# HIER

# Harvard Institute of Economic Research

Discussion Paper Number 2106

## LOG(RANK-l/2): A SIMPLE WAY TO IMPROVE THE OLS ESTIMATION OF TAIL EXPONENTS

by

Xavier Gabaix
and
Rustam Ibragimov

February 2006

HARVARD UNIVERSITY
Cambridge, Massachusetts

This paper can be downloaded without charge from:
http://post.economics.harvard.edu/hier/2006papers/2006list.html

The Social Science Research Network Electronic Paper Collection:
http://ssrn.com/abstract=881759

# LOG(RANK−1/2): A SIMPLE WAY TO IMPROVE THE OLS ESTIMATION OF TAIL EXPONENTS[1]

Xavier Gabaix

*Department of Economics, MIT, and NBER*

Rustam Ibragimov

*Department of Economics, Harvard University*

January 4, 2006

## ABSTRACT

A popular way to estimate a Pareto exponent is to run an OLS regression: $\log(\text{Rank}) = c - b\log(\text{Size})$, and take $b$ as an estimate of the Pareto exponent. Unfortunately, this procedure is strongly biased in small samples. We provide a simple practical remedy for this bias, and argue that, if one wants to use an OLS regression, one should use the *Rank* $-1/2$, and run $\log(\text{Rank} - 1/2) = c - b\log(\text{Size})$. The shift of 1/2 is optimal, and cancels the bias to a leading order. The standard error on the Pareto exponent is not the OLS standard error, but is asymptotically $(2/n)^{1/2}b$. To obtain this result, we provide asymptotic expansions for the OLS estimate in such log-log rank-size regression with arbitrary shifts in the ranks. The arguments for the asymptotic expansions rely on strong approximations to martingales with the optimal rate and demonstrate that martingale convergence methods provide a natural and conceptually simple framework for deriving the asymptotics of the tail index estimates using the log-log rank-size regressions.

*KEYWORDS:* power law, heavy-tailedness, OLS log-log rank-size regression, standard errors, strong approximation, martingale convergence, $U-$statistics, bilinear forms

*JEL Classification:* C13, C14, C16

---

[1]xgabaix@mit.edu, ribragim@fas.harvard.edu. The authors are grateful to Gary Chamberlain, Victor Chernozhukov, Graham Elliott, Yoko Konishi, Hu McCulloch, Marcelo Moreira, Yoshihiko Nishiyama, Peter Phillips, Sidney Resnick, Jim Stock, Charles van Marrewijk and the participants at seminars at the Departments of Economics at Harvard University and Massachusetts Institute of Technology for helpful comments and discussions.

# 1 Introduction

## 1.1 Objectives and key results

Last four decades have witnessed rapid expansion of the study of heavy-tailedness phenomena in economics and finance. Following the pioneering work by Mandelbrot (1963) (see also the papers in Mandelbrot, 1997; Fama, 1965), numerous studies have documented that time series encountered in many fields in economics and finance are typically thick-tailed and can be well approximated using distributions with tails exhibiting the power law decline

$$P\Big(Z > x\Big) \sim kx^{-\xi}, \quad k, x > 0. \tag{1}$$

with a tail index $\xi > 0$ (see the discussion in Jansen and de Vries, 1991; Loretan and Phillips, 1994; Brakman, Garretsen, van den Berg and van Marrewijk, 1999; Meerschaert and Scheffler, 2000; Gabaix, Gopikrishnan, Plerou and Stanley, 2003; Ibragimov, 2004, 2005*a*; Nishiyama and Osada, 2005, and references therein).[2][3]

Let

$$Z_{(1)} \geq ... \geq Z_{(n)} \tag{2}$$

be decreasingly ordered observations from a population satisfying power law (1). A popular and commonly used approach to estimating the heavy-tailedness parameter $\xi$ is the one based on the following OLS log-log rank-size regression with $\gamma = 0$:

$$\log\,(t - \gamma) = a - b \log\; Z_{(t)}, \tag{3}$$

or, in other words, calling $t$ the rank of an observation, and $Z_{(t)}$ its size:

$$\log\,(\text{Rank} - \gamma) = a - b \log\,(\text{Size}) \tag{4}$$

(here and throughout the paper, $\log(\cdot)$ stands for the natural logarithm).[4]

[2]Here $f(x) \sim g(x)$ means that $f(x) = g(x)(1 + o(1))$ as $x \to \infty$.

[3]The reader is referred to, among other works, Duffie and Pan (1997); Uchaikin and Zolotarev (1999); Glasserman, Heidelberger and Shahabuddin (2002); De Vries (2005); Ibragimov (2004, 2005*a*); Hartmann, Straetmans and De Vries (2005); Ortobelli and Rachev (2005); Rachev, Jašic, Stoyanov and Fabozzi (2005); Ibragimov and Walden (2006) for the analysis of a number of economic and financial models under heavy-tailed assumptions, and to Loretan and Phillips (1994); Quintos, Fan and Phillips (2001); Chen, Härdle and Spokoiny (2005); Chernozhukov (2005*a*,*b*); Dufour, Khalaf, Kurz-Kim and Beaulieu (2005); Garcia, Renault and Veredas (2005); Haas, Mittnik, Paolella and Steude (2005); Haug, Klüppelberg, Lindner and Zapp (2005); Ibragimov (2005*b*); Kurz-Kim, Rachev, Samorodnitsky and Stoyanov (2005); Levy and Taqqu (2005); Lombardini and Calzolari (2005); McCulloch (2005); Nishiyama and Osada (2005); Nolan (2005); Samarakoon and Knight (2005) for recent studies of problems of statistical inference and modeling for data from thick-tailed populations and extremal phenomena.

[4]In this paper, we also consider the asymptotic asymptotic expansions using the dual to (4) regressions $\log\,(\text{Size}) = c - d \log\,(\text{Rank} - \gamma)$ with logarithms of ordered sizes regressed on logarithms of shifted ranks. As follows from Theorem 1, the approaches to the tail index inference using regressions (4) and their above dual analogues are equivalent in terms of the small sample biases and standard errors of the estimates.

Let $\hat{b}_n$ denote the usual OLS estimate of the tail index $\xi$ using regression (3) with $\gamma = 0$ and let $\hat{b}_n^\gamma$ denote OLS estimate of $\xi$ in general regression (3).

It is known that the OLS estimate $\hat{b}_n$ in the usual regression (3) with $\gamma = 0$ is consistent for $\xi$. However, the standard OLS procedure has an important bias. This paper shows that the bias is canceled (up to leading order terms) with $\gamma = 1/2$. Hence, we propose that always, if one uses a log-log regression, one should use $\log(Rank - 1/2)$ rather than log(Rank). Also, extending the work of Kratz and Resnick (1996) and Csörgő and Viharos (1997) (see also Viharos, 1999; Csörgő and Viharos, 2006) for the case $\gamma = 0$, we show that the standard error of the OLS estimate $\hat{b}_n^\gamma$ of the tail index $\xi$ in general regression (3) equals to $(2/n)^{1/2}\xi$.

The following tables provide comparisons of the standard errors of the traditional OLS estimate $\hat{b}_n$ of the tail index in standard regression (3) with $\gamma = 0$ with that of the estimate $\hat{b}_n^\gamma$ with $\gamma = 1/2$ recommended in the present paper.

**Table 1. Behavior of the traditional OLS estimate $\hat{b}_n$ in the regression** $\log(\textbf{Rank}) = a - b\log(\textbf{Size})$

| $n$ | 50 | 100 | 200 | 500 |
|---|---|---|---|---|
| Mean $\hat{b}_n^{\gamma=0}$ | 0.92 | 0.94 | 0.96 | 0.98 |
| Nom. s.e. | 0.023 | 0.013 | 0.0078 | 0.0037 |
| True s.e. | 0.20 | 0.14 | 0.098 | 0.063 |

Caption: Monte Carlo simulations, based on a power law distribution $P(Z > x) = x^{-1}$ for $x > 1$. For a general power law distribution with exponent $\xi$, one multiplies all the numbers in the table by $\xi$. "Mean $\hat{b}_n^{\gamma=0}$" is the mean of the estimate $\hat{b}_n$ in the regression: $\log(\text{Rank}) = a - b\log(\text{Size})$. "Nom s.e." is the nominal standard error from the OLS regression.

**Table 2. Behavior of the OLS estimate $\hat{b}_n^\gamma$ with $\gamma = 1/2$ in the regression** $\log\left(\textbf{Rank} - \frac{1}{2}\right) = a - b\log(\textbf{Size})$

| $n$ | 50 | 100 | 200 | 500 |
|---|---|---|---|---|
| Mean $\hat{b}_n^{\gamma=1/2}$ | 1.01 | 1.00 | 1.00 | 1.00 |
| Mean $\sqrt{2/n}\hat{b}_n^\gamma$ | 0.20 | 0.14 | 0.100 | 0.063 |
| True s.e. | 0.20 | 0.14 | 0.098 | 0.063 |

Caption: Monte Carlo simulations, based on a power law distribution $P(Z > x) = x^{-1}$ with $\xi = 1$ for $x > 1$. For a general power law distribution with exponent $\xi$, one multiplies all the numbers in the table by $\xi$. "Mean $\hat{b}_n^{\gamma=1/2}$" is the mean of the estimate $\hat{b}_n^{\gamma}$ with $\gamma = 1/2$ in the regression: $\log(\text{Rank} - 1/2) = a - b\log(\text{Size})$.

The OLS estimates $\hat{b}_n$ of $b$ using the regression $\log(\text{Rank}) = a - b\log(\text{Size})$ reported in Table 1 are significantly different from 1, which means that $\hat{b}_n$ is biased in small samples. According to the same table, the nominal standard errors in the regression $\log(\text{Rank}) = a - b\log(\text{Size})$ are consistently lower than the true standard errors (see Nishiyama and Osada 2005 for an analysis of this phenomenon). On the other hand, the estimates $\hat{b}_n^{\gamma}$ with $\gamma = 1/2$ in the regression $\log(\text{Rank} - 1/2) = a - b\log(\text{Size})$ reported in Table 2 are close to 1. That is, with the shifted rank we recommend, the bias of the usual OLS estimate $\hat{b}_n$ is considerably reduced. Furthermore, the asymptotic standard errors in the regression with the shifts $\gamma = 1/2$ in ranks are very close to the true standard errors.

The 1/2 shift actually comes from a more systematic result, in Theorem 1, which shows that it is optimal and further demonstrates that the following asymptotic expansion holds for the general OLS estimate $\hat{b}_n^{\gamma}$:

$$\hat{b}_n^{\gamma}/\xi = 1 + \sqrt{\frac{2}{n}}\mathcal{N}(0,1) + \frac{(\log n)^2(2\gamma - 1)}{4n} + o_P\left(\frac{\log^2 n}{n}\right)$$

(here and throughout the paper, $\mathcal{N}(0,1)$ stands for a standard normal r.v.).

We conclude that, for estimation of the tail index $\xi$, one should always use the regression

$$\log(\text{Rank} - 1/2) \quad = \quad a - b\log(\text{Size})$$

with the standard error of the OLS estimate of the slope given by $\sqrt{\frac{2}{n}}\xi$.

According to the results obtained in the paper, the nominal standard errors reported in OLS log-log rank-size regressions (3) considerably underestimate the true standard errors. Consequently, taking the OLS estimates of the standard errors at the face value will lead one to reject the true numerical value of the tail index too often. The reason for these conclusions is that ordering of the observations in (2) naturally generates *dependence* among the regressors in log-log rank-size model (3) employed to estimate the true tail index $\xi$. The standard asymptotics for the OLS regressors, on the other hand, treats them as being *independent*. The asymptotic expansions obtained in the paper also provide the correct confidence intervals for the tail index estimates. They further show that the OLS approach to the inference on the tail shape parameter is more robust relative to Hill's estimator in the case of deviations from the exact power law.

The rest of the paper formalizes the above claims. The arguments for the results in the paper are based, in a large part, on strong approximations to partial sums with the optimal rate (see Remark 3). They are also closely related to the martingale approaches to obtaining asymptotic results for bilinear forms and $U-$statistics recently developed in Ibragimov and Phillips (2004) (see Remark 4). This relation has an independent interest.

## 1.2 Estimation of the tail index and the OLS regression

At present, several approaches to the inference about the tail index $\xi$ of heavy-tailed distributions are available (see, among others, the review in Embrechts, Klupperberg and Mikosch, 1997). The two most commonly used ones are Hill's estimator introduced in Hill (1975) and the OLS approach using the log-log rank-size regression described in the previous subsection whose graphical analogues go back to Pareto.

It was reported in a number of studies that inference on the tail index using Hill's estimator suffers from several problems. For instance, it is well-known that Hill's estimator tends to severely overestimate the true tail index in relatively small stable samples. For instance, as reported in, e.g., McCulloch (1997) and Weron (2001), Hill's estimates of the tail index $\xi$ greater than two are to be expected for heavy-tailed stable distributions with infinite variances (for which $\xi < 2$) for relatively small samples of sizes $n < 1000$. In addition, Hill's estimator is very sensitive to the dependence in the data (see Embrechts et al., 1997, Chapter 6). Naturally, since it is based on the extreme observations (the greatest order statistics) in the data, the estimator requires very large sample sizes to exhibit convergence to the gaussian distribution. In fact, as indicated in Embrechts et al. (1997), the rate of convergence of Hill's estimator can be *arbitrary* slow and, furthermore, the choice of the number of the order statistics to be included in the estimator is problematic because of the important bias-variance tradeoff involved. The reader is referred to Hall (1982); Hall and Welsh (1984); Haeusler and Teugels (1985); Csörgő and Mason (1985); Beirlant and Teugels (1989); Embrechts, Klupperberg and Mikosch (1997); Danielsson, de Haan, Peng and de Vries (2001); Beirlant, Goegebeur, Teugels and Segers (2004) and the works cited therein for the discussion of the asymptotics of Hill's estimator.

Motivated by the above problems with Hill's estimator, several studies have focused on the alternative approaches to the tail index estimation. For instance, Huisman, Koedijk, Kool and Palm (2001) proposed a weighted analogue of Hill's estimator that was reported to correct the small sample bias of the latter statistical technique for sample sizes less than 1000. Recently, using Monte-Carlo methods, Dufour and Kurz-Kim (2005) obtained exact finite sample confidence intervals for Hill's estimator. Dufour and Kurz-Kim (2005) further proposed Monte-Carlo based estimation and exact testing procedures for the stability parameter of $\alpha-$stable distributions.

Embrechts et al. (1997), Beirlant, Dierckx, Goegebeur and Matthys (1999*b*), Jansen and de Vries (1991) and Feurverger and Hall (1999) advocated sophisticated non-linear procedures to the tail parameter estimation based on the direct estimation of the coefficients in expansions for the tail probabilities involving terms of smaller order in addition to the one in (1).

Despite of the availability of the above sophisticated approaches to the inference on the tail parameter, the OLS log-log rank-size regression (3) with $\gamma = 0$ remains the most widely applied alternative to Hill's estimator, for the most part due to its simplicity and its robustness.[5] The simplicity of the OLS approach to the tail index estimation, together with its easy implementability on virtually any statistical software is very appealing. This, together with the widespread of the OLS approach to the tail index estimation among the practitioners motivates the analysis of the statistical properties of the estimators $\hat{b}_n^\gamma$.

In recent years, several studies have focused on the analysis of normality of the OLS tail index estimate in a "dual" to (3) with $\gamma = 0$ model with logarithms of ordered observations $log(Z_{(t)})$ regressed on logarithms of ranks. Such approach to estimation of the tail shape parameters was introduced by Kratz and Resnick (1996) who refer to it as QQ-estimator. Nishiyama and Osada (2005) discussed asymptotic normality of the OLS tail index estimate in the regression of $log(Z_{(t)})$ on $\log\ t$ and also proposed more efficient than the OLS estimation procedures based on generalized least squares method and on a trimmed least squares regression. Schultze and Steinebach (1999) considered closely related problems of least-squares approaches to estimation for data with exponential tails (see also Aban and Meerschaert, 2004, who discuss efficient OLS estimation of parameters in shifted and scaled exponential models). Kratz and Resnick (1996) establish consistency and asymptotic normality of the QQ-estimator in the case of populations with regularly varying tails. Their results demonstrate that in the case of populations in the domain of attraction of power law (1), the standard error of the QQ-estimator of the inverse $1/\xi$ of the tail index based on $n$ largest observations is $\sqrt{2}\xi/\sqrt{n}$. Csörgő and Viharos (1997) prove asymptotic normality of the OLS estimates of the tail index (see also Viharos, 1999; Csörgő and Viharos, 2006). The latter OLS approach to the tail index estimation is closely related to kernel smoothed Hill's estimators studied in Csörgő, Deheuvels and Mason (1985) and Groeneboom, Lopuhaä and de Wolf (2003). Bias reduction using shifted log-ranks $\log(Rank - 1/2)$ in (4) proposed in this paper is close in spirit to approaches to small sample bias correction in econometric models based on Edgeworth expansions and bootstrap (see, among other works, Nishiyama and Robinson, 2000, 2005).

The above-discussed dual to (3) models with logarithms of order statistics $x_t$ regressed on logarithms of ranks $y_t$ fall into the framework of regressions with slowly integrated regressors. Therefore, the asymptotics for it can be derived using the approach and the results developed

[5]See, e.g., examples and discussion in Gabaix et al. (2003) and Gabaix and Ioannides (2004).

in Phillips (2001).

However, to our knowledge, the complete asymptotic theory for the OLS estimates for the tail estimation using regressions (3) with logarithms of ordered observations as dependent variables is not available in the literature. The asymptotic theory for such most commonly used log-log rank-size regressions is necessarily much more complicated than that for its above analogues due to the dependence in regressors. This dependence implies that the statistics of interest involve $U-$statistics and bilinear forms in weighted independent r.v.'s rather than their sums (see Section 2). In particular, the CLT's for sums of independent r.v.'s cannot be applied in the present instance and one has to appeal to more involved approaches to deriving asymptotics under dependence in random summands.

In fact, according to the Rényi representation theorem (see the proof of Theorem 1), under the null hypothesis that (unordered) observations $Z_t'$s are from a population with power-law distribution (1), the structure of the above dependence is similar to that in a regression with integrated regressors (see Park and Phillips, 1999, 2001; Ibragimov and Phillips, 2004). Namely, similar to the wide class of nonlinear statistics whose asymptotics was studied in Ibragimov and Phillips (2004), under the null, the numerator of the OLS estimates $\hat{b}_n^\gamma$ has the form of a second-order $U-$statistic, more precisely, a bilinear form in weighted exponential r.v.'s. This makes the power of strong approximations to partial sums of martingale-differences and the new martingale convergence machinery for the analysis of weak convergence of econometric estimators recently developed in Ibragimov and Phillips (2004) applicable in the treatment of the asymptotics of estimates in (3) and in its more general analogues (see Remarks 3 and 4).

One of the important contributions of the present paper is that our theoretical results demonstrate, in fact, that the strong approximations to martingales and, more generally, martingale convergence methods provide a natural and conceptually simple framework for deriving the asymptotics of the tail index estimates using the log-log rank-size regression. One needs to emphasize here that, under the null hypothesis, the $U-$statistic in the numerator of the OLS estimates $\hat{b}_n^\gamma$ converges to a deterministic integral of the Wiener process, but not to a stochastic one, as its typically the case for the typical bilinear forms $\sum_{t=1}^n \left(\sum_{i=1}^{t-1} v_i\right) u_t$ in (possibly correlated) linear processes $u_t$ and $v_t$ (see Ibragimov and Phillips, 2004). The main reason is that the decline of the coefficients in the representation of $Z_{(t)}'$s as weighted sums of exponential r.v.'s assure that the second predictable characteristic of the statistic of interest tends to a deterministic process, namely to a deterministic function given by the quadratic variation of the deterministic integral of a Wiener process.

## 2 Asymptotic expansions for the log-log rank-size tail estimator via strong approximations to partial sums

Let $Z_{(1)} \geq Z_{(2)} \geq ... \geq Z_{(n)}$ be the order statistics for a sample from the population with the distribution satisfying power law

$$P(Z > x) = \frac{1}{x^{\xi}},\ x \geq 1. \tag{5}$$

Denote $y_t = \log(t-\gamma)$ and $x_t = \log(Z_{(t)})$. Let us consider the OLS estimate $\hat{b}_n^{\gamma}$ of the slope parameter $b$ in log-log rank-size regression (3) with $\gamma < 1$, that is, in the model

$$y_t = a - bx_t \tag{6}$$

with logarithms of ordered observations regressed on logarithms of shifted ranks:

$$\hat{b}_n^{\gamma} = -\frac{\sum_{t=1}^{n}(x_t - \overline{x}_n)(y_t - \overline{y}_n)}{\sum_{t=1}^{n}(x_t - \overline{x}_n)^2} = -\frac{A_n^{\gamma}}{B_n}. \tag{7}$$

We will also consider the OLS estimate $\hat{d}_n^{\gamma}$ of slope in the dual to (3) regression

$$x_t = c - dy_t \tag{8}$$

with logarithms of ordered sizes regressed on logarithms of shifted ranks:

$$\hat{d}_n^{\gamma} = -\frac{\sum_{t=1}^{n}(x_t - \overline{x}_n)(y_t - \overline{y}_n)}{\sum_{t=1}^{n}(y_t - \overline{y}_n)^2} = -\frac{A_n^{\gamma}}{D_n}. \tag{9}$$

The following theorem provides the main result of the present paper.

**Theorem 1** *For any $\gamma < 1$, the following expansions hold:*

$$\hat{b}_n^{\gamma}/\xi = 1 + \sqrt{\frac{2}{n}}\mathcal{N}(0,1) + \frac{(\log n)^2\,(2\gamma - 1)}{4n} + o_P\Big(\frac{\log^2 n}{n}\Big), \tag{10}$$

$$\xi\hat{d}_n^{\gamma} = 1 + \sqrt{\frac{2}{n}}\mathcal{N}(0,1) + \frac{(\log n)^2\,(1 - 2\gamma)}{4n} + o_P\Big(\frac{\log^2 n}{n}\Big). \tag{11}$$

**Remark 1** *As follows from asymptotic expansions (10) and (11), the small sample biases of the OLS estimates* $\hat{b}_n^\gamma$ *and* $\hat{d}_n^\gamma$ *in regressions* $y_t = a - bx_t$ *and* $x_t = c - dy_t$ *involving logarithms of shifted ranks are both minimized under the choice* $\gamma = 1/2$.[6]

**Remark 2** *As discussed in the introduction, from Theorem 1 it follows that the nominal standard errors in the log-log rank-size regression and in its more general analogues (10) and (11) considerably underestimate the true standard errors. Consequently, making the inferences on the tail parameter values on the basis of the OLS estimates of the standard errors will lead one to reject the true value of the tail index too often.*

Throughout the paper, $W$ denotes the standard Brownian motion.

**Remark 3** *The proof of asymptotic expansions (10) and (11) is based, in a large part, on strong approximations to partial sums of independent r.v.'s* $S_t = \sum_{i=1}^t \tau_i$. *Furthermore, as follows from the argument for Theorem 1, derivation of the expansion is based on an application of strong approximations with the best possible rate* $\sup_{1\le t\le n}\left|\frac{S_{t-1}}{\sqrt{n}} - W\left(\frac{t-1}{n}\right)\right| = O\left(\frac{\log n}{n}\right)$ *(a.s.) Such approximations hold under the assumption of the existence of moment generating function in a neighborhood of zero and are provided by the results in Komlós, Major and Tusnády (1975, 1976); Csörgö and Révész (1981, Theorem 2.6.1). Due to the Rényi representation theorem (see Subsection 1.2 and relation (19) in the proof of Theorem 1 below), the ordered observations from a population with the distribution satisfying power law can be represented as weighted sums of i.i.d. exponential r.v.'s. The fact that these r.v.'s have exponential moments ensures that the above strong approximations with best error order are applicable in the study of asymptotic expansions for the tail index estimates. One should also emphasize here that approximations in the form* $\sup_{1\le t\le n}\left|\frac{S_{t-1}}{\sqrt{n}} - W\left(\frac{t-1}{n}\right)\right| = O(n^{1/p-1/2})$, $p > 2$ *(a.s.) hold for partial sums* $S_t = \sum_{i=1}^t \xi_i$ *of i.i.d. r.v.'s* $\xi_i$ *under the weaker assumption than above, namely, under the condition* $E|\xi_i|^{2p} < \infty$. *The latter approximations have been used in a number of recent works (see, among others, Phillips, 1999; Ibragimov and Phillips, 2004, and references therein) to develop asymptotic results for various functionals of independent r.v.'s and linear processes. As follows from the proof of Theorem 1, these approximation results will not work in the case of asymptotic expansions in the form (10) and (11) as the argument for them requires the use of error bounds of better*

[6] Beirlant et al. (1999*a*) and Aban and Meerschaert (2004) indicate the possibility of modification of the QQ-estimator discussed in Subsection 1.2 in which logarithms of ordered observations $\log(Z_{(t)})$ regressed on $\log(t-1/2)$. Aban and Meerschaert (2004) mention in a remark without providing a proof that regressing logarithms of observations from a heavy-tailed population on logarithms of their ranks shifted by 1/2 reduces the bias of the QQ-estimator. Their remark seems motivated by simulations, not by the systematic understanding that Theorem 1 provides, in particular they do not indicate that a shift of 1/2 is the best shift.

*order.*[7]

**Remark 4** *Due to the the Rényi representation theorem, the OLS log-log rank-size approach to the inference on the tail index in model (3) is similar to the problem of statistical inference in a regression with integrated regressors (see Park and Phillips, 1999, 2001; Ibragimov and Phillips, 2004). Furthermore, the normalized estimation errors* $\hat{b}_n^\gamma - 1 = -(A_n^\gamma + B_n)/B_n$ *in OLS log-log rank-size regression (3) can be approximated by martingale processes. This allows one to derive asymptotic results such as*

$$\sqrt{\frac{2}{n}}\Big(\hat{b}_n^\gamma - 1\Big) \to_d \mathcal{N}(0,1) \tag{12}$$

*using the general martingale convergence approach to deriving weak convergence of bilinear forms and* $U-$*statistics to stochastic integrals (here and in what follows,* $\to_d$ *denotes convergence in distribution; further, the notation* $X =_d Y$ *for two r.v.'s* $X$ *and* $Y$ *means that their distributions are the same). Such a new approach to developing asymptotic results was developed recently in Ibragimov and Phillips (2004) and was applied by the authors to develop unification of the treatment of the asymptotics for stationary autoregression, autoregression with roots at or near unity and for a wide range of nonlinear econometric models. The martingale convergence approach to obtaining normal convergence for* $\hat{b}_n^\gamma$ *can be based on the easy to establish approximation*

$$\frac{1}{\sqrt{n}}(\xi A_n + \xi^2 B_n) = M_n\Big(\frac{n-1}{n}\Big) + o_P(1), \tag{13}$$

*where* $M_n(s)$ *is the continuous-time martingale given by the sum of* $U-$*statistics*

$$M_n(s) = \frac{1}{\sqrt{n}}\sum_{t=1}^{[ns]}\Big(\tau_t - 1\Big)\Big[-\log\,(t/n) + \frac{2}{t}\sum_{i=1}^{t-1}\Big(\tau_i - 1\Big) - \frac{2}{n}\sum_{i=1}^{t-1}\Big(\tau_i - 1\Big)\Big] \tag{14}$$

*in i.i.d. exponential r.v.'s* $\tau_t$ *with parameter 1. The methods presented and discussed in details in Ibragimov and Phillips (2004) allows one to reduce the study of distributional convergence of the process* $M_n(s)$ *to the analysis of convergence in probability of its predictable characteristics. The study of the asymptotics for the predictable characteristics of* $M_n(s)$ *allows one to show, similar to the results in Sections 5 and 6 in Ibragimov and Phillips (2004) that*

$$M_n(s) \to_d M(s) = \int_0^s \log u\, dW(u). \tag{15}$$

[7] In particular, the strong approximations with the best order of error are crucial, together with the results on the modulus of continuity of Brownian sample paths, in obtaining the term of order $O_P\Big(\frac{(\log\ n)^{3/2}}{\sqrt{n}}\Big) = o_P\Big(\frac{(\log\ n)^2}{\sqrt{n}}\Big)$ in approximation (44) in the proof of Theorem 1.

*Using (15), together with representation (13) and the fact that* $\frac{n-1}{n} \to 1$ *as* $n \to 1$*, one obtains*

$$\frac{1}{\sqrt{n}}(\xi A_n + \xi^2 B_n) \to_d M(1) = \int_0^1 \log u \, dW(u) =_d \mathcal{N}(0,2)$$

. *The latter relation, together with the easy to establish convergence* $\frac{\xi^2 B_n}{n} \to_P 1$*, implies that (12) indeed holds. It is important to note here that the martingale approximating the numerator of the normalized errors in the OLS log-log rank-size regression converges to a deterministic integral* $\int_0^s \log u \, dW(u)$ *of the Wiener process, but not to a stochastic one, as its typically the case for the typical bilinear forms* $\sum_{t=1}^n f\Big(\sum_{i=1}^{t-1} v_i\Big) u_t$ *in (possibly correlated) linear processes* $u_t$ *and* $v_t$ *(see Ibragimov and Phillips, 2004). The latter bilinear forms typically have stochastic integrals such as* $\int_0^1 f(V(s))dW(s)$ *with (possibly correlated) Wiener processes* $V$ *and* $W$ *as their weak limits. The main reason for the above is that the fast decline of the coefficients in the representation of the regressors* $x_t'$*s in (3) as weighted sums of exponential r.v.'s assure that the second predictable characteristic of the martingale of interest tends to a deterministic process. Namely, it converges to a deterministic function* $\int_0^s \log^2 u \, du$ *which is the quadratic variation of the continuous process* $\int_0^s \log u \, dW(u)$*.*

Proof of Theorem 1.

It is enough to prove the Theorem in the case $\xi = 1$. For the general case, one applies the theorem to $Z' = Z^\xi$, whose power law exponent is 1.

Relation (10) for $\xi = 1$ is a consequence of (7) and the following asymptotic expansions for the statistics $A_n^\gamma$ and $B_n$ under $\xi = 1$ that we establish in turn:

$$\frac{1}{\sqrt{n}}(A_n^\gamma + B_n) = \mathcal{N}(0,2) + \frac{(\log n)^2 (1-2\gamma)}{4\sqrt{n}} + o_P\Big(\frac{\log^2 n}{\sqrt{n}}\Big) \tag{16}$$

and

$$\frac{B_n}{n} = 1 + O_P\Big(\frac{\log\ n}{\sqrt{n}}\Big). \tag{17}$$

We first focus on proving relation (16).

By the Rényi representation theorem, one has that, for the logarithms $x_t = \log\ S_{(t)}$ of ordered observations from a population with the distribution satisfying power law (5), the transformations

$$\tau_t = t\big(x_t - x_{t+1}\big), \quad t = 1, ..., n-1, \tag{18}$$

are i.i.d. exponential r.v.'s with parameter 1: $P(\tau_t > x) = \exp(-x)$, $x \geq 0$. That is one has one can represent the regressors in (3) as weighted sums of exponential r.v.'s in the following way:

$$x_t = x_n + z_t,\ t = 1, ..., n, \tag{19}$$

where $z_n = 0$ and $z_t = \sum_{i=t}^{n-1} \frac{\tau_i}{i}$, $t = 1, ..., n-1$.

We, therefore, get that

$$B_n = \sum_{t=1}^{n} (x_t - \overline{x}_n)^2 = \sum_{t=1}^{n} (x_n + z_t - x_n - \overline{z}_n)^2 = \sum_{t=1}^{n} (z_t - \overline{z}_n)^2 = \sum_{t=1}^{n-1} z_t^2 - n\overline{z}_n^2, \tag{20}$$

and, similarly,

$$A_n^\gamma = \sum_{t=1}^{n} (x_t - \overline{x}_n)(y_t - \overline{y}_n) = \sum_{t=1}^{n} (z_t - \overline{z}_n)(y_t - \overline{y}_n) = \sum_{t=1}^{n-1} z_t y_t - n\overline{z}_n \overline{y}_n. \tag{21}$$

We further have

$$\sum_{t=1}^{n-1} z_t^2 = \sum_{t=1}^{n-1} \Big( \sum_{i=t}^{n-1} \frac{\tau_i}{i} \Big)^2 = \sum_{t=1}^{n-1} \sum_{i=t}^{n-1} \frac{\tau_i^2}{i^2} + 2 \sum_{t=1}^{n-1} \sum_{i=t}^{n-2} \frac{\tau_i}{i} \sum_{j=i+1}^{n-1} \frac{\tau_j}{j} =$$
$$\sum_{i=1}^{n-1} \frac{\tau_i^2}{i} + 2 \sum_{j=2}^{n-1} \frac{\tau_j}{j} \sum_{i=1}^{j-1} \tau_i. \tag{22}$$

In addition,

$$n\overline{z}_n^2 = \frac{1}{n} \Big( \sum_{t=1}^{n-1} \sum_{i=t}^{n-1} \frac{\tau_i}{i} \Big)^2 = \frac{1}{n} \Big( \sum_{i=1}^{n-1} \tau_i \Big)^2 = \frac{1}{n} \sum_{i=1}^{n-1} \tau_i^2 + \frac{2}{n} \sum_{i=2}^{n-1} \tau_i \sum_{j=1}^{i-1} \tau_j. \tag{23}$$

By (20), (22) and (23) we get that

$$B_n = \sum_{t=1}^{n} (x_t - \overline{x}_n)^2 = \Big( \sum_{t=1}^{n-1} z_t^2 - n\overline{z}_n^2 \Big) =$$
$$\sum_{i=1}^{n-1} \frac{\tau_i^2}{i} + 2 \sum_{j=2}^{n-1} \frac{\tau_j}{j} \sum_{i=1}^{j-1} \tau_i - \frac{1}{n} \sum_{i=1}^{n-1} \tau_i^2 - \frac{2}{n} \sum_{i=2}^{n-1} \tau_i \sum_{j=1}^{i-1} \tau_j. \tag{24}$$

Similar to the above derivations, we have

$$\sum_{t=1}^{n-1} z_t y_t = \sum_{t=1}^{n-1} \log\,(t-\gamma)\Big(\sum_{i=t}^{n-1} \frac{\tau_i}{i}\Big) = \sum_{t=1}^{n-1} \frac{\tau_t}{t}\Big(\sum_{i=1}^{t} \log\,(i-\gamma)\Big), \tag{25}$$

and

$$n\overline{z}_n\overline{y}_n = \Big(\sum_{t=1}^{n-1}\sum_{i=t}^{n-1} \frac{\tau_i}{i}\Big)\Big(\frac{1}{n}\sum_{t=1}^{n} \log\,(t-\gamma)\Big) = \Big(\sum_{t=1}^{n-1} \tau_t\Big)\Big(\frac{1}{n}\sum_{t=1}^{n} \log\,(t-\gamma)\Big). \tag{26}$$

Relations (21), (25) and (26) imply

$$A_n^{\gamma} = \sum_{t=1}^{n-1} \frac{\tau_t}{t}\Big(\sum_{i=1}^{t} \log\,(i-\gamma)\Big) - \Big(\sum_{t=1}^{n-1} \tau_t\Big)\Big(\frac{1}{n}\sum_{t=1}^{n} \log\,(t-\gamma)\Big). \tag{27}$$

From (24) and (27) we get

$$\begin{aligned}\frac{1}{\sqrt{n}}(A_n^{\gamma}+B_n) = \frac{1}{\sqrt{n}}\Big[\sum_{t=1}^{n-1} \frac{\tau_t}{t}\Big(\sum_{i=1}^{t} \log\,(i-\gamma)\Big) - \Big(\sum_{t=1}^{n-1} \tau_t\Big)\Big(\frac{1}{n}\sum_{t=1}^{n} \log\,(t-\gamma)\Big) +\\ \sum_{i=1}^{n-1} \frac{\tau_i^2}{i} + 2\sum_{j=2}^{n-1} \frac{\tau_j}{j}\sum_{i=1}^{j-1} \tau_i - \frac{1}{n}\sum_{i=1}^{n-1} \tau_i^2 - \frac{2}{n}\sum_{i=2}^{n-1} \tau_i \sum_{j=1}^{i-1} \tau_j\Big].\end{aligned} \tag{28}$$

Using (28) and elementary algebraic manipulations, it is not difficult to show that the following representation holds:

$$\begin{aligned}\frac{1}{\sqrt{n}}(A_n^{\gamma}+B_n) = \frac{1}{\sqrt{n}}\sum_{t=1}^{n-1}(\tau_t-1)\Big[\frac{1}{t}\Big(\sum_{i=1}^{t} \log\,(i-\gamma)\Big) - \Big(\frac{1}{n}\sum_{i=1}^{n} \log\,(i-\gamma)\Big) +\\ 2\frac{1}{t}\sum_{i=1}^{t-1}(\tau_i-1) - \frac{2}{n}\sum_{i=1}^{t-1}(\tau_i-1) + 2\sum_{j=t+1}^{n-1}\frac{1}{j}\Big] +\\ \frac{1}{\sqrt{n}}\Big[n + \sum_{t=1}^{n} \frac{1}{t}\Big(\sum_{i=1}^{t} \log\,(i-\gamma)\Big) - \Big(\sum_{t=1}^{n} \log\,(t-\gamma)\Big)\Big] + \Big[\frac{1}{\sqrt{n}}\sum_{i=1}^{n-1} \frac{\tau_i^2}{i} - \frac{1}{n^{3/2}}\sum_{i=1}^{n-1} \tau_i^2\Big].\end{aligned} \tag{29}$$

We have

$$E\Big[\sum_{i=1}^{n-1}\frac{\tau_i^2}{i}\Big] = O(\log\ n)$$

and

$$E\Big[\sum_{i=1}^{n-1}\tau_i^2\Big] = O(n)$$

that imply

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n-1}\frac{\tau_i^2}{i} = O_P\Big(\frac{\log\ n}{\sqrt{n}}\Big) \tag{30}$$

and

$$\frac{1}{n^{3/2}}\sum_{i=1}^{n-1}\sum_{i=1}^{n-1}\tau_i^2 = O_P\Big(\frac{1}{\sqrt{n}}\Big). \tag{31}$$

In addition, it is not difficult to see that

$$Var\Big[\sum_{t=1}^{n-1}\frac{\tau_t - 1}{t}\sum_{i=1}^{t-1}(\tau_i - 1)\Big] = O\Big(\sum_{t=1}^{n}\frac{1}{t}\Big) = O(\log\ n).$$

This implies that

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n-1}\frac{\tau_t - 1}{t}\sum_{i=1}^{t-1}(\tau_i - 1) = O\Big(\frac{(\log\ n)^{1/2}}{\sqrt{n}}\Big). \tag{32}$$

Similarly, since, evidently,

$$Var\Big[\sum_{t=1}^{n-1}(\tau_t - 1)\sum_{i=1}^{t-1}(\tau_i - 1)\Big] = O(n^2)$$

we get

$$\frac{1}{n^{3/2}}\sum_{t=1}^{n-1}(\tau_t - 1)\sum_{i=1}^{t-1}(\tau_i - 1) = O\Big(\frac{1}{\sqrt{n}}\Big). \tag{33}$$

Let us now consider the term

$$G_n = \frac{1}{\sqrt{n}}\Big[n + \sum_{t=1}^{n} \frac{1}{t}\Big(\sum_{i=1}^{t} \log\,(i-\gamma)\Big) - \Big(\sum_{t=1}^{n} \log\,(t-\gamma)\Big)\Big]$$

appearing in (29).

Using Euler-Maclaurin summation formula (see, e.g., Havil, 2003, p. 86), we have

$$\sum_{i=1}^{t} \log\,(i-\gamma) = \int_1^t \log\,(x-\gamma)dx + \frac{\log\,(t-\gamma)}{2} + O(1) =$$
$$t\log\,(t-\gamma) - t + \Big(\frac{1}{2}-\gamma\Big)\log\,(t-\gamma) + O(1). \tag{34}$$

From (34) we get

$$G_n = \frac{1}{\sqrt{n}}\Big[n + \sum_{t=1}^{n} \log\,(t-\gamma) - n + \Big(\frac{1}{2}-\gamma\Big)\sum_{t=1}^{n}\frac{\log\,(t-\gamma)}{t} -$$
$$n\log\,(n-\gamma) + n - \Big(\frac{1}{2}-\gamma\Big)\log\,(n-\gamma) + O(\log\ n)\Big] =$$
$$\frac{1}{\sqrt{n}}\Big[n\log\,(n-\gamma) - n + \Big(\frac{1}{2}-\gamma\Big)\log\,(n-\gamma) + \Big(\frac{1}{2}-\gamma\Big)\sum_{t=1}^{n}\frac{\log\,(t-\gamma)}{t} -$$
$$n\log\,(n-\gamma) + n - \Big(\frac{1}{2}-\gamma\Big)\log\,(n-\gamma) + O(\log\ n)\Big] =$$
$$\frac{1}{\sqrt{n}}\Big(\frac{1}{2}-\gamma\Big)\sum_{t=1}^{n}\frac{\log\,(t-\gamma)}{t} + O\Big(\frac{\log\ n}{\sqrt{n}}\Big).$$

Applying integral approximations to partial sums, it is easy to see that, for all $\gamma < 1$,

$$\sum_{t=1}^{n}\frac{\log\,(t-\gamma)}{t} = \frac{(\log\ n)^2}{2} + o((\log\ n)^2).$$

The last relations imply that

$$G_n = \frac{(1-2\gamma)}{4}\frac{(\log\ n)^2}{\sqrt{n}} + o\Big(\frac{(\log\ n)^2}{\sqrt{n}}\Big). \tag{35}$$

From relations (29)-(33) and (35) it follows that the following asymptotic approximation holds:

$$\frac{1}{\sqrt{n}}(A_n^\gamma + B_n) = \frac{1}{\sqrt{n}}\sum_{t=1}^{n-1}(\tau_t - 1)\Big[\frac{1}{t}\Big(\sum_{i=1}^{t}\log\,(i-\gamma)\Big) -$$

$$\Big(\frac{1}{n}\sum_{i=1}^{n}\log\,(i-\gamma)\Big) + 2\sum_{j=t+1}^{n-1}\frac{1}{j}\Big] + \frac{(1-2\gamma)}{4}\frac{(\log\ n)^2}{\sqrt{n}} + o_P(\frac{(\log\ n)^2}{\sqrt{n}}) =$$

$$-\frac{1}{\sqrt{n}}\sum_{t=1}^{n-1}(\tau_t - 1)\log\,(t/n) + \frac{1}{\sqrt{n}}\sum_{t=1}^{n-1}(\tau_t - 1)\Big[\frac{1}{t}\Big(\sum_{i=1}^{t}\log\,(i-\gamma)\Big) -$$

$$\Big(\frac{1}{n}\sum_{i=1}^{n}\log\,(i-\gamma)\Big) + 2\sum_{j=t+1}^{n-1}\frac{1}{j} + \log\,(t/n)\Big] +$$

$$\frac{(1-2\gamma)}{4}\frac{(\log\ n)^2}{\sqrt{n}} + o_P\Big(\frac{(\log\ n)^2}{\sqrt{n}}\Big). \tag{36}$$

Let us show that

$$U_n = \frac{1}{\sqrt{n}}\sum_{t=1}^{n-1}(\tau_t - 1)\Big[\frac{1}{t}\Big(\sum_{i=1}^{t}\log\,(i-\gamma)\Big) -$$

$$\Big(\frac{1}{n}\sum_{i=1}^{n}\log\,(i-\gamma)\Big) + 2\sum_{j=t+1}^{n-1}\frac{1}{j} + \log\,(t/n)\Big] = o_P\Big(\frac{1}{\sqrt{n}}\Big). \tag{37}$$

We have

$$Var(nU_n) = \sum_{t=1}^{n-1}\Big[\frac{1}{t}\Big(\sum_{i=1}^{t}\log\,(i-\gamma)\Big) -$$

$$\Big(\frac{1}{n}\sum_{i=1}^{n}\log\,(i-\gamma)\Big) + 2\sum_{j=t+1}^{n-1}\frac{1}{j} + \log\,(t/n)\Big]^2. \tag{38}$$

Using relation (34), Taylor expansion for the logarithm and the fact that

$$\sum_{j=t+1}^{n-1}\frac{1}{j} = -\log\,(t/n) + O\Big(\frac{1}{t}\Big), \tag{39}$$

it is not difficult to see that

$$\sum_{t=1}^{n-1}\Big[\frac{1}{t}\Big(\sum_{i=1}^{t}\log\,(i-\gamma)\Big)-\Big(\frac{1}{n}\sum_{i=1}^{n}\log\,(i-\gamma)\Big)+2\sum_{j=t+1}^{n-1}\frac{1}{j}+\log\,(t/n)\Big]^2=$$

$$\sum_{t=1}^{n-1}\Big[\Big(\frac{1}{2}-\gamma\Big)\frac{\log\,(t-\gamma)}{t}+\Big(\frac{1}{2}-\gamma\Big)\frac{\log\,(n-\gamma)}{n}+O\Big(\frac{1}{t}\Big)\Big]^2=O(1). \tag{40}$$

Relations (38) and (40) imply that (37) indeed holds.

We now provide the argument for the convergence

$$-\frac{1}{\sqrt{n}}\sum_{t=1}^{n-1}(\tau_t-1)\log\,(t/n)\to_d\sqrt{2}\mathcal{N}(0,1) \tag{41}$$

using strong approximations to partial sums of r.v.'s by Brownian motion.

Using partial summation similar to the proof of Lemma 2.3 in Phillips (2001), we get (below, $S_t=\sum_{i=1}^{t}u_i$ and $u_i=\tau_i-1$)

$$-\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\log\,(t/n)u_t=-\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\log\;tu_t+\log\;n\frac{1}{\sqrt{n}}\sum_{t=1}^{n}u_t=$$

$$\Big[-\log\;n\frac{S_n}{\sqrt{n}}+\sum_{t=2}^{n}\big(\log\;t-\log\,(t-1)\big)\frac{S_{t-1}}{\sqrt{n}}\Big]+\log\;n\frac{S_n}{\sqrt{n}}=$$

$$\sum_{t=2}^{n}\big(\log\;t-\log\,(t-1)\big)\frac{S_{t-1}}{\sqrt{n}}.$$

Using (the best theoretically possible) strong approximation to partial sums of independent r.v.'s that holds under the assumption of the existence of moment generating function in a neighborhood of zero (see, e.g., Komlós, Major and Tusnády, 1975, 1976; Csörgö and Révész, 1981, Theorem 2.6.1) we get

$$\sup_{1\le t\le n}\Big|\frac{S_{t-1}}{\sqrt{n}}-W\big(\frac{t-1}{n}\big)\Big|=O\Big(\frac{\log\;n}{n}\Big)\;\;(a.s.). \tag{42}$$

From the above relations we get

$$\sum_{t=2}^{n}\Big(\log\;t-\log\,(t-1)\Big)\frac{S_{t-1}}{\sqrt{n}}=\sum_{t=2}^{n}\Big(\log\;t-\log\,(t-1)\Big)W\Big(\frac{t-1}{n}\Big)+$$

$$O\Big(\frac{\log\ n}{n}\Big)\sum_{t=2}^{n}\Big(\log\ t-\log\ (t-1)\Big)=\sum_{t=2}^{n}\Big(\log\ t-\log\ (t-1)\Big)W\Big(\frac{t-1}{n}\Big)+O\Big(\frac{\log^2\ n}{n}\Big).$$

Let us consider the difference between

$$\sum_{t=2}^{n}\Big(\log\ t-\log\ (t-1)\Big)W\Big(\frac{t-1}{n}\Big)=\sum_{t=2}^{n}\Big[\log\ \Big(n\frac{t}{n}\Big)-\log\ \Big(n\frac{t-1}{n}\Big)\Big]W\Big(\frac{t-1}{n}\Big)$$

and

$$\int_0^1 W(r)d\ \log\ (nr).$$

We have

$$\Big|\sum_{t=2}^{n}\Big(\log\ t-\log\ (t-1)\Big)W\Big(\frac{t-1}{n}\Big)-\int_0^1 W(r)d\ \log\ (nr)\Big|=$$

$$\Big|\sum_{t=2}^{n}\Big[\Big(\log\ t-\log\ (t-1)\Big)W\Big(\frac{t-1}{n}\Big)-\int_{(t-1)/n}^{t/n}W(r)d\log\ (nr)\Big]\Big|\leq$$

$$\sum_{t=2}^{n}\int_{(t-1)/n}^{t/n}\Big|W(r)-W\Big(\frac{t-1}{n}\Big)\Big|d\log(nr)\leq$$

$$\sup_{t_1,t_2:|t_2-t_1|\leq 1/n}\big|W(t_2)-W(t_1)\big|\sum_{t=2}^{n}\int_{(t-1)/n}^{t/n}d\log(nr)=$$

$$\sup_{t_1,t_2:|t_2-t_1|\leq 1/n}\big|W(t_2)-W(t_1)\big|\sum_{t=2}^{n}\big(\log\ t-\log\ (t-1)\big)=\log\ n\sup_{t_1,t_2:|t_2-t_1|\leq 1/n}\big|W(t_2)-W(t_1)\big|.$$

According to the results on the modulus continuity for Brownian sample paths (Karatzas and Shreve, 1991, pp. 114-116)

$$\sup_{t_1,t_2:|t_2-t_1|\leq 1/n}\big|W(t_2)-W(t_1)\big|=O\Big(\frac{\sqrt{\log\ n}}{\sqrt{n}}\Big). \tag{43}$$

This implies that

$$-\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\log\ (t/n)u_t=\int_0^1 W(r)d\ \log\ (nr)+O\Big(\frac{\sqrt{\log\ n}}{\sqrt{n}}\Big)=$$
$$-\int_0^1\log\ sdW(s)+O_P\Big(\frac{(\log\ n)^{3/2}}{\sqrt{n}}\Big) \tag{44}$$

Since

$$\int_0^1 \log\ s dW(s) =_d W\Big(\int_0^1 \log^2\ s ds\Big) = W(2),$$

we get that (37) indeed holds. Relations (36), (37) and (37) imply (16).

We now turn to proving (17). By (20),

$$\frac{B_n}{n} = \frac{1}{n}\sum_{t=1}^{n-1} z_t^2 - \overline{z}_n^2 = \frac{2}{n}\sum_{j=2}^{n-1}\frac{\tau_j}{j}\sum_{i=1}^{j-1}\tau_i - \overline{z}_n^2. \tag{45}$$

By the fact that $Var(\overline{z}_n) = O\Big(\frac{1}{n}\Big)$ (or by the central limit theorem for $\overline{z}_n$) we have

$$\overline{z}_n = 1 + O_P\Big(\frac{1}{\sqrt{n}}\Big). \tag{46}$$

In addition,

$$\frac{2}{n}\sum_{j=2}^{n-1}\frac{\tau_j}{j}\sum_{i=1}^{j-1}\tau_i = \frac{2}{n}\sum_{t=2}^{n-1}\frac{\tau_t - 1}{t}\sum_{i=1}^{t-1}\tau_i + \frac{2}{n}\sum_{t=2}^{n-1}\frac{1}{t}\sum_{i=1}^{t-1}(\tau_i - 1) + \frac{2}{n}\sum_{t=2}^{n-1}\frac{1}{t}\sum_{i=1}^{t-1} 1 =$$
$$\frac{2}{n}\sum_{t=1}^{n-1}\frac{\tau_t - 1}{t}\sum_{i=1}^{t-1}\tau_i + \frac{2}{n}\sum_{t=1}^{n-1}\frac{1}{t}\sum_{i=1}^{t-1}(\tau_i - 1) + 2 + O_P\Big(\frac{\log\ n}{n}\Big) =$$
$$F_n^{(1)} + F_n^{(2)} + 2 + O_P\Big(\frac{\log\ n}{n}\Big). \tag{47}$$

It is easy to see that

$$Var(F_n^{(1)}) = O\Big(\frac{1}{n^2}\sum_{t=1}^{n}\frac{1}{t^2}E\Big(\sum_{i=1}^{t-1}\tau_i\Big)^2\Big) = O\Big(\frac{1}{n}\Big)$$

and, thus,

$$F_n^{(1)} = O_P\Big(\frac{1}{\sqrt{n}}\Big). \tag{48}$$

Besides, as it is not difficult to observe,

$$Var(F_n^{(2)}) = O\Big(\frac{1}{n^2}\sum_{t=1}^{n}\Big(\sum_{i=t}^{n}\frac{1}{i}\Big)^2\Big) = O\Big(\frac{\log^2\ n}{n}\Big)$$

and, consequently,

$$F_n^{(2)} = O_P\Big(\frac{\log\ n}{\sqrt{n}}\Big). \quad (49)$$

From (45)-(49) it clearly follows that (17) indeed holds.

Similar to the proof of relation (16), one can show that, in the case $\xi = 1$,

$$\frac{1}{\sqrt{n}}(A_n^\gamma + D_n) = \mathcal{N}(0,2) + \frac{(\log n)^2\,(1-2\gamma)}{4\sqrt{n}} + o_P\Big(\frac{\log^2 n}{\sqrt{n}}\Big). \quad (50)$$

Indeed, as in (36), we get

$$\frac{1}{\sqrt{n}}(A_n^\gamma + D_n) = \frac{1}{\sqrt{n}}\sum_{t=1}^{n-1}(\tau_t - 1)\log(t/n) + \frac{1}{\sqrt{n}}\sum_{t=1}^{n-1}(\tau_t - 1)\Big[\frac{1}{t}\Big(\sum_{i=1}^{t}\log\ (i-\gamma)\Big) - \Big(\frac{1}{n}\sum_{t=1}^{n}\log\ (t-\gamma)\Big) - \log(t/n)\Big] + \frac{1}{\sqrt{n}}\Big[\sum_{t=1}^{n}\log^2\ (t-\gamma) - \frac{1}{n}\Big(\sum_{t=1}^{n}\log\ (t-\gamma)\Big)^2 + \sum_{t=1}^{n}\frac{1}{t}\sum_{i=1}^{t}\log\ (i-\gamma) - \sum_{t=1}^{n}\log\ (t-\gamma)\Big]. \quad (51)$$

Similar to the arguments for (37), we have that variance of the statistic

$$V_n = \frac{1}{\sqrt{n}}\sum_{t=1}^{n-1}(\tau_t - 1)\Big[\frac{1}{t}\Big(\sum_{i=1}^{t}\log\ (i-\gamma)\Big) - \Big(\frac{1}{n}\sum_{t=1}^{n}\log\ (t-\gamma)\Big) - \log(t/n)\Big]$$

satisfies

$$Var(nV_n) = \sum_{t=1}^{n-1}\Big[\frac{1}{t}\Big(\sum_{i=1}^{t}\log\ (i-\gamma)\Big) - \Big(\frac{1}{n}\sum_{i=1}^{n}\log\ (i-\gamma)\Big) - \log\ (t/n)\Big]^2 = \sum_{t=1}^{n-1}\Big[\Big(\frac{1}{2}-\gamma\Big)\frac{\log\ (t-\gamma)}{t} + \Big(\frac{1}{2}-\gamma\Big)\frac{\log\ (n-\gamma)}{n} + O\Big(\frac{1}{t}\Big)\Big]^2 = O(1) \quad (52)$$

and, thus,

$$V_n \to_P 0. \quad (53)$$

In addition, Using Euler-Maclaurin summation formula similar to (34), it is not difficult to show that

$$\sum_{t=1}^{n}\log^2\ (t-\gamma) = (n-\gamma)\log^2\ (n-\gamma) - 2(n-\gamma)\log\ (n-\gamma) + 2n + \frac{\log^2\ (n-\gamma)}{2} + O(1). \quad (54)$$

It is not difficult to get, using (54), together with (34) that

$$\frac{1}{\sqrt{n}}\Big[\sum_{t=1}^{n}\log^2\,(t-\gamma)-\frac{1}{n}\Big(\sum_{t=1}^{n}\log\,(t-\gamma)\Big)^2+$$

$$\sum_{t=1}^{n}\frac{1}{t}\sum_{i=1}^{t}\log\,(i-\gamma)-\sum_{t=1}^{n}\log\,(t-\gamma)\Big]=\frac{1}{2}\big(\frac{1}{2}-\gamma\big)\frac{\log^2\,n}{\sqrt{n}}+o\big(\frac{\log^2\,n}{\sqrt{n}}\big). \tag{55}$$

Relations (41), (51), (53) and (55) imply (50).

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n-1}(\tau_t-1)\log\,(t/n)\to_d\sqrt{2}\mathcal{N}(0,1) \tag{56}$$

Using (34) and (55), it is not difficult to get that

$$\frac{D_n}{n}=1+O\Big(\frac{\log^2\,n}{n}\Big). \tag{57}$$

Relations (9), (50) and (57) imply asymptotic expansion (11). ■

**Remark 5** *As follows from the proof of Theorem 1, the order of the error terms in asymptotic expansions (6) and (8) is, in fact,* $O_P\Big(\frac{\log^{3/2}n}{n}\Big)$.

**Remark 6** *Denote* $H(t)=\sum_{i=1}^{t}\frac{1}{i}$, $t\geq 1$, $H(0)=0$. *Consider the analogues of regressions (6) and (8) that involve logarithms of ordered sizes* $y_t=\log(Z_{(t)})$ *and the functions* $\tilde{x}_t=H(t-1)$ *of ranks of observations:*

$$y_t=a'-b'\tilde{x}_t; \tag{58}$$

$$\tilde{x}_t=c'-d'y_t. \tag{59}$$

*Similar to the proof of Theorem 1, one can show that the following asymptotic expansions hold for the tail index estimates* $\hat{b}'_n$ *and* $\hat{d}'_n$ *using models (58) and (59):*

$$\hat{b}'_n/\xi=1+\sqrt{\frac{2}{n}}\mathcal{N}(0,1)+O_P\Big(\frac{\log n}{n}\Big); \tag{60}$$

$$\xi\hat{d}'_n=1+\sqrt{\frac{2}{n}}\mathcal{N}(0,1)+O_P\Big(\frac{\log n}{n}\Big). \tag{61}$$

*Comparison of expansions (60) and (61) with asymptotic expansions (10) and (11) shows that,* ceteris paribus, *tail index estimation using regressions involving partial sums of harmonic series is to be preferred, in terms of the small sample bias, to that based on regressions (6) and (8) involving logarithms of shifted ranks* $\log(t-\gamma)$ *for any* $\gamma$. *On the other hand, regressions (6) and (8) are simpler to implement, more visual and much more commonly used than estimating procedures based on (60) and (61). Comparison of the asymptotic expansions for the tail index estimates using models (58) and (59) with the OLS tail parameter estimates in log-log rank-size regressions (6) and (8) also sheds light on the main driving force behind the small bias improvements using logarithms of shifted ranks* $\log(Rank-1/2)$. *This driving force is, essentially, the fact that* $\log(n-1/2)$ *provides better approximation for the partial sums* $H(n-1)$ *than does* $\log(n)$ *and, more generally, than* $\log(n+\alpha)$, $\alpha \geq -n$ *(see Havil, 2003, p. 75-79).*

**Remark 7** *Consider a population whose distribution exhibits deviation from power law (5) in the form*

$$P(Z > x) = x^{-\xi}\left(1 + c\left(x^{-\alpha\xi} - 1\right)\right), \; x \geq 1, \;\; \alpha > 0. \tag{62}$$

*Heuristic arguments suggest that the estimate of the tail index* $\xi$ *in OLS log-log rank-size regression (3) satisfies the following asymptotic expansion in the case of model (62):*

$$\frac{\hat{b}_n^\gamma}{\xi} = 1 + \frac{\alpha}{(1+\alpha)^2}c + \sqrt{\frac{2}{n}}\mathcal{N}(0,1) + \frac{(\log n)^2(2\gamma-1)}{4n} + o_P\Big(\frac{\log^2 n}{n}\Big) + o(c). \tag{63}$$

*That is, the term* $c\left(x^{-\alpha\xi}-1\right)$ *modeling the deviations from the exact power law in (62) creates a bias in the estimate. Analogously, according to the heuristic arguments, the following expansion holds for the Hill estimator*

$$\xi_{Hill} = \frac{n}{\sum_{t=1}^{n}\Big(\log Z_{(t)} - \log Z_{(n+1)}\Big)} = \frac{n}{\sum_{t=1}^{n}\tau_t/\xi}$$

*in the case of a population with the distribution satisfying (62):*

$$\frac{\xi_{Hill}}{\xi} = 1 + \frac{\alpha}{1+\alpha}c + \sqrt{\frac{1}{n}}\mathcal{N}(0,1) + o_P\Big(\frac{\log^2 n}{n}\Big) + o(c). \tag{64}$$

*According to asymptotic expansions (63) and (64), although the standard error of the Hill estimator is less than that of the OLS log-log rank-size tail index estimate* $\hat{b}_n^\gamma$, *its (small sample) bias in the case of deviations from the power law is (typically) greater than that of* $\hat{b}_n^\gamma$. *That is, the OLS estimate* $\hat{b}_n^\gamma$ *of the tail index in regression (3) is more robust to deviations from the exact power law than Hill's estimator* $\xi_{Hill}$.

# References


Aban, I. B. and Meerschaert, M. M. (2004), 'Generalized least-squares estimators for the thickness of heavy tails', *Journal of Statistical Planning and Inference* **119**, 341–352.

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

Beirlant, J., Goegebeur, Y., Teugels, J. and Segers, J. (2004), *Statistics of extremes*, Wiley Series in Probability and Statistics, John Wiley & Sons Ltd., Chichester. Theory and applications, With contributions from Daniel De Waal and Chris Ferro.

Beirlant, J. and Teugels, J. L. (1989), Asymptotic normality of Hill's estimator, *in* 'Extreme value theory (Oberwolfach, 1987)', Vol. 51 of *Lecture Notes in Statist.*, Springer, New York, pp. 148–155.

Beirlant, P., Dierckx, G., Goegebeur, Y. and Matthys, G. (1999*b*), 'Tail index estimation and an exponential regression model', *Extremes* **2177**, 177–200.

Brakman, S., Garretsen, H., van den Berg, M. and van Marrewijk, C. (1999), 'The return of zipf: towards a further understanding of the rank-size distribution', *Journal of Regional Science* **39**, 182–213.

Chen, Y., Härdle, W. and Spokoiny, V. (2005), 'Ghica - risk analysis with gh distributions and independent components', *Working paper, Humboldt Universität Berlin* .

Chernozhukov, V. (2005*a*), 'Extremal quantile regression', *Annals of Statistics* **33**, 806–839.

Chernozhukov, V. (2005*b*), 'Inference for extremal conditional quantile models', *Working paper, Massachusetts Institute of Technology* .

Csörgö, M. and Révész, P. (1981), *Strong approximations in probability and statistics*, Academic Press, New York.

Csörgő, S., Deheuvels, P. and Mason, D. (1985), 'Kernel estimates of the tail index of a distribution', *Annals of Statistics* **13**, 1050–1077.

Csörgő, S. and Mason, D. M. (1985), 'Central limit theorems for sums of extreme values', *Mathematical Proceedings of the Cambridge Philosophical Society* **98**, 547–558.

Csörgő, S. and Viharos, L. (1997), 'Asymptotic normality of least-squares estimators of tail indices', *Bernoulli* **3**, 351–370.

Csörgő, S. and Viharos, L. (2006), 'Testing for small bias of tail index estimators', *Journal of Computational and Applied Mathematics* **186**, 232–252.

Danielsson, J., de Haan, L., Peng, L. and de Vries, C. G. (2001), 'Using a bootstrap method to choose the sample fraction in tail index estimation', *Journal of Multivariate Analysis* **76**, 226–248.

De Vries, C. G. (2005), 'The simple economics of bank fragility', *Working paper, Erasmus Universiteit Rotterdam* .

Duffie, D. and Pan, J. (1997), 'An overview of value at risk', *Journal of Derivatives* **4**, 7–49.

Dufour, J.-M., Khalaf, L., Kurz-Kim, J.-R. and Beaulieu, M.-C. (2005), 'Exact confidence set estimation and goodness-of-fit test methods for asymmetric heavy tailed stable distributions', *Working paper, Université de Montréal* .

Dufour, J.-M. and Kurz-Kim, J.-R. (2005), 'Exact inference and optimal invatriant estimation for the tail coefficient of symmetric $\alpha-$stable distributions', *Working paper, Université de Montréal* .

Embrechts, P., Klupperberg, C. and Mikosch, T. (1997), *Modelling extremal events for insurance and finance*, Springer, New York.

Fama, E. (1965), 'The behavior of stock market prices', *Journal of Business* **38**, 34–105.

Feurverger, A. and Hall, P. (1999), 'Estimating a tail exponent by modelling departure from a pareto distribution', *Annals of Statistics* **27**, 760–781.

Gabaix, X., Gopikrishnan, P., Plerou, V. and Stanley, H. E. (2003), 'A theory of power-law distributions in financial market fluctuations', *Nature* **423**, 267–270.

Gabaix, X. and Ioannides, Y. (2004), The evolution of city size distributions, *in* V. Henderson and J.-F. Thisse, eds, 'Handbook of Regional and Urban Economics', Vol. 4, North-Holland, pp. 2341–2378.

Garcia, R., Renault, E. and Veredas, D. (2005), 'Estimation of stable distributions by indirect inference', *Working paper, Université de Montréal* .

Glasserman, P., Heidelberger, P. and Shahabuddin, P. (2002), 'Portfolio value-at-risk with heavy-tailed risk factors', *Mathematical Finance* **12**, 239–269.

Groeneboom, P., Lopuhaä, H. P. and de Wolf, P. P. (2003), 'Kernel-type estimators for the extreme value index', *Annals of Statistics* **31**, 1956–1995.

Haas, M., Mittnik, S., Paolella, M. S. and Steude, S. C. (2005), 'Stable mixture garch models', *Working paper, University of Munich* .

Haeusler, E. and Teugels, J. L. (1985), 'On asymptotic normality of Hill's estimator for the exponent of regular variation', *Annals of Statistics* **13**, 743–756.

Hall, P. (1982), 'On some simple estimates of an exponent of regular variation', *Journal of the Royal Statistical Society. Series B.* **44**, 37–42.

Hall, P. and Welsh, A. H. (1984), 'Best attainable rates of convergence for estimates of parameters of regular variation', *Annals of Statistics* **12**, 1079–1084.

Hartmann, P., Straetmans, S. and De Vries, C. G. (2005), 'A typology of currency crises', *Working paper, Erasmus Universiteit Rotterdam* .

Haug, S., Klüppelberg, C., Lindner, A. and Zapp, M. (2005), 'Estimating the cogarch(1, 1) model - a first go', *Working paper, Munich University of Technology* .

Havil, J. (2003), *Gamma: Exploring Euler's constant*, Princeton University Press, Princeton, NJ.

Hill, B. M. (1975), 'A simple general approach to inference about the tail of a distribution', *Ann. Statist.* **3**, 1163–1174.

Huisman, R., Koedijk, K. G., Kool, C. J. M. and Palm, F. (2001), 'Tail-index estimates in small samples', *Journal of Business and Economic Statistics* **19**, 208–216.

Ibragimov, R. (2004), 'On the robustness of economic models to heavy-tailedness assumptions', *Mimeo, Yale University* . Available at http://pantheon.yale.edu/~ri23/Papers/HeavyTails.pdf.

Ibragimov, R. (2005*a*), *New majorization theory in economics and martingale convergence results in econometrics*. Ph.D. dissertation, Yale University.

Ibragimov, R. (2005*b*), 'On efficiency of linear estimators under heavy-tailedness', *Harvard Institute of Economic Research Discussion Paper No. 2085* . Available at http://post.economics.harvard.edu/hier/2005papers/HIER2085.pdf.

Ibragimov, R. and Phillips, P. C. B. (2004), 'Regression asymptotics using martingale convergence methods', *Cowles Foundation Discussion Paper 1473* . Available at http://cowles.econ.yale.edu/P/cd/d14b/d1473.pdf.

Ibragimov, R. and Walden, J. (2006), 'The limits of diversification when losses may be large', *Harvard Institute of Economic Research Discussion Paper No. 2104* . Available at http://post.economics.harvard.edu/hier/2006papers/HIER2104.pdf.

Jansen, D. W. and de Vries, C. G. (1991), 'On the frequency of large stock returns: Putting booms and busts into perspective', *Review of Economics and Statistics* **73**, 18–32.

Karatzas, I. and Shreve, S. E. (1991), *Brownian motion and stochastic calculus*, Springer-Verlag, New York.

Komlós, J., Major, P. and Tusnády, G. (1975), 'An approximation of partial sums of independent rv's and the sample df. i.', *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **32**, 111–131.

Komlós, J., Major, P. and Tusnády, G. (1976), 'An approximation of partial sums of independent rv's and the sample df. i.', *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **34**, 33–58.

Kratz, M. and Resnick, S. I. (1996), 'The QQ-estimator and heavy tails', *Communications in Statistics. Stochastic Models* **12**, 699–724.

Kurz-Kim, J.-R., Rachev, S. T., Samorodnitsky, G. and Stoyanov, S. (2005), 'Asymptotic distribution of unbiased linear estimators in the presence of heavy-tailed stochastic regtressors and residuals', *Working paper, Deutsche Bundesbank and University of California at Santa Barbara* .

Levy, J. B. and Taqqu, M. S. (2005), 'The asymptotic codifference and covariation of log-fractional stable noise', *Working paper, Boston University* .

Lombardini, M. J. and Calzolari, G. (2005), 'Indirect estimation of $\alpha-$stable stochastic volatility models', *Working paper, European Central Bank* .

Loretan, M. and Phillips, P. C. B. (1994), 'Testing the covariance stationarity of heavy-tailed time series', *Journal of Empirical Finance* **1**, 211–248.

Mandelbrot, B. (1963), 'The variation of certain speculative prices', *Journal of Business* **36**, 394–419.

Mandelbrot, B. (1997), *Fractals and scaling in finance. Discontinuity, concentration, risk*, Springer-Verlag, New York.

McCulloch, J. H. (1997), 'Measuring tail thickness to estimate the stable index alpha: A critique', *Journal of Business and Economic Statistics* **15**, 74–81.

McCulloch, J. H. (2005), 'The risk-neutral measure and option pricing under log-stable uncertainty', *Working paper, Ohio State University* .

Meerschaert, M. M. and Scheffler, H.-P. (2000), Portfolio modelling with heavy tailed random vectors, *in* S. T. Rachev, ed., 'Handbook of Heavy-Tailed Distributions in Finance', Elsevier, North-Holland.

Nishiyama, Y. and Osada, S. (2005), 'Statistical theory of rank size rule regression under pareto distribution', *Working paper, Kyoto Institute of Economic Research* .

Nishiyama, Y. and Robinson, P. M. (2000), 'Edgeworth expansions for semiparametric averaged derivatives', *Econometrica* **68**, 931–979.

Nishiyama, Y. and Robinson, P. M. (2005), 'The bootstrap and the edgeworth correction for semiparametric averaged derivatives', *Econometrica* **73**, 903–948.

Nolan, J. P. (2005), 'Multivariate stable densities and distribution functions: general and elliptical case', *Working paper, American University* .

Ortobelli, S. and Rachev, S. (2005), 'Risk management and dynamic portfolio selection with stable paretian distributions', *Working paper, University of Berdamo and University of California at Santa Barbara* .

Park, J. Y. and Phillips, P. C. B. (1999), 'Asymptotics for nonlinear transformations of integrated time series', *Econometric Theory* **15**, 269–298.

Park, J. Y. and Phillips, P. C. B. (2001), 'Nonlinear regressions with integrated time series', *Econometrica* **69**, 117–161.

Phillips, P. C. B. (1999), 'Unit root log periodogram regression', *Cowles Foundation Discussion Paper 1244* . Available at http://cowles.econ.yale.edu/P/cd/d12a/d1244.pdf.

Phillips, P. C. B. (2001), 'Regression with slowly varying regressors', *Cowles Foundation Discussion Paper 1310* . Available at http://cowles.econ.yale.edu/P/cd/d13a/d1310.pdf.

Quintos, C. E., Fan, Z. and Phillips, P. (2001), 'Structural change in tail behavior and the asian financial crisis', *Review of Economic Studies* .

Rachev, S., Jašic, T., Stoyanov, S. and Fabozzi, F. J. (2005), 'Devising zero-investment strategies using reward-risk stock selection criteria', *Working paper, University of Karlsruhe* .

Samarakoon, D. M. M. and Knight, K. (2005), 'A note on unit root tests with infinite variance noise', *Working paper, University of Toronto* .

Schultze, J. and Steinebach, J. (1999), 'On least squares estimates of an exponential tail coefficient', *Statistcs & Extremes* **2**, 177–200.

Uchaikin, V. V. and Zolotarev, V. M. (1999), *Chance and stability. Stable distributions and their applications*, VSP, Utrecht.

Viharos, L. (1999), 'Weighted least-squares estimators of tail indices', *Probability and Mathematical Statistics* **19**, 249–265.

Weron, R. (2001), 'Levy-stable distributions revisited: Tail index ¿ 2 does not exclude the levy-stable regime', *International Journal of Modern Physics, C* **12**, 209–223.