PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 27, Fasc. 2 (2007), pp. 275–302

ASYMPTOTIC DISTRIBUTION OF UNBIASED LINEAR ESTIMATORS IN THE PRESENCE OF HEAVY-TAILED STOCHASTIC REGRESSORS AND RESIDUALS*

BY

GENNADY SAMORODNITSKY (ITHACA, NY), SVETLOZAR T. RACHEV (KARLSRUHE), JEONG-RYEOL KURZ-KIM (FRANKFURT AM MAIN) AND STOYAN V. STOYANOV (SANTA BARBARA, CA)

Abstract. Under the symmetric α-stable distributional assumption for the disturbances, Blattberg and Sargent [3] consider unbiased linear estimators for a regression model with non-stochastic regressors. We study both the rate of convergence to the true value and the asymptotic distribution of the normalized error of the linear unbiased estimators. By doing this, we allow the regressors to be stochastic and disturbances to be heavy-tailed with either finite or infinite variances, where the tail-thickness parameters of the regressors and disturbances may be different.

2000 AMS Mathematics Subject Classification: Primary: 62J05; Secondary: 60E07.

Key words and phrases: Asymptotic distribution, rate of convergence, stochastic regressor, stable non-Gaussian, finite or infinite variance, heavy tails.

1. INTRODUCTION

For the estimation of the coefficients of a regression model one typically applies ordinary least squares (OLS), which is equivalent to the maximum likelihood estimation if the disturbances are normally distributed. Further-

* The views expressed in this paper are those of the authors and not necessarily those of the Deutsche Bundesbank or other institutions involved.

Samorodnitsky's research was partially supported by the Deutschen Forschungsgemeinschaft and by NSA grant MSPF-02G-183 and NSF grant DMS-0071073 at Cornell University.

Rachev's research was supported by grants from Division of Mathematical, Life and Physical Sciences, College of Letters and Science, University of California, Santa Barbara, and the Deutschen Forschungsgemeinschaft.

Kurz-Kim gratefully acknowledges research support from the Alexander von Humboldt Foundation.

more, according to the Gauss–Markov theorem, the OLS estimator has the minimum variance of all linear unbiased estimators if the disturbances follow a distribution with finite variance. However, if the disturbances follow a distribution with infinite variance, but with finite mean, the OLS estimator is still unbiased but no longer a minimum variance estimator.

Relaxing the normality assumption by allowing disturbances to have a symmetric α-stable distribution with infinite variance ($1 < \alpha < 2$), Blattberg and Sargent [3] generalize the OLS estimator to a different linear unbiased estimator that minimizes the α-stable scale of the estimator. The generalization is performed in the framework of a regression model in which the independent variable is assumed to be non-stochastic.

We consider both the rate of convergence to the true value and the asymptotic distribution of the normalized error of the linear unbiased estimators of coefficients in the regression model with both stochastic regressors and disturbances being heavy-tailed with either finite or infinite variances,[1] and the tail-thickness parameters of the regressors and disturbances may be different. Even though our distributional assumptions are more general than the assumptions of α-stability, the *limiting distributions* of the estimators will often be expressed through stable random variables (r.v.'s).

A random variable (r.v.) X is *stable* if for all A, $B > 0$ there are $C > 0$ and D real such that $AX_1 + BX_2 \stackrel{d}{=} CX + D$, where X_1 and X_2 are independent copies of X. For a stable random variable X there is a number $\alpha \in (0, 2]$ such that $C^\alpha = A^\alpha + B^\alpha$. Exponent α is called the *stability parameter*. A r.v. with exponent α is said to be *α-stable* distributed. Closed-form expressions of α-stable distributions exist only for a few special cases. However, the logarithm of the characteristic function of the α-stable distribution can be written as (see [15] and [14] for more details on α-stable distributions)

$$\ln \varphi(t) = \begin{cases} -\sigma^\alpha |t|^\alpha \left[1 - i\beta \operatorname{sign}(t) \tan \frac{\pi\alpha}{2}\right] + i\mu t & \text{for } \alpha \neq 1, \\ -\sigma |t| \left[1 + i\beta \frac{\pi}{2} \operatorname{sign}(t) \ln |t|\right] + i\mu t & \text{for } \alpha = 1, \end{cases}$$

where α is the *stability parameter* (or tail-thickness parameter), σ is the *scale parameter*, $\beta \in [-1, 1]$ is the *skewness parameter*, and μ is the *location parameter*. If $\beta = 0$, the distribution is symmetric. The shape of the symmetric α-stable distribution ($S\alpha S$) is determined by the tail-thickness parameter $\alpha \in (0, 2]$. For $0 < \alpha < 2$ the tails of the distribution are thicker than those of the normal distribution; and the tail-thickness increases as α decreases. When $\alpha = 2$, the $S\alpha S$

[1] There is some controversy on whether the variance of financial returns is always infinite. We avoid this controversy by using a heavy-tailed model that allows for both finite or infinite variance.

distribution coincides with the normal distribution with variance $2\sigma^2$, the only member of the family with finite variance. When $\alpha = 1$, the $S\alpha S$ distribution reduces to the Cauchy distribution. If $\alpha < 2$, moments of order α or higher do not exist, which means the variance is infinite. If X is an α-stable random variable, $0 < \alpha < 2$, with scale σ, skewness β, and location μ, then a common notation is $X \sim S_\alpha(\sigma, \beta, \mu)$. In that case the tails of X are given by

$$P(\pm X > \lambda) \sim C_\alpha \frac{1 \pm \beta}{2} \sigma^\alpha \lambda^{-\alpha} \tag{1.1}$$

as $\lambda \to \infty$, where

$$C_\alpha = \Big(\int_0^\infty x^{-\alpha} \sin x \, dx\Big)^{-1}. \tag{1.2}$$

Some more basic information and notation on stable random variables we use, unless otherwise specified, can be found in [14].

One distinct example for a possible application of our results in the paper can be found in financial market analysis. For an econometric analysis based on a dynamic capital asset pricing model, distributional assumptions on the disturbance must be made. Since publishing the paper by Bachelier [1], the traditional and most widely adopted distributional assumption on financial return process has been the Gaussian assumption. Due to the influential works of Mandelbrot [10] and Fama [6], however, the α-stability with $0 < \alpha < 2$ has often been considered to be a more realistic distribution assumption for asset returns than that of a normal distribution, because asset returns are typically fat-tailed and excessively peaked around zero – phenomena that can be captured by α-stable distributions with $\alpha < 2$. This is the so-called *stable Paretian assumption.* In a certain sense the stable Paretian assumption is a generalization rather than an alternative to the Gaussian assumption. Indeed, according to the generalized central limit theorem, the limiting distribution of the sum of a large number of iid r.v.'s is α-stable with $0 < \alpha \leqslant 2$; see [15]. For more applications of the α-stable distributions in economics and finance, see [11] and [12].

In this paper we introduce a new family of unbiased estimators of the coefficient in a linear regression model in the presence of heavy tails, generalizing the estimator introduced by Blattberg and Sargent [3]. In Section 2 we present our new estimator and analyze the asymptotic distribution of the estimator. In Section 3 we summarize various scenarios for different tail indices for regressors and disturbances, and discuss the choice of the optimal estimator in the proposed family of estimators, both analytically and numerically. Section 4 contains some concluding remarks.

2. RATE OF CONVERGENCE AND THE LIMITING DISTRIBUTION FOR THE REGRESSION COEFFICIENT ESTIMATOR

Consider a simple regression model like the following:

$$Y_j = \beta X_j + U_j, \quad j = 1, 2, \ldots \tag{2.1}$$

We assume that the regressors $\{X_j\}$ are iid random variables with polynomially decaying tails. Specifically,

$$P(|X_1| > \lambda) \sim D_1 \lambda^{-\alpha_x}, \ \lambda \to \infty, \quad \text{for some } \alpha_x > 0 \text{ and } D_1 > 0. \tag{2.2}$$

Furthermore, we assume that the noise (disturbances) $\{U_j\}$ are also iid random variables, which we assume to be symmetric, with

$$P(|U_1| > \lambda) \sim D_2 \lambda^{-\alpha_u}, \ \lambda \to \infty, \quad \text{for some } \alpha_u > 0 \text{ and } D_2 > 0. \tag{2.3}$$

We assume, further, that the sequences $\{X_j\}$ and $\{U_j\}$ are independent.

Note that no assumptions on the symmetry of either dependent observations or regressors are made. We remark, further, that it is relatively straightforward (at least, away from the boundary cases) to extend the results below to the case where the tails of the regressors and noise variables are regularly varying (i.e. adding slowly varying factors in (2.2) and (2.3)). Since such slowly varying functions are not practically observable, we decided against including extra technical arguments in an already highly technical paper. Finally, we allow values of α_x and α_u in the interval $(0, 1]$ as well, since our methods cover those cases equally well.

The goal is to estimate the regression coefficient β in (2.1), and our estimator is

$$\hat{\beta}_{\theta,n} = \frac{\sum_{j=1}^{n} X_j^{\langle 1/(\theta-1)\rangle} Y_j}{\sum_{j=1}^{n} |X_j|^{\theta/(\theta-1)}} \tag{2.4}$$

for some $\theta > 1$ with $\langle\cdot\rangle$ defined as a signed power.[2] Note that the OLS estimator corresponds to $\theta = 2$ in (2.4).

Our immediate task is to understand the behavior of the difference

$$\Delta_n := \hat{\beta}_{\theta,n} - \beta = \frac{\sum_{j=1}^{n} X_j^{\langle 1/(\theta-1)\rangle} U_j}{\sum_{j=1}^{n} |X_j|^{\theta/(\theta-1)}} \stackrel{d}{=} \frac{\sum_{j=1}^{n} |X_j|^{1/(\theta-1)} U_j}{\sum_{j=1}^{n} |X_j|^{\theta/(\theta-1)}}, \tag{2.5}$$

where the last distributional equality follows from the symmetry of the noise. That is, we are interested in the rate of convergence of the estimator $\hat{\beta}_{\theta,n}$ to the true value depending on the choice of θ. When such convergence actually takes place, this will also establish consistency (in probability) of our estimator.

---

[2] We mean that $a^{\langle p\rangle} = |a|^{p-1} a$.

It is clear that the rate of convergence to zero of the difference Δ_n depends significantly on the tail exponents α_x and α_u, and on the choice of θ. What is interesting is that we will see below that there are basically 7 different cases of possible values of α_x and α_u, in each of which the rate of convergence is a different function of θ.

A common feature of our results will be the existence of an exponent d such that

$$n^d \Delta_n \Rightarrow W \quad \text{as } n \to \infty \tag{2.6}$$

for some non-degenerate weak limit W. Occasionally, on certain boundaries we will have to modify (2.6) to allow for a slowly varying factor on the left-hand side. That is, we will have

$$n^d L(n) \Delta_n \Rightarrow W \quad \text{as } n \to \infty, \tag{2.7}$$

where L is a slowly varying function.[3] In any case we will view the exponent d in either (2.6) or (2.7) as measuring the rate of convergence. In particular, the exponent d turns out to be a different function of θ in the 7 different cases of possible values of α_x and α_u we mentioned above.

The reader will find it easier to follow the different technical detail below after noticing the existence of several critical boundaries. The first boundary is that

$$\frac{\theta - 1}{\theta} \alpha_x > 1. \tag{2.8}$$

Note that on one side of that boundary $|X_j|^{\theta/(\theta-1)}$ has a finite mean, and hence the denominator in (2.5) is governed by the law of large numbers (LLN). On the other side of that boundary, $|X_j|^{\theta/(\theta-1)}$ is in the domain of attraction of a positive stable law and the corresponding heavy-tailed central limit theorem governs the behavior of the denominator in (2.5). On the boundary itself, the mean is infinite, but the (weak) LLN is still in force.

The second critical boundary is that of

$$\min((\theta - 1)\alpha_x, \alpha_u) > 2. \tag{2.9}$$

Here on one side of the boundary the random variables $|X_j|^{1/(\theta-1)} U_j$ have a finite variance, and hence the Gaussian central limit theorem (CLT) governs the behavior of the numerator in the second expression in (2.5). On the other side of that boundary these random variables are in the domain of attraction of a symmetric stable non-Gaussian law, and hence the corresponding CLT will be responsible for the behavior of the numerator. On the boundary itself the variance is infinite, but the CLT will still be in force.

[3] $L(x)$ is a slowly varying function as $x \to \infty$, if for every constant $c > 0$ and $\lim_{x \to \infty} L(cx)/L(x)$ exists and is equal to 1.

We now proceed to consider the different ranges of α_x and α_u mentioned above.

Scenario 1. Suppose that

$$0 < \alpha_x \leqslant 1 \quad \text{and} \quad \alpha_u \geqslant 2. \tag{2.10}$$

Note that under this scenario (2.8) fails independently of θ. On the other hand, one can be on either side of the other the critical boundary (2.9) under this scenario, and the various possible cases are described in the following theorem.

THEOREM 2.1. *Assume that* (2.10) *holds.*

(a) *Suppose that*

$$\theta \geqslant \frac{2+\alpha_x}{\alpha_x}.$$

Then the exponent d in (2.7) *is given by*

$$d = \frac{2\theta-(\theta-1)\alpha_x}{2(\theta-1)\alpha_x}. \tag{2.11}$$

Specifically, in the non-boundary case

$$\alpha_u > 2 \quad \textit{and} \quad \theta > \frac{2+\alpha_x}{\alpha_x} \tag{2.12}$$

we have

$$n^d \Delta_n \Rightarrow \frac{(\mathrm{E}\,U_1^2)^{1/2}(\mathrm{E}\,|X_1|^{2/(\theta-1)})^{1/2}}{C_{\alpha_x(\theta-1)/\theta}^{-\theta/((\theta-1)\alpha_x)}\,D_1^{\theta/\alpha_x(\theta-1)}}\,\frac{N(0,1)}{S_{(\theta-1)\alpha_x/\theta}(1,1,0)} \tag{2.13}$$

weakly, where $N(0, 1)$ and $S_{(\theta-1)\alpha_x/\theta}(1, 1, 0)$ on the right-hand side above are independent. Furthermore, D_1 is the constant in the tail in (2.2), *and the constant C_α for $0 < \alpha < 2$ is given in* (1.2).

In the boundary case

$$\alpha_u = 2 \quad \textit{and} \quad \theta > \frac{2+\alpha_x}{\alpha_x} \tag{2.14}$$

we have

$$\frac{n^d}{(\log n)^{1/2}}\Delta_n \Rightarrow \frac{(D_2\,\mathrm{E}\,[|X_1|^{2/(\theta-1)}])^{1/2}}{C_{\alpha_x(\theta-1)/\theta}^{-\theta/\alpha_x(\theta-1)}\,D_1^{\theta/\alpha_x(\theta-1)}}\,\frac{N(0,1)}{S_{(\theta-1)\alpha_x/\theta}(1,1,0)}, \tag{2.15}$$

the random variables on the right-hand side of (2.15) *being, once again, independent. Here D_2 is the tail constant in* (2.3).

In the second boundary case

$$\alpha_u > 2 \quad \textit{and} \quad \theta > \frac{2+\alpha_x}{\alpha_x} \tag{2.16}$$

we have

$$\frac{n^d}{(\log n)^{1/2}} \Delta_n \Rightarrow \frac{(D_1 \mathrm{E}[|U_1|^2])^{1/2}}{C_{\alpha_x(\theta-1)/\theta}^{-\theta/\alpha_x(\theta-1)} D_1^{\theta/\alpha_x(\theta-1)}} \frac{N(0, 1)}{S_{(\theta-1)\alpha_x/\theta}(1, 1, 0)}, \tag{2.17}$$

and the random variables on the right-hand side of (2.17) *are independent.*

In the third boundary case

$$\alpha_u = 2 \quad \text{and} \quad \theta > \frac{2+\alpha_x}{\alpha_x} \tag{2.18}$$

we obtain

$$\frac{n^d}{\log n} \Delta_n \Rightarrow \frac{(D_1 D_2)^{1/2}}{C_{\alpha_x(\theta-1)/\theta}^{-\theta/\alpha_x(\theta-1)} D_1^{\theta/\alpha_x(\theta-1)}} \frac{N(0, 1)}{S_{(\theta-1)\alpha_x/\theta}(1, 1, 0)}, \tag{2.19}$$

and the random variables on the right-hand side of (2.19) *are independent.*

(b) *Suppose that*

$$(\theta-1)\alpha_x < 2 \quad \text{or, equivalently,} \quad \theta < \frac{2+\alpha_x}{\alpha_x}. \tag{2.20}$$

Then the exponent d in (2.7) *is given by*

$$d = \frac{1}{\alpha_x}; \tag{2.21}$$

specifically, we have

$$n^d \Delta_n \Rightarrow D_1^{-1/\alpha_x} \frac{\sum_{j=1}^{\infty} \Gamma_j^{-1/((\theta-1)\alpha_x)} U_j}{\sum_{j=1}^{\infty} \Gamma_j^{-\theta/((\theta-1)\alpha_x)}}. \tag{2.22}$$

Here, as usual, (Γ_j) represents the arrival times of a unit rate homogeneous Poisson process on $(0, \infty)$, *independent of the sequence* (U_j) (*here the numerator and the denominator on the right-hand side of* (2.22) *are not independent*).

Proof. (a) We start with the non-boundary case. Here

$$n^d \Delta_n = \frac{n^{-1/2} \sum_{j=1}^{n} |X_j|^{1/(\theta-1)} U_j}{n^{-\theta/(\theta-1)\alpha_x} \sum_{j=1}^{n} |X_j|^{\theta/(\theta-1)}}. \tag{2.23}$$

Let $\varepsilon > 0$, and

$$K_n(\varepsilon) = \{j = 1, 2, \ldots, n : |X_j| > \varepsilon n^{1/\alpha_x}\}. \tag{2.24}$$

Note that

$$k_n(\varepsilon) := \mathrm{Card}(K_n(\varepsilon)) \Rightarrow \mathrm{Poiss}(D_1 \varepsilon^{-\alpha_x}) \quad \text{as } n \to \infty \tag{2.25}$$

weakly, where $\mathrm{Poiss}(\mu)$ stands for a Poisson random variable with mean μ.

Write (by giving names to the numerator and denominator on the right-hand side of (2.23))

$$n^d \Delta_n = \frac{NU_n}{DE_n}, \tag{2.26}$$

and let

$$\hat{N}_n = n^{-1/2} \sum_{j=1}^{n} |\hat{X}_j|^{1/(\theta-1)} U_j, \tag{2.27}$$

$$\hat{D}_n = n^{-\theta/(\theta-1)\alpha_x} \sum_{j \in K_n(\varepsilon)} |X_j|^{\theta/(\theta-1)}, \tag{2.28}$$

where

$$\hat{X}_j = \begin{cases} X_j & \text{if } j = \{1, 2, \ldots, n\} \backslash K_n(\varepsilon), \\ \tilde{X}_j & \text{if } j \in K_n(\varepsilon). \end{cases} \tag{2.29}$$

In (2.29), $\{\tilde{X}_j\}$ is an iid sequence with a common law $\mathbf{P}(X_j \in \cdot \mid |X_j| \leqslant \varepsilon n^{1/\alpha_x})$, and independent of the sequences $\{X_j\}$ and $\{U_j\}$.

Note that $\hat{N}_n$ and $\hat{D}_n$ are independent. By the CLT for triangular arrays we see that

$$\hat{N}_n \Rightarrow N(0, \mathrm{E}[|X_1|^{2/(\theta-1)}]\,\mathrm{E}[U_1^2]) \tag{2.30}$$

(e.g. Theorem 5.1.2 in [9]).

Furthermore, we claim that

$$\hat{D}_n \Rightarrow \sum_{j=1}^{N_\varepsilon} Z_j(\varepsilon), \tag{2.31}$$

where N_ε is Poiss$(D_1 \varepsilon^{-\alpha_x})$, independent of an iid sequence $\{Z_j(\varepsilon)\}$ with a common law

$$\mathbf{P}(Z_j(\varepsilon) > \lambda) = \left(\frac{\lambda^{(\theta-1)/\theta}}{\varepsilon}\right)^{-\alpha_x}, \qquad \lambda \geqslant \varepsilon^{\theta/(\theta-1)}.$$

This is, however, clear because of (2.25) and the fact that, for all λ,

$$\mathbf{P}(n^{-\theta/(\theta-1)\alpha_x} |X_1|^{\theta/(\theta-1)} > \lambda \mid |X_1| > \varepsilon n^{1/\alpha_x}) \Rightarrow \mathbf{P}(Z_1(\varepsilon) > \lambda) \quad \text{as } n \to \infty,$$

and the independence between N_ε and $\{Z_j(\varepsilon)\}$ follows from the fact that the terms in the sum in (2.28) are independent of the random number of terms $K_n(\varepsilon)$. We conclude that

$$\frac{\hat{N}_n}{\hat{D}_n} \Rightarrow (\mathrm{E}[|X_1|^{2/(\theta-1)}]\,\mathrm{E}[U_1^2])^{1/2} \cdot \frac{N(0, 1)}{\sum_{j=1}^{N_\varepsilon} Z_j(\varepsilon)}, \tag{2.32}$$

with the numerator and the denominator on the right-hand side of (2.32) being independent.

Note that

$$\text{(2.33)}\quad \mathrm{E}\left[\exp\{-\theta \sum_{j=1}^{N_\varepsilon} Z_j(\varepsilon)\}\right]$$

$$\to \exp\left\{-D_1 \int_0^\infty (1-e^{-\theta x}) \frac{(\theta-1)\alpha_x}{\theta} x^{-(\theta-1)\alpha_x/\theta-1}\,dx\right\}$$

$$= \mathrm{E}\left[\exp\{-\theta C_{(\theta-1)\alpha_x/\theta}^{-\theta/((\theta-1)\alpha_x)} D_1^{\theta/((\theta-1)\alpha_x)} S_{(\theta-1)\alpha_x/\theta}(1,\,1,\,0)\}\right]$$

for $\theta > 0$ as $\varepsilon \to 0$. Therefore, (2.13) will follow once we show that for all $\delta > 0$

$$\lim_{\varepsilon\to 0}\overline{\lim_{n\to\infty}}\, \mathbf{P}\left(\left|\frac{NU_n}{DE_n} - \frac{\hat{N}_n}{\hat{D}_n}\right| > \delta\right) = 0;$$

see Theorem 3.2 in [2]. To this end, it is enough to prove that

$$\text{(2.34)}\qquad \lim_{\varepsilon\to 0}\overline{\lim_{n\to\infty}}\, \mathbf{P}\left(\left|\frac{NU_n}{DE_n} - \frac{\hat{N}_n}{DE_n}\right| > \delta\right) = 0,$$

and

$$\text{(2.35)}\qquad \lim_{\varepsilon\to 0}\overline{\lim_{n\to\infty}}\, \mathbf{P}\left(\left|\frac{\hat{N}_n}{\hat{D}_n} - \frac{\hat{N}_n}{DE_n}\right| > \delta\right) = 0.$$

We will start with (2.34). Since $(1/DE_n)$ is tight, it is enough to prove that for every $\delta > 0$

$$\lim_{\varepsilon\to 0}\overline{\lim_{n\to\infty}}\, \mathbf{P}(|NU_n - \hat{N}_n| > \delta) = 0.$$

We have

$$\text{(2.36)}\qquad |NU_n - \hat{N}_n| = n^{-1/2}\left|\sum_{j\in K_n(\varepsilon)} (|X_j|^{1/(\theta-1)} - |\tilde{X}_j|^{1/(\theta-1)})\, U_j\right|$$

$$\leqslant n^{-1/2} \sum_{j\in K_n(\varepsilon)} |X_j|^{1/(\theta-1)} U_j + n^{-1/2} \sum_{j\in K_n(\varepsilon)} |\tilde{X}_j|^{1/(\theta-1)} U_j,$$

and so (2.34) will follow once we show that for all $\delta > 0$

$$\text{(2.37)}\qquad \lim_{\varepsilon\to 0}\overline{\lim_{n\to\infty}}\, \mathbf{P}\big(n^{-1/2} \sum_{j\in K_n(\varepsilon)} |X_j|^{1/(\theta-1)} U_j > \delta\big) = 0,$$

$$\text{(2.38)}\qquad \lim_{\varepsilon\to 0}\overline{\lim_{n\to\infty}}\, \mathbf{P}\big(n^{-1/2} \sum_{j\in K_n(\varepsilon)} |\tilde{X}_j|^{1/(\theta-1)} U_j > \delta\big) = 0.$$

Note that as $n \to \infty$

$$\mathrm{E}\big(n^{-1/2} \sum_{j\in K_n(\varepsilon)} |X_j|^{1/(\theta-1)} U_j\big)^2$$

$$= n^{-1}\big(\mathrm{E}[k_n(\varepsilon)]\big)\,\mathrm{E}[U_1^2]\,\mathrm{E}[|X_1|^{2/(\theta-1)} \mid |X_1| > \varepsilon n^{1/\alpha_x}]$$

$$= \frac{1}{n\mathbf{P}(|X_1| > \varepsilon n^{1/\alpha_x})} \mathrm{E}[k_n(\varepsilon)] \mathrm{E}[U_1^2] \mathrm{E}[|X_1|^{2/(\theta-1)} \mathbf{1}(|X_1| > \varepsilon n^{1/\alpha_x})]$$

$$\sim C_1^{-1} \varepsilon^{-\alpha_x} D_1 \varepsilon^{-\alpha_x} \mathrm{E}[U_1^2] \cdot 0 = 0,$$

and (2.37) follows. The proof of (2.38) is similar and even easier. Hence we have established (2.34).

We now switch to proving (2.35). Since $(\hat{N}_n)$, $(1/\hat{D}_n)$ and $(1/DE_n)$ are all tight, it is enough to prove that

$$\lim_{\varepsilon \to 0} \overline{\lim_{n \to \infty}} \mathbf{P}(|DE_n - \hat{D}_n| > \delta) = 0. \tag{2.39}$$

Notice that

$$DE_n - \hat{D}_n = n^{-\theta/((\theta-1)\alpha_x)} \sum_{j = \{1,2,\ldots,n\} \setminus K_n(\varepsilon)} |X_j|^{\theta/(\theta-1)},$$

and hence

$$DE_n - \hat{D}_n \Rightarrow \int_0^{\varepsilon^{\theta/(\theta-1)}} x N_*(dx) \quad \text{as } n \to \infty,$$

where

$$N_* = \sum_{j=1}^{\infty} \delta_{\{D_1^{\theta/((\theta-1)\alpha_x)} \Gamma_j^{-\theta/(\theta-1)\alpha_x}\}}$$

is the appropriate Poisson random measure. Here (Γ_j) represents the arrival times of a unit rate homogeneous Poisson process on $(0, \infty)$. See, e.g., problem 4.4.2.8 in [13]. Since

$$\int_0^{\varepsilon^{\theta/(\theta-1)}} x N_*(dx) \to 0 \text{ a.s.} \quad \text{as } \varepsilon \to 0,$$

we have established (2.39), and so we have proved (2.35). That completes the proof of (2.13).

The argument in the first boundary case (2.14) is very similar to that of (2.13) above, but instead of the CLT for triangular arrays with a finite variance it uses the general CLT for triangular arrays as, for example, in Theorem 5.3.2 in [9]). For example, the statement (2.30) now reads

$$\hat{N}_n \Rightarrow N(0, \mathrm{E}[|X_1|^{2/(\theta-1)}] D_2)$$

with

$$\hat{N}_n = (n \log n)^{-1/2} \sum_{j=1}^{n} |\hat{X}_j|^{1/(\theta-1)} U_j.$$

Indeed, by the symmetry of (U_i), the third condition in the above-cited theorem holds, and the first condition is weaker than the second one. The second condition reduces to the claim

$$\frac{1}{\log n}\mathrm{E}\left[U_1^2|X_1|^{2/(\theta-1)}\mathbf{1}\left(|X_1|\leqslant \varepsilon n^{1/\alpha_x},\ |U_1|\,|X_1|^{1/(\theta-1)}\leqslant (n\log n)^{1/2}\right)\right]$$
$$\to \mathrm{E}\left[|X_1|^{2/(\theta-1)}\right]D_2 \quad \text{as } n\to\infty,$$

which easily follows from the fact that

$$P(U_1|X_1|^{1/(\theta-1)}>y)\sim \mathrm{E}\left[|X_1|^{2/(\theta-1)}\right]D_2\,y^{-2} \quad \text{as } y\to\infty.$$

The second boundary case (2.16) is entirely similar. The third boundary case (2.18) is also similar, but there is an extra power of the logarithm in that case. This is due to the fact that in this boundary case

$$\mathbf{P}(|X_1|^{2/(\theta-1)}U_1^2>y)\sim D_1 D_2\frac{\log y}{y} \quad \text{as } y\to\infty;$$

see e.g. [4].

We now switch to the second part of the theorem. Here

$$n^d\Delta_n=\frac{n^{-1/((\theta-1)\alpha_x)}\sum_{j=1}^n |X_j|^{1/(\theta-1)}U_j}{n^{-\theta/((\theta-1)\alpha_x)}\sum_{j=1}^n |X_j|^{\theta/(\theta-1)}}.$$

Let $\varepsilon>0$ and write

$$n^d\Delta_n=\frac{\sum_{j=1}^n |n^{-1/\alpha_x}X_j|^{1/(\theta-1)}\mathbf{1}(|n^{-1/\alpha_x}X_j|>\varepsilon)U_j}{\sum_{j=1}^n |n^{-1/\alpha_x}X_j|^{\theta/(\theta-1)}}$$
$$+\frac{\sum_{j=1}^n |n^{-1/\alpha_x}X_j|^{1/(\theta-1)}\mathbf{1}(|n^{-1/\alpha_x}X_j|\leqslant\varepsilon)U_j}{\sum_{j=1}^n |n^{-1/\alpha_x}X_j|^{\theta/(\theta-1)}}$$
$$:=M_n(\varepsilon)+R_n(\varepsilon), \quad n=1,2,\ldots$$

Note that

$$M_n(\varepsilon)\stackrel{d}{=}\frac{\sum_{i=1}^{N_n((\varepsilon,\infty))}K_{i,n}^{1/(\theta-1)}U_i}{\int_0^\infty x^{\theta/(\theta-1)}N_n(dx)}, \quad n=1,2,\ldots,$$

where

$$N_n=\sum_{j=1}^n \delta_{\{n^{-1/\alpha_x}|X_j|\}}, \quad n=1,2,\ldots,$$

and $K_{1,n}\geqslant K_{2,n}\geqslant\ldots\geqslant K_{n,n}$ are the size-ordered points of N_n.

Recalling (see, once again, [13]) that

$$N_n\Rightarrow\sum_{j=1}^\infty \delta_{\{D_1^{1/\alpha_x}\Gamma_j^{-1/\alpha_x}\}}:=N \quad \text{as } n\to\infty$$

weakly in $[-\infty, +\infty]\setminus\{0\}$, we see that

$$M_n(\varepsilon) \Rightarrow \frac{\sum_{i=1}^{N((\varepsilon,\infty))} K_i^{1/(\theta-1)} U_i}{\int_0^\infty x^{\theta/(\theta-1)} N(dx)}$$

$$= D_1^{-1/\alpha_x} \frac{\sum_{j=1}^\infty \Gamma_j^{-1/((\theta-1)\alpha_x)} \mathbf{1}(\Gamma_j^{-1/\alpha_x} > \varepsilon) U_j}{\sum_{j=1}^\infty \Gamma_j^{-\theta/((\theta-1)\alpha_x)}} := L(\varepsilon),$$

weakly as $n \to \infty$, where (K_i) stands for the size-ordered points of N. Note that, almost surely,

$$L(\varepsilon) \to D_1^{-1/\alpha_x} \frac{\sum_{j=1}^\infty \Gamma_j^{-1/(\theta-1)\alpha_x} U_j}{\sum_{j=1}^\infty \Gamma_j^{-\theta/((\theta-1)\alpha_x)}} := L \quad \text{as } \varepsilon \to 0,$$

the right-hand side of (2.22). Therefore, an appeal to Theorem 3.2 in [2] shows that, to prove the latter, it remains to be demonstrated that for any $\lambda > 0$

$$\lim_{\varepsilon \to \infty} \overline{\lim_{n \to \infty}} \, \mathrm{P}(|R_n(\varepsilon)| > \lambda) = 0. \tag{2.40}$$

Clearly, the sequence $\{(\sum_{j=1}^n |n^{-1/\alpha_x} X_j|^{\theta/(\theta-1)})^{-1}\}$ is (asymptotically) tight. Given $\delta > 0$ we can choose $M > 0$ and n_0 such that

$$\mathrm{P}\Big(\sum_{j=1}^n |n^{-1/\alpha_x} X_j|^{\theta/(\theta-1)} \leqslant M\Big) \leqslant \delta \qquad \text{for all } n \geqslant n_0.$$

Then for all $n \geqslant n_0$ and $\lambda > 0$

$$\mathrm{P}(|R_n(\varepsilon)| > \lambda) \leqslant \delta + \mathrm{P}\Big(\Big|\sum_{j=1}^n |n^{-1/\alpha_x} X_j|^{1/(\theta-1)} \mathbf{1}(n^{-1/\alpha_x}|X_j| \leqslant \varepsilon) U_j\Big| > \lambda M\Big).$$

For $K > 0$ we have

$$\mathrm{P}\Big(\Big|\sum_{j=1}^n |n^{-1/\alpha_x} X_j|^{1/(\theta-1)} \mathbf{1}(n^{-1/\alpha_x}|X_j| \leqslant \varepsilon) U_j\Big| > \lambda M\Big) \tag{2.41}$$

$$\leqslant \mathrm{P}\Big(\Big|\sum_{j=1}^n |n^{-1/\alpha_x} X_j|^{1/(\theta-1)} \mathbf{1}(n^{-1/\alpha_x}|X_j| \leqslant \varepsilon,\, n^{-1/\alpha_x}|X_j|\,|U_j|^{\theta-1} \leqslant K) U_j\Big| > \lambda M\Big)$$
$$+ n\mathrm{P}(|X_1|\,|U_1|^{\theta-1} > Kn^{1/\alpha_x}).$$

Keeping K fixed, we have by symmetry, using the equivalence of different moments of Bernoulli random variables (see e.g. Proposition 3.4.1 in [8]), also known as the Khintchine inequalities,

$$\mathrm{P}\Big(\Big|\sum_{j=1}^n |n^{-1/\alpha_x} X_j|^{1/(\theta-1)} \mathbf{1}(n^{-1/\alpha_x}|X_j| \leqslant \varepsilon,\, n^{-1/\alpha_x}|X_j|\,|U_j|^{\theta-1} \leqslant K) U_j\Big| > \lambda M\Big)$$

$$\leqslant \frac{1}{\lambda M} \mathrm{E}\Big[\Big|\sum_{j=1}^n |n^{-1/\alpha_x} X_j|^{1/(\theta-1)} \mathbf{1}(n^{-1/\alpha_x}|X_j| \leqslant \varepsilon,\, n^{-1/\alpha_x}|X_j|\,|U_j|^{\theta-1} \leqslant K) U_j\Big|\Big]$$

$$\leqslant cn^{-1/((\theta-1)\alpha_x)} \mathrm{E}\Big[\sum_{j=1}^{n} |X_j|^{2/(\theta-1)} \mathbf{1}(|X_j| \leqslant \varepsilon n^{1/\alpha_x}, |X_j||U_j|^{\theta-1} \leqslant Kn^{1/\alpha_x}) U_j^2\Big]^{1/2}$$

$$\leqslant cn^{1/2-1/((\theta-1)\alpha_x)} \big(\mathrm{E}\big[|X_1|^{2/(\theta-1)} U_1^2 \mathbf{1}(|X_1| \leqslant \varepsilon n^{1/\alpha_x}, |X_1||U_1|^{\theta-1} \leqslant Kn^{1/\alpha_x})\big]\big)^{1/2}.$$

Here and in the sequel, c is an arbitrary finite and positive constant that does not have to be the same every time it appears. By the assumption (2.20) we have

$$\mathrm{E}\big[|X_1|^{2/(\theta-1)} U_1^2 \mathbf{1}(|X_1| \leqslant \varepsilon n^{1/\alpha_x}, |X_1||U_1|^{\theta-1} \leqslant Kn^{1/\alpha_x})\big]$$
$$\sim a(\varepsilon)(n^{1/\alpha_x})^{2/(\theta-1)-\alpha_x} := a(\varepsilon) n^{\rho} \quad \text{as } n \to \infty,$$

where $a(\varepsilon) \to 0$ as $\varepsilon \to 0$, so that the right-hand side of (2.41) is less than or equal to

$$c\big(a(\varepsilon)\big)^{1/2} + nP(|X_1||U_1|^{\theta-1} > Kn^{1/\alpha_x}),$$

since $\alpha_u \geqslant 2$. Now (2.40) follows after letting $K \to \infty$ (we are using, once again, the fact that $\alpha_u \geqslant 2$), and so we have proved (2.22). ∎

Scenario 2. Suppose that

$$(2.42) \qquad 0 < \alpha_x \leqslant 1, \quad 0 < \alpha_u < 2, \quad \text{and} \quad \alpha_u \geqslant \alpha_x.$$

We are now on one side of both critical boundaries (2.8) and (2.9), and the different ranges of θ appear here depending on which of the two elements under the minimum in (2.9) is smaller. The possibilities are described in the following theorem.

THEOREM 2.2. *Assume that* (2.42) *holds.*

(a) *Suppose that*

$$\theta \geqslant \frac{\alpha_u}{\alpha_x} + 1.$$

Then the exponent d in (2.7) *is given by*

$$(2.43) \qquad d = \frac{\alpha_u \theta - (\theta-1)\alpha_x}{(\theta-1)\alpha_u \alpha_x}.$$

Specifically, in the non-boundary case

$$(2.44) \qquad \theta > \frac{\alpha_u}{\alpha_x} + 1$$

we have

$$(2.45) \qquad n^d \Delta_n \Rightarrow \frac{C_{\alpha_u}^{-1/\alpha_u} D_2^{1/\alpha_u} (\mathrm{E}[|X_1|^{\alpha_u/(\theta-1)}])^{1/\alpha_u}}{C_{(\theta-1)\alpha_x/\theta}^{-\theta/((\theta-1)\alpha_x)} D_1^{\theta/((\theta-1)\alpha_x)}} \frac{S_{\alpha_u}(1, 0, 0)}{S_{(\theta-1)\alpha_x/\theta}(1, 1, 0)}$$

weakly, with the random variables on the right-hand side being independent.

In the boundary case

$$(2.46) \qquad \theta = \frac{\alpha_u}{\alpha_x} + 1$$

we have

$$(2.47)\qquad \frac{n^d}{(\log n)^{1/\alpha_u}}\Delta_n \Rightarrow \frac{C_{\alpha_u}^{-1/\alpha_u}(D_1 D_2)^{1/\alpha_u}}{C_{(\theta-1)\alpha_x/\theta}^{-\theta/((\theta-1)\alpha_x)} D_1^{\theta/((\theta-1)\alpha_x)}} \frac{S_{\alpha_u}(1, 0, 0)}{S_{(\theta-1)\alpha_x/\theta}(1, 1, 0)}$$

weakly, the random variables on the right-hand side above still being independent.

(b) *Suppose that*

$$\theta < \frac{\alpha_u}{\alpha_x} + 1.$$

Then the exponent d *in* (2.7) *is given by* (2.21), *and* (2.22) *holds.*

Proof. (a) In the non-boundary case (2.44) we have

$$(2.48)\qquad n^d \Delta_n = \frac{n^{-1/\alpha_u}\sum_{j=1}^n |X_j|^{1/(\theta-1)} U_j}{n^{-\theta/((\theta-1)\alpha_x)}\sum_{j=1}^n |X_j|^{\theta/(\theta-1)}}.$$

The proof is parallel to that of (2.13). We use the notation of (2.24), (2.25), (2.26) and (2.28), while instead of (2.27) we use, obviously,

$$\hat{N}_n = n^{-1/\alpha_u} \sum_{j=1}^{n} |\hat{X}_j|^{1/(\theta-1)} U_j,$$

with $\{\hat{X}_j\}$ given by (2.29). In particular, (2.31) still holds. We will show now that

$$(2.49)\qquad \hat{N}_n \Rightarrow C_{\alpha_u}^{-1/\alpha_u} D_2^{1/\alpha_u} (\mathrm{E}[|X_1|^{\alpha_u/(\theta-1)}])^{1/\alpha_u} S_{\alpha_u}(1, 0, 0)$$

weakly as $n \to \infty$. Since, by the CLT,

$$(2.50)\qquad N_n \Rightarrow C_{\alpha_u}^{-1/\alpha_u} D_2^{1/\alpha_u} (\mathrm{E}[|X_1|^{\alpha_u/(\theta-1)}])^{1/\alpha_u} S_{\alpha_u}(1, 0, 0)$$

weakly as $n \to \infty$ (see e.g. Chapter XVII in [7]), (2.49) will follow if we check that

$$(2.51)\qquad \hat{N}_n - N_n \to 0 \text{ in probability} \quad \text{as } n \to \infty.$$

Now,

$$(2.52)\qquad \hat{N}_n - N_n = n^{-1/\alpha_u} \sum_{j \in K_n(\varepsilon)} U_j(|\hat{X}_j|^{1/(\theta-1)} - |X_j|^{1/(\theta-1)})$$

$$= n^{-1/\alpha_u} \sum_{j \in K_n(\varepsilon)} U_j |\hat{X}_j|^{1/(\theta-1)} - n^{-1/\alpha_u} \sum_{j \in K_n(\varepsilon)} U_j |X_j|^{1/(\theta-1)}.$$

Hence, (2.51) will follow once we prove that

$$(2.53)\qquad n^{-1/\alpha_u} \sum_{j \in K_n(\varepsilon)} U_j |X_j|^{1/(\theta-1)} \to 0 \text{ in probability} \quad \text{as } n \to \infty,$$

$$(2.54)\qquad n^{-1/\alpha_u} \sum_{j \in K_n(\varepsilon)} U_j |\hat{X}_j|^{1/(\theta-1)} \to 0 \text{ in probability} \quad \text{as } n \to \infty.$$

Consider (2.53). Let $0 < p < 1 \wedge \alpha_u$. This gives us

$$\begin{aligned} &\mathrm{E}\big[\big|n^{-1/\alpha_u} \sum_{j\in K_n(\varepsilon)} U_j |X_j|^{1/(\theta-1)}\big|^p\big] \\ &\quad \leqslant n^{-p/\alpha_u} \cdot \mathrm{E}[k_n(\varepsilon)]\,\mathrm{E}[U_1]^p\,\mathrm{E}[|X_1|^{p/(\theta-1)} \mathbf{1}(|X_1| > \varepsilon n^{1/\alpha_x})] \\ &\quad = n^{-p/\alpha_u}\big(n\mathbf{P}(|X_1| > \varepsilon n^{1/\alpha_x})\big)\,\mathrm{E}[U_1]^p \\ &\qquad \times \left(\frac{1}{\mathbf{P}(|X_1| > \varepsilon n^{1/\alpha_x})}\mathrm{E}[|X_1|^{p/(\theta-1)} \mathbf{1}(|X_1| > \varepsilon n^{1/\alpha_x})]\right) \\ &\quad \sim cn^{-p/\alpha_u+1}\, n^{-1+p/((\theta-1)\alpha_x)} = cn^{-p(1/\alpha_u - 1/((\theta-1)\alpha_x))} \to 0 \quad \text{as } n\to\infty. \end{aligned}$$

Hence (2.53) holds, and the proof of (2.54) is the same, but easier. The rest of the proof of (2.54) is the same as that of (2.13) above.

The proof of (2.47) in the boundary case (2.46) is the same as that of (2.45), except that (2.50) is now replaced by

$$(2.55)\quad n^{-1/\alpha_u}(\log n)^{-1/\alpha_u} \sum_{j=1}^{n} |X_j|^{1/(\theta-1)} U_j \Rightarrow C_{\alpha_u}^{-1/\alpha_u}(D_1 D_2)^{1/\alpha_u} S_{\alpha_u}(1, 0, 0)$$

weakly as $n\to\infty$ (see [7]).

The argument for the second part of the theorem is identical to that of the second part of Theorem 2.1. ∎

Scenario 3. The next scenario is

$$(2.56)\qquad 0 < \alpha_u \leqslant 1 \quad \text{and} \quad \alpha_u < \alpha_x.$$

The possibilities provided by the different ranges of θ are described in the following theorem.

THEOREM 2.3. *Assume that* (2.56) *holds.*

(a) *Suppose that*

$$\theta \geqslant \frac{\alpha_x}{\alpha_x - \alpha_u}.$$

In this case the estimator (2.4) *is not consistent.*

(b) *Suppose that*

$$\frac{\alpha_u}{\alpha_x} + 1 \leqslant \theta < \frac{\alpha_x}{\alpha_x - \alpha_u}.$$

Here d is given by (2.43), *and in the non-boundary case* (2.45) *holds. Furthermore, in the only boundary case* (2.47) *holds.*

(c) *Suppose that*

$$\theta < \frac{\alpha_u}{\alpha_x} + 1.$$

Here d is given by (2.21), *and* (2.22) *holds.*

Proof. For part (a) we claim that Δ_n does not converge in probability to 0. Indeed, let d be given by (2.43), and notice that now $d \leqslant 0$. Since the reciprocal of the fraction on the right-hand side of (2.48) is clearly tight, we see that Δ_n cannot converge to zero (it is not even tight if $\theta > \alpha_x/(\alpha_x - \alpha_u)$).

The proof of part (b) is identical to the proof of part (a) of Theorem 2.2, while the proof of part (c) of the present theorem is identical to the proof of part (b) of Theorem 2.1. ∎

Scenario 4. Suppose now that

$$1 < \alpha_x \leqslant 2 \quad \text{and} \quad \alpha_u \geqslant 2. \tag{2.57}$$

The two ranges of θ we consider are on different sides of the boundary (2.8), as described in the following theorem.

THEOREM 2.4. *Assume that* (2.57) *holds.*

(a) *Suppose that*

$$\theta \geqslant \frac{\alpha_x}{\alpha_x - 1}.$$

Then the exponent d is

$$d = \tfrac{1}{2}. \tag{2.58}$$

Specifically, in the non-boundary case

$$\alpha_u > 2 \quad \textit{and} \quad \theta > \frac{\alpha_x}{\alpha_x - 1}, \tag{2.59}$$

we have

$$n^d \Delta_n \Rightarrow \frac{(\mathrm{E}[|X_1|^{2/(\theta-1)}])^{1/2} (\mathrm{E}[U_1^2])^{1/2}}{\mathrm{E}[|X_1|^{\theta/(\theta-1)}]} N(0, 1) \tag{2.60}$$

weakly as $n \to \infty$.

In the first boundary case

$$\alpha_u = 2 \quad \textit{and} \quad \theta > \frac{\alpha_x}{\alpha_x - 1} \tag{2.61}$$

we have

$$\frac{n^d}{(\log n)^{1/2}} \Delta_n \Rightarrow \frac{(D_2 \mathrm{E}[|X_1|^{2/(\theta-1)}])^{1/2}}{\mathrm{E}[|X_1|^{\theta/(\theta-1)}]} N(0, 1). \tag{2.62}$$

In the second boundary case

$$\alpha_u > 2, \quad \alpha_x > 2 \quad \textit{and} \quad \theta = \frac{\alpha_x}{\alpha_x - 1} \tag{2.63}$$

we have

$$(n^d \log n)\, \Delta_n \Rightarrow \frac{(\mathrm{E}\,[|X_1|^{2/(\theta-1)}])^{1/2}\,(\mathrm{E}\,[U_1^2])^{1/2}}{D_1} N(0,\, 1)$$

weakly as $n \to \infty$.

In the boundary case

$$\alpha_u > 2, \quad \alpha_x = 2 \quad \textit{and} \quad \theta = \frac{\alpha_x}{\alpha_x - 1} \tag{2.64}$$

we have

$$n^d (\log n)^{1/2}\, \Delta_n \Rightarrow \frac{(\mathrm{E}\,[|U_1|^2])^{1/2}}{D_1^{1/2}} N(0,\, 1).$$

Next, if

$$\alpha_u = 2, \quad \alpha_x > 2 \quad \textit{and} \quad \theta = \frac{\alpha_x}{\alpha_x - 1} \tag{2.65}$$

we have

$$n^d (\log n)^{1/2}\, \Delta_n \Rightarrow \frac{(D_2\, \mathrm{E}\,[|X_1|^{2/(\theta-1)}])^{1/2}}{D_1} N(0,\, 1).$$

Finally, in the boundary case

$$\alpha_u = 2, \quad \alpha_x = 2 \quad \textit{and} \quad \theta = \frac{\alpha_x}{\alpha_x - 1} \tag{2.66}$$

we will have

$$n^d\, \Delta_n \Rightarrow \frac{D_2^{1/2}}{D_1^{1/2}} N(0,\, 1).$$

(b) *Suppose that*

$$\frac{2+\alpha_x}{\alpha_x} \leqslant \theta < \frac{\alpha_x}{\alpha_x - 1}.$$

The exponent d here is given by (2.11), *and, in fact, in the non-boundary case* (2.12), *the convergence result* (2.13) *holds, whereas in the three boundary cases* (2.14), (2.16) *and* (2.18), *the convergence results* (2.15), (2.17) *and* (2.19), *respectively, hold.*

(c) *Suppose that*

$$\theta < \frac{2+\alpha_x}{\alpha_x}.$$

Here, the exponent d is given by (2.21), *and the convergence result* (2.22) *holds.*

Proof. (a) We start with the non-boundary case (2.59). Here

$$n^d\, \Delta_n = \frac{n^{-1/2} \sum_{j=1}^{n} |X_j|^{1/(\theta-1)}\, U_j}{n^{-1} \sum_{j=1}^{n} |X_j|^{\theta/(\theta-1)}}, \tag{2.67}$$

and the strong LLN applies in the denominator, while the CLT for iid random variables with a finite variance applies in the numerator. In the first boundary case (2.61) we use the general CLT for iid random variables in the numerator (see e.g. Proposition 5.3.3 in [9]).

In the second boundary case (2.63) the CLT for iid random variables with a finite variance still applies in the numerator in (2.67), but the finite law means LLN no longer holds in the denominator. Instead, we use the weak LLN

$$\frac{1}{n\log n}\sum_{j=1}^{n}|X_j|^{\theta/(\theta-1)} \to D_1 \tag{2.68}$$

in probability as $n \to \infty$; see Theorem VII.7.2 in [7]. To see why this is true, note that the function μ in (7.10), page 236 of [7], satisfies, in the case of iid random variables $|X_i|^{1/\alpha_x}$, $\mu(s) \sim D_1 \log s$ as $s \to \infty$, and so the corresponding sequence (s_n) satisfies $s_n \sim D_1 n \log n$, from which (2.68) follows.

In the third boundary case (2.64) we similarly use both the general CLT for iid random variables and the weak LLN (2.64).

The argument is similar in the fourth boundary case (2.65). Also similar is the last boundary case, (2.66), and we refer the reader to the discussion after (2.19).

For the parts (b) and (c) the argument is identical to that of parts (a) and (b), respectively, of Theorem 2.1. ■

Scenario 5. Suppose that

$$\alpha_x > 1, \quad 1 < \alpha_u < 2 \quad \text{and} \quad \alpha_u \leqslant \frac{\alpha_x}{\alpha_x - 1}. \tag{2.69}$$

The various possibilities detailing the side of the critical boundary (2.8) we are on, and the tail relationships between the random variables involved are described in the following theorem.

THEOREM 2.5. *Assume that* (2.69) *holds.*

(a) *Suppose that*

$$\theta \geqslant \frac{\alpha_x}{\alpha_x - 1}.$$

Here the exponent d *is given by*

$$d = 1 - \frac{1}{\alpha_u}. \tag{2.70}$$

Specifically, in the non-boundary case

$$\theta > \frac{\alpha_x}{\alpha_x - 1} \tag{2.71}$$

we have

$$n^d \Delta_n \Rightarrow \frac{C_{\alpha_u}^{-1/\alpha_u} D_2^{1/\alpha_u} (\mathrm{E}[|X_1|^{\alpha_u/(\theta-1)}])^{1/\alpha_u}}{\mathrm{E}[|X_1|^{\theta/(\theta-1)}]} S_{\alpha_u}(1, 0, 0). \tag{2.72}$$

In the boundary case

$$\theta = \frac{\alpha_x}{\alpha_x - 1} \quad \text{and} \quad \alpha_u < \frac{\alpha_x}{\alpha_x - 1} \tag{2.73}$$

we have

$$(n^d \log n) \Delta_n \Rightarrow \frac{C_{\alpha_u}^{-1/\alpha_u} D_2^{1/\alpha_u} (\mathrm{E}[|X_1|^{\alpha_u/(\theta-1)}])^{1/\alpha_u}}{D_1} S_{\alpha_u}(1, 0, 0)$$

weakly as $n \to \infty$, *whereas in the second boundary case*

$$\theta = \frac{\alpha_x}{\alpha_x - 1} \quad \text{and} \quad \alpha_u = \frac{\alpha_x}{\alpha_x - 1} \tag{2.74}$$

we have

$$n^d (\log n)^{1-1/\alpha_u} \Delta_n \Rightarrow \frac{C_{\alpha_u}^{-1/\alpha_u} (D_1 D_2)^{1/\alpha_u}}{D_1} S_{\alpha_u}(1, 0, 0)$$

weakly as $n \to \infty$.

(b) *Suppose that*

$$\frac{\alpha_u}{\alpha_x} + 1 \leqslant \theta < \frac{\alpha_x}{\alpha_x - 1}.$$

Here the exponent d is given by (2.43), *and the convergence results* (2.45) *and* (2.47) *still hold, in the non-boundary and boundary cases, respectively.*

(c) *Suppose that*

$$\theta < \frac{\alpha_u}{\alpha_x} + 1.$$

Here the exponent d is again given by (2.21), *and the convergence result* (2.22) *holds.*

Proof. (a) Since in the non-boundary case (2.71) we have

$$n^d \Delta_n = \frac{n^{-1/\alpha_u} \sum_{j=1}^{n} |X_j|^{1/(\theta-1)} U_j}{n^{-1} \sum_{j=1}^{n} |X_j|^{\theta/(\theta-1)}},$$

we can use the CLT in the numerator and the LLN in the denominator to obtain the result. In the first boundary case (2.73) we can still apply the CLT in the numerator, but this time we need to use the weak LLN (2.68) in the denominator. We treat the denominator in the same way in the second boundary case (2.74), but this time we use the version of the CLT given in (2.55) for the numerator.

The argument for the part (b) is identical to that for part (a) of Theorem 2.2, while the argument for part (c) is the same as that of part (b) of Theorem 2.1. ■

Scenario 6. Suppose that

$$\alpha_x > 2 \quad \text{and} \quad \alpha_u \geqslant 2. \tag{2.75}$$

Here the different choices of θ place the situation on different sides of the critical boundaries (2.8) and (2.9), as described in the following theorem.

THEOREM 2.6. *Assume that* (2.75) *holds.*

(a) *Suppose that*

$$\theta \geqslant \frac{2+\alpha_x}{\alpha_x}.$$

Here d is given by (2.58), *and in the non-boundary case*

$$\theta > \frac{2+\alpha_x}{\alpha_x} \quad \text{and} \quad \alpha_u > 2$$

we have the convergence result (2.60). *In the boundary case*

$$\theta > \frac{2+\alpha_x}{\alpha_x} \quad \text{and} \quad \alpha_u = 2$$

we have the weak convergence in (2.62), *and in the boundary case*

$$\theta = \frac{2+\alpha_x}{\alpha_x} \quad \text{and} \quad \alpha_u = 2$$

we have

$$\frac{n^d}{\log n} \Delta_n \Rightarrow \frac{(D_1 D_2)^{1/2}}{\mathrm{E}[|X_1|^{\theta/(\theta-1)}]} N(0, 1)$$

weakly as $n \to \infty$.

(b) *Suppose that*

$$\frac{\alpha_x}{\alpha_x - 1} \leqslant \theta < \frac{2+\alpha_x}{\alpha_x}.$$

In this case the exponent d is given by

$$d = \frac{(\theta-1)\alpha_x - 1}{(\theta-1)\alpha_x}. \tag{2.76}$$

Specifically, in the non-boundary case

$$\theta > \frac{\alpha_x}{\alpha_x - 1} \tag{2.77}$$

we have

(2.78) $$n^d \Delta_n \Rightarrow \frac{C_{(\theta-1)\alpha_x}^{-1/((\theta-1)\alpha_x)} D_1^{1/((\theta-1)\alpha_x)} (\mathrm{E}[|U_1|^{(\theta-1)\alpha_x}])^{1/((\theta-1)\alpha_x)}}{\mathrm{E}[|X_1|^{\theta/(\theta-1)}]} \times S_{(\theta-1)\alpha_x}(1, 0, 0)$$

weakly as $n \to \infty$.

In the boundary case

(2.79) $$\theta = \frac{\alpha_x}{\alpha_x - 1}$$

we have

(2.80) $$(n^d \log n)\Delta_n \Rightarrow \frac{C_{(\theta-1)\alpha_x}^{-1/((\theta-1)\alpha_x)} D_1^{1/((\theta-1)\alpha_x)} (\mathrm{E}[|U_1|^{(\theta-1)\alpha_x}])^{1/((\theta-1)\alpha_x)}}{D_1} S_{(\theta-1)\alpha_x}(1, 0, 0)$$

weakly as $n \to \infty$.

(c) *Suppose that*

(2.81) $$\theta < \frac{\alpha_x}{\alpha_x - 1}.$$

Here d *is given by* (2.21), *and the weak convergence* (2.22) *holds.*

Proof. The proof of part (a) is the same as that of part (a) of Theorem 2.4; for the last boundary case see the discussion after (2.19).

In part (b) note that

(2.82) $$n^d \Delta_n = \frac{n^{-1/(\theta-1)\alpha_x} \sum_{j=1}^n |X_j|^{1/(\theta-1)} U_j}{n^{-1} \sum_{j=1}^n |X_j|^{\theta/(\theta-1)}}.$$

In the non-boundary case (2.77) we use the CLT for sums of iid random variables with a finite variance in the numerator and the LLN for iid random variables with a finite mean in the denominator to obtain the result. In the boundary case (2.79) the argument is similar, but this time we use the general weak LLN in the denominator.

The argument in part (c) is the same as that of part (b) of Theorem 2.1. ∎

Scenario 7. Suppose that

(2.83) $$\alpha_x > 2 \quad \text{and} \quad \frac{\alpha_x}{\alpha_x - 1} < \alpha_u < 2.$$

Here the different ranges of θ will determine which of the two elements under the minimum in (2.9) is smaller, and which side of the critical boundary (2.8) we are on.

THEOREM 2.7. *Assume that* (2.83) *holds.*

(a) *Suppose that*

$$\theta \geqslant \frac{\alpha_u}{\alpha_x}+1.$$

Here d *is given by* (2.70), *and in the non-boundary case*

$$\theta > \frac{\alpha_u}{\alpha_x}+1$$

the weak convergence in (2.72) *holds. In the boundary case*

$$\theta = \frac{\alpha_u}{\alpha_x}+1 \tag{2.84}$$

we have

$$\frac{n^d}{(\log n)^{1/\alpha_u}}\Delta_n \Rightarrow \frac{C_{\alpha_u}^{-1/\alpha_u}(D_1 D_2)^{1/\alpha_u}}{\mathrm{E}\,[|X_1|^{\theta/(\theta-1)}]} S_{\alpha_u}(1, 0, 0)$$

weakly as $n \to \infty$.

(b) *Suppose that*

$$\frac{\alpha_x}{\alpha_x-1} \leqslant \theta < \frac{\alpha_u}{\alpha_x}+1.$$

Here d *is given by* (2.76), *and in the non-boundary case*

$$\theta > \frac{\alpha_x}{\alpha_x-1}$$

the weak convergence in (2.78) *holds. In the boundary case*

$$\theta = \frac{\alpha_x}{\alpha_x-1}$$

the weak convergence in (2.80) *holds.*

(c) *Suppose that*

$$\theta < \frac{\alpha_x}{\alpha_x-1}.$$

Here d *is given by* (2.21), *and we have the weak convergence in* (2.22).

Proof. In part (a) the proof in the non-boundary case is the same as that in part (a) of Theorem 2.5. In the boundary case (2.84) in the numerator we appeal to the general CLT for iid summands to obtain the result.

The argument for part (b) is the same as that for part (b) of Theorem 2.6, while the argument for part (c) of the present theorem is the same as that for part (b) of Theorem 2.1. ∎

Recall that the exponent d describes the rate of convergence of the estimator (2.4); see (2.6) and (2.7). Under each one of the seven scenarios this exponent is a different function of the parameter θ.

3. WHAT θ SHOULD ONE USE?

We start with a plot showing how the scenarios partition the positive quadrant.

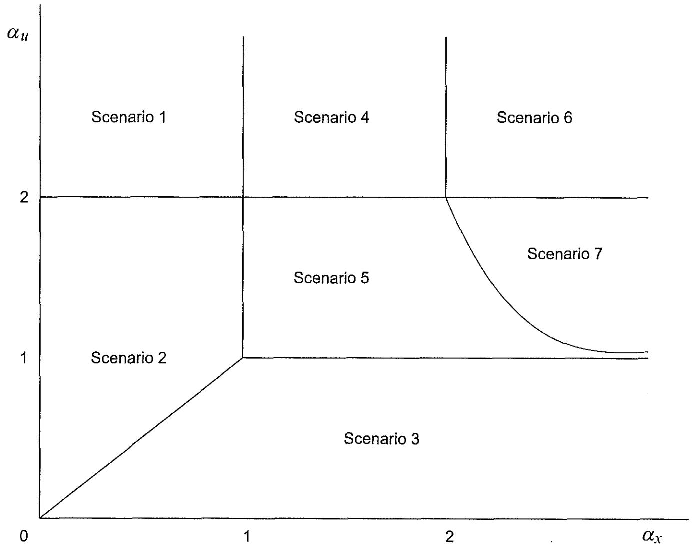

FIGURE 1. All possible scenarios

The exponent d describes the rate of convergence of the estimator (2.4); see (2.6) and (2.7). Under each one of the seven scenarios this exponent is a different function of the parameter θ. The knowledge of α_x and α_u allows us to select the θ that leads to the highest possible rate of decay of Δ_n, i.e. the highest possible d. Recall that $d(\theta)$ is a *non-increasing* function of θ in Scenarios 1 through 5, and a *non-decreasing* function of θ in Scenarios 6 and 7.

What θ do we choose if the α_x and α_u are unknown or rather we do not know them precisely? This is a common situation since the precision of even the best non-parametric estimators of the tail exponents is not very high; see e.g. [5].

Clearly, the tighter bounds on α_x and α_u we have, the easier it is to select a good θ. In this section we will consider several possible situations. The reader is invited to consider additional possibilities. We will only consider the cases $\alpha_x \geqslant 1$ and $\alpha_u \geqslant 1$ here, as those are of relevance in empirical analysis.

Suppose first that we know that

$$\alpha_x \geqslant 1 \quad \text{and} \quad \alpha_u \geqslant 2.$$

Then the choice of $\theta = 2$ always leads to the highest possible rate of decay of Δ_n, i.e. the highest possible d. Indeed, if $\alpha_x \leqslant 2$, then Scenario 4 is in force (the boundary case $\alpha_x = 1$ does not distinguish between Scenarios 1 and 4), and since

$$\frac{2+\alpha_x}{\alpha_x} = 1+\frac{2}{\alpha_x} \geqslant 2,$$

we obtain the optimal $d = 1/\alpha_x$. On the other hand, if $\alpha_x > 2$, then Scenario 6 is in force, and since

$$\frac{2+\alpha_x}{\alpha_x} = 1-\frac{2}{\alpha_x} < 2,$$

we obtain the optimal $d = 1/2$.

On the other hand, suppose we know that

$$1 \leqslant \alpha_x < 2 \quad \text{and} \quad 1 \leqslant \alpha_u < 2.$$

Then any choice of

$$1 < \theta \leqslant 3/2 \tag{3.1}$$

always leads to the highest possible rate of decay of Δ_n (highest possible d). Indeed, Scenario 5 is in force and

$$\frac{\alpha_u}{\alpha_x}+1 > \frac{1}{2}+1 = \frac{3}{2} \geqslant \theta,$$

and we obtain the highest possible value of $d = 1/\alpha_x$.

Note that in the above case, and with the choice of θ we are recommending, we will always have $d \geqslant 1/2$.

Unfortunately, in the range $1 \leqslant \alpha_u < 2$, if α_x can be bigger than 2, no such efficiency is possible.

To measure the relative efficiency of a given choice of θ, let us introduce the notation

$$R(\theta;\, \alpha_x,\, \alpha_u) = \frac{d(\theta;\, \alpha_x,\, \alpha_u)}{d^*(\alpha_x,\, \alpha_u)},$$

where $d(\theta;\, \alpha_x,\, \alpha_u)$ is the value of d corresponding to θ, α_x, α_u and

$$d^*(\alpha_x,\, \alpha_u) = \max_{\theta > 1} d(\theta;\, \alpha_x,\, \alpha_u).$$

For a set A of (α_x, α_u) let

$$\mathfrak{R}_A(\theta) = \inf_{(\alpha_x,\alpha_u)\in A} R(\theta;\, \alpha_x,\, \alpha_u)$$

be the worst efficiency of a given choice of θ. We may then look for a *maxmin* value θ_A such that

$$\mathfrak{R}_A(\theta_A) = \max_{\theta > 1} \inf_{(\alpha_x, \alpha_u) \in A} R(\theta; \alpha_x, \alpha_u) := R_A.$$

If $A \supseteq (2, \infty) \times (1, 2)$, then $R_A = 0$.

Indeed, for a given $\theta > 1$, choose α_x so large that $\alpha_x/(\alpha_x - 1) < \theta$ and let $\alpha_u \downarrow 1$. Then we will be eventually within Scenario 5, and

$$R(\theta; \alpha_x, \alpha_u) = \frac{1 - 1/\alpha_u}{1/\alpha_x} \to 0.$$

Hence $\mathfrak{R}_A(\theta) = 0$ for all $\theta > 1$, and so $R_A = 0$, as claimed.

If, however, α_x cannot be arbitrarily large, then things are better.

Let $\alpha^* > 2$, and $A = [1, \alpha^*] \times [1, 2)$. Then any θ in the range

$$1 < \theta \leqslant \frac{\alpha^{*2} + 2}{\alpha^{*2} - \alpha^* + 2} \tag{3.2}$$

is a θ_A. Furthermore,

$$R_A = 2/\alpha^*. \tag{3.3}$$

To prove this, consider first the range

$$\theta > \frac{\alpha^*}{\alpha^* - 1}.$$

Here, as in the case $A = (2, \infty) \times (1, 2)$, we see that $\mathfrak{R}_A(\theta) = 0$. Next, we consider the range

$$\frac{\alpha^* + 1}{\alpha^*} < \theta \leqslant \frac{\alpha^*}{\alpha^* - 1}. \tag{3.4}$$

Note that

$$\inf_{\substack{(\alpha_x, \alpha_u) \times A \\ \alpha_u \leqslant \alpha_x/(\alpha_x - 1)}} R(\theta; \alpha_x, \alpha_u) = \frac{\theta}{\theta - 1} - \alpha^*$$

and is achieved when $\alpha_x \uparrow \alpha^*$ and $\alpha_u \downarrow 1$. On the other hand,

$$\inf_{\substack{(\alpha_x, \alpha_u) \times A \\ \alpha_u > \alpha_x/(\alpha_x - 1)}} R(\theta; \alpha_x, \alpha_u) = \frac{2}{\alpha^*}$$

and is achieved when $\alpha_x \uparrow \alpha^*$ and $\alpha_u \uparrow 2$. Therefore, in the range (3.4) we have

$$\mathfrak{R}_A(\theta) = \min\left(\frac{\theta}{\theta - 1} - \alpha^*, \frac{2}{\alpha^*}\right) = \frac{2}{\alpha^*}$$

if $(\alpha^* + 1)/\alpha^* < \theta \leqslant (\alpha_x^2 + 2)/(\alpha_x^2 - \alpha_x + 2)$. Furthermore, we have $\mathfrak{R}_A(\theta) < 2/\alpha^*$ if $(\alpha_x^2 + 2)/(\alpha_x^2 - \alpha_x + 2) < \theta \leqslant \alpha^*/(\alpha^* - 1)$.

Similarly, in the range

$$1 < \theta \leqslant \frac{\alpha^* + 1}{\alpha^*},$$

we obtain

$$\mathfrak{R}_A(\theta) = 2/\alpha^*.$$

Therefore, both (3.2) and (3.3) follow. In this situation we can guarantee $d \geqslant 2/(\alpha^*)^2$ with the choice of θ recommended above.

The above discussion of the ways to select the parameter θ focuses on the rate of convergence to the true value, which is clearly the single most important criterion. With the rate of convergence kept fixed, however, other things become important. Among them is the spread of the limiting distribution. To compare such spreads, and hence to be able to tell more about good ways to select θ, we performed a simulations study.

Design of simulation. From the viewpoint of empirical evidence, we consider $\alpha_x \in [1, 2)$ and $\alpha_u \in [1, 2)$. To implement data-generating processes, we have selected $\alpha_x, \alpha_u = 1, 1.2, 1.4, 1.6, 1.8, 1.99$.[4] For sample size we choose

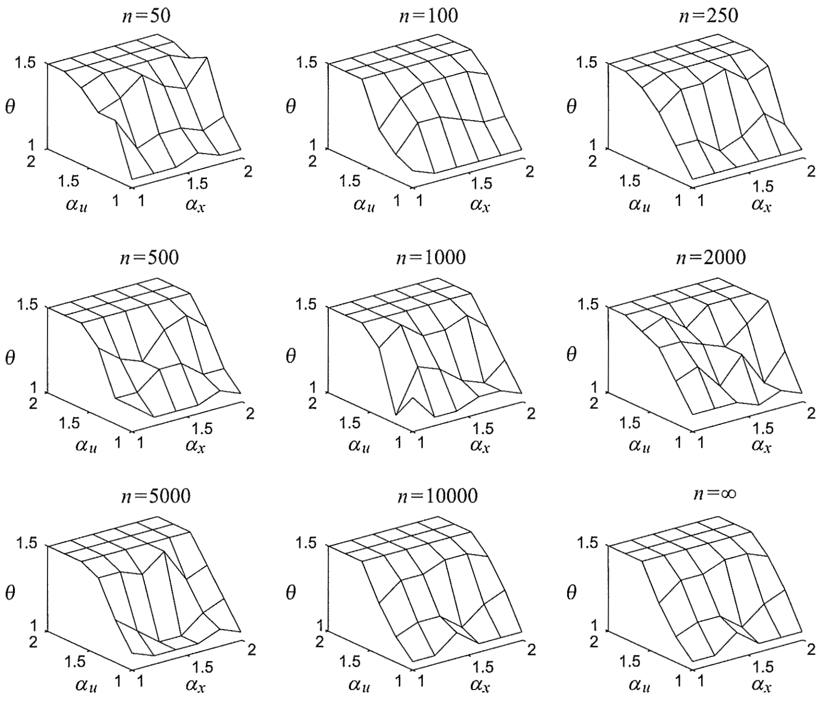

FIGURE 2. θ for selected α_x and α_u

[4] In order to better see the behavior of the estimates near the boundary points 1 and 2, a more detailed selection was used in another simulation. The results show that the transition from 1 to numbers bigger than 1 (1.01 and 1.05 were additionally chosen in the simulation) and from 2 to numbers smaller than 2 (1.99 and 1.95) is smooth.

$n = 50$, 100, 250, 500, 1000, 2000, 5000, 10 000 and ∞, where the limiting distributions are calculated from Scenario 5. We use a length of quantile $\xi_{0.975} - \xi_{0.025}$ as a spread measure, where ξ_p is the pth quantile of the simulated distribution of (2.5).[5] According to the recommendation in (3.1), we use $\theta = (1, 1.5]$. For implementation we have selected $\theta = 1.05$, 1.1, 1.15, 1.2, 1.25, 1.3, 1.35, 1.4, 1.45, 1.5. To determine simulated densities for each estimate, 10 000 replications were made. Figure 2 shows the θ minimizing the spread for selected α_x, α_u and sample sizes.

The selected θ shows noticeable irregularity, even for large samples. Nevertheless, some useful rules for choosing the θ can be formulated as

$$\theta = \begin{cases} \alpha_u & \text{if } \alpha_u < 1.5, \\ 1.5 & \text{if } \alpha_u \geqslant 1.5. \end{cases}$$

Here, the parameter α_u plays a key role, while the role of α_x and the sample size seem to be less important.

4. CONCLUDING REMARKS

One can see that blindly using the OLS approach $\theta = 2$ can lead to very inefficient estimators of the regression coefficient. A much better approach is to take the tails into account. Even if the tails of the regressors and disturbances are known only approximately, this can still provide valuable information for selecting a good value of θ, and hence constructing a more efficient estimator. Iterated procedures in which the tails and the regression coefficient are estimated simultaneously should be considered.

REFERENCES

[1] L. Bachelier, *Théorie de la spéculation*, Ann. Sci. Ecole Norm. Sup. III-17 (1900), pp. 21–86. Translated in: *The Random Character of Stock Market Prices*, P.H. Cootner (Ed.), MIT Press, Cambridge 1964, pp. 17–78.

[2] P. Billingsley, *Convergence of Probability Measures*, 2nd edition, Wiley, New York 1999.

[3] R. Blattberg and T. Sargent, *Regression with non-Gaussian stable disturbances: some sampling results*, Econometrica 39 (1971), pp. 501–510.

[4] D. Cline, *Convolutions tails, product tails and domain of attraction*, Probab. Theory Related Fields 72 (1986), pp. 529–557.

[5] P. Embrechts, C. Klüppelberg and T. Mikosch, *Modelling Extremal Events for Insurance and Finance*, Springer, Berlin 1997.

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

[7] W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. 2, 1st edition, Wiley, New York 1966.

---

[5] The results are very robust against taking other quantiles and other spread measure such as variation; see [14], Chapter 2, for an explanation of the variation.

[8] S. Kwapień and W. Woyczyński, *Random Series and Stochastic Integrals: Single and Multiple*, Birkhäuser, Boston 1992.
[9] R. Laha and V. Rohatgi, *Probability Theory*, Wiley, New York 1979.
[10] B. Mandelbrot, *The variation of certain speculative prices*, Journal of Business 26 (1963), pp. 394–419.
[11] S. Rachev, J.-R. Kim and S. Mittnik, *Stable Paretian econometrics, Parts I and II*, Math. Sci. 24 (1999), pp. 24–25 and pp. 113–127.
[12] S. Rachev and S. Mittnik, *Stable Paretian Models in Finance*, Wiley, 2000.
[13] S. Resnick, *Extreme Values, Regular Variation and Point Processes*, Springer, New York 1987.
[14] G. Samorodnitsky and M. Taqqu, *Stable Non-Gaussian Random Processes*, Chapman and Hall, New York 1994.
[15] V. Zolotarev, *One-dimensional Stable Distributions*, Transl. Math. Monogr., Vol. 65, 1986, American Mathematical Society. Translation from the original 1983 Russian edition.

Gennady Samorodnisky
School of Operations Research
and Industrial Engineering
and
Cornell University
Ithaca, NY 14853, USA
E-mail: gennady@orie.cornell.edu

Svetlozar T. Rachev
Institut für Statistik
und Mathematische Wirtschaftstheorie
Universität Karlsruhe
Kollegium am Schloss
D-76128 Karlsruhe, Germany
E-mail: rachev@pstat.ucsb.edu

Jeong-Ryeol Kurz-Kim
Deutsche Bundesbank
Wilhelm-Epstein-Strasse 14, D-60431
Frankfurt am Main, Germany
E-mail: jeong-ryeol.kurz-kim@bundesbank.de

Stoyan V. Stoyanov
FinAnalytica Inc.
440-F Camino del Remedio Santa Barbara
CA 93110, USA
E-mail: stoyan.stoyanov@finanalytica.com

Received on 20.7.2006;
revised version on 31.3.2007