e_n(t) = \sqrt{n}(F_n(Q(t)) - F(Q(t)))$ *is the uniform empirical process.*

**Result 3.** *For any $\epsilon > 0$ as small as desired,*

$$\sup_{x\in\mathbb{R}} \frac{\sqrt{n}\,(F_n(x) - F(x))}{F(x)^{(1/2)-\epsilon}(1 - F(x))^{(1/2)-\epsilon}} = O_{\mathbf{P}}(1),$$

**Result 4.** *(Mason, 1982) If $E(X^r) < \infty$ for some $r > 1$, then*

$$\sup_{0<s<1} (1 - s)^{1/r} |Q_n(s) - Q(s)| = o_{\mathbf{P}}(1).$$

Acknowledgments

The authors are grateful to Prof. Ričardas Zitikis of the University of Western Ontario (Canada), for providing invaluable help and advice in the initial phase of this research.

Table 4.1. Simulation results: estimates of some quantiles of the plug-in estimator for the standard errors and estimated coverage accuracies of confidence intervals for Zenga's index. The "average k" column reports the average of the estimates for the optimal sample fraction while the "max cov" column reports the estimates of the probability that the Hill estimator takes on a value in the interval $(0.5, 1)$.

| | plug-in estimator | | | | new estimator | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Zenga distribution with** $\alpha = 1.8594$ **and** $\theta = 29.3769 \Rightarrow \gamma = 0.3497$ **and** $Z = 0.9603$ | | | | | | | | | |
| Estimated standard errors | | | | | | | | | |
| n | min | median | max | mean | min | median | max | mean | average k |
| 1000 | 0.0024 | 0.0037 | 0.0114 | 0.0039 | 0.0028 | 0.0046 | 0.1612 | 0.0055 | 94 |
| 2000 | 0.0019 | 0.0027 | 0.0113 | 0.0028 | 0.0022 | 0.0032 | 0.1629 | 0.0036 | 33 |
| 4000 | 0.0015 | 0.0020 | 0.0102 | 0.0020 | 0.0017 | 0.0023 | 0.0393 | 0.0024 | 42 |
| Estimated coverage probabilites | | | | | | | | | |
| n | 0.9000 | 0.9500 | 0.9750 | 0.9900 | 0.9000 | 0.9500 | 0.9750 | 0.9900 | max cov |
| 1000 | 0.8694 | 0.9234 | 0.9635 | 0.9787 | 0.9187 | 0.9598 | 0.9824 | 0.9896 | 0.9973 |
| 2000 | 0.8773 | 0.9312 | 0.9691 | 0.9849 | 0.9229 | 0.9646 | 0.9856 | 0.9927 | 0.9986 |
| 4000 | 0.8802 | 0.9369 | 0.9721 | 0.9852 | 0.9204 | 0.9645 | 0.9866 | 0.9937 | 0.9995 |
| **Zenga distribution with** $\alpha = 0.9068$ **and** $\theta = 7.0462 \Rightarrow \gamma = 0.5244$ **and** $Z = 0.9588$ | | | | | | | | | |
| Estimated standard errors | | | | | | | | | |
| n | min | median | max | mean | min | median | max | mean | average k |
| 1000 | 0.0034 | 0.0053 | 0.0182 | 0.0058 | 0.0036 | 0.0071 | 0.2355 | 0.0076 | 80 |
| 2000 | 0.0027 | 0.0041 | 0.0159 | 0.0045 | 0.0029 | 0.0052 | 0.1098 | 0.0054 | 111 |
| 4000 | 0.0021 | 0.0031 | 0.0148 | 0.0035 | 0.0023 | 0.0037 | 0.0115 | 0.0038 | 138 |
| Estimated coverage probabilites | | | | | | | | | |
| n | 0.9000 | 0.9500 | 0.9750 | 0.9900 | 0.9000 | 0.9500 | 0.9750 | 0.9900 | max cov |
| 1000 | 0.8141 | 0.8833 | 0.9314 | 0.9542 | 0.8683 | 0.9280 | 0.9655 | 0.9774 | 0.9988 |
| 2000 | 0.8296 | 0.8949 | 0.9414 | 0.9605 | 0.8828 | 0.9394 | 0.9735 | 0.9851 | 0.9999 |
| 4000 | 0.8440 | 0.9054 | 0.9457 | 0.9630 | 0.8896 | 0.9437 | 0.9789 | 0.9891 | 1.0000 |
| **Zenga distribution with** $\alpha = 0.4113$ **and** $\theta = 8.7133 \Rightarrow \gamma = 0.7086$ **and** $Z = 0.9887$ | | | | | | | | | |
| Estimated standard errors | | | | | | | | | |
| n | min | median | max | mean | min | median | max | mean | average k |
| 1000 | 0.0002 | 0.0023 | 0.0057 | 0.0024 | 0.0013 | 0.0046 | 0.0434 | 0.0052 | 187 |
| 2000 | 0.0000 | 0.0018 | 0.0046 | 0.0020 | 0.0012 | 0.0036 | 0.0516 | 0.0041 | 64 |
| 4000 | 0.0002 | 0.0015 | 0.0043 | 0.0016 | 0.0009 | 0.0027 | 0.0302 | 0.0032 | 71 |
| Estimated coverage probabilites | | | | | | | | | |
| n | 0.9000 | 0.9500 | 0.9750 | 0.9900 | 0.9000 | 0.9500 | 0.9750 | 0.9900 | max cov |
| 1000 | 0.6555 | 0.7359 | 0.8158 | 0.8567 | 0.7544 | 0.7815 | 0.8041 | 0.8155 | 0.8503 |
| 2000 | 0.6598 | 0.7431 | 0.8189 | 0.8620 | 0.8402 | 0.8738 | 0.8975 | 0.9073 | 0.9403 |
| 4000 | 0.6723 | 0.7478 | 0.8173 | 0.8562 | 0.8809 | 0.9119 | 0.9347 | 0.9459 | 0.9799 |

## References

Cheng S. and Peng L. (2001). Confidence Intervals for the Tail Index. *Bernoulli*, 7, 751-760.

Csrgo, M., Csrg, S., Horvth, L., Mason, D. M.(1986). Weighted empirical and quantile processes. *Ann. Probab.*, 14 (1), 31–85.

de Haan L. and Ferreira A. (2006). *Extreme Value Theory: An Introduction.* Springer-Verlag, New York.

Greselin, F. and Pasquazzi, L. (2009). Asymptotic confidence intervals for a new inequality measure. *Communications in Statistics: Computation and Simulation*, 38(8), 17-42.

Greselin, F., Puri, M.L. and Zitikis, R. (2009). $L$-functions, processes, and statistics in measuring economic inequality and actuarial risks. *Statistics and Its Interface*, 2, 227–245.

Mason D. M. (1982). Laws of Large Numbers for Sums of Extreme Values *The Annals of Probability* 10 (3), 754-764.

Necir, A., Rassoul, A. and Zitikis, R. (2010). Estimating the conditional tail expectation in the case of heavy-tailed losses. *Journal of Probability and Statistics* (Special Issue on "Actuarial and Financial Risks: Models, Statistical Inference, and Case Studies").

Weissman, I. (1978). Estimation of parameters and larger quantiles based on the k largest observations. *Journal of the American Statistical Society*, 73 (364), 812-815.

Zenga, M. (2007). Inequality curve and inequality index based on the ratios between lower and upper arithmetic means. *Statistica & Applicazioni* 5, 3–27.

Zenga, M. (2010). Mixture of Polisicchios truncated Pareto distributions with beta weights. *Statistica & Applicazioni* 6, 3–25.

Zenga, M., Pasquazzi, L. and Zenga, M.A. (2010a). First applications of a new three-parameter distribution for non-negative variables. *Technical report*, Dipartimento di Metodi Quantitativi per le Scienze Economiche e Aziendali, Universita' di Milano Bicocca Italy. http://hdl.handle.net/10281/17464

Zenga, M., Polisicchio, M., Zenga, M.A., and Pasquazzi, L. (2010b). More on M.M. Zengas new three-parameter distribution for non-negative variables *Technical report*, Dipartimento di Metodi Quantitativi per le Scienze Economiche e Aziendali, Universita' di Milano Bicocca Italy. http://hdl.handle.net/10281/17465

Zitikis, R. (1998). The Vervaat process. In: *Asymptotic Methods in Probability and Statistics*, pp. 667–694. North-Holland, Amsterdam.