Statistical Inference in Regression with Heavy-tailed Integrated Variables

Stefan Mittnik

Institute of Statistics and Econometrics

University of Kiel

Vygantas Paulauskas

Department of Mathematics

Vilnius University, Lithuania

and

Institut of Mathematics and Informatics

Vilnius, Lithuania

Svetlozar T. Rachev

Department of Statistics and Applied Probability

University of California, Santa Barbara

1. Introduction

The concept of cointegration was introduced in the seminal work of Granger (1981) and further developed by Engle and Granger (1987),see also Engle and Yoo (1987), Stock and Watson (1988), Park and Phillips (1988), Phillips (1991), Park (1992), and Johansen (1988, 1991). Studies in empirical macroeconomics typically involve non-stationary, integrated and cointegrated variables, such as prices, consumption, money demand, exchange rates, etc. A basic cointegration model for two macroeconomic variables, x_t and y_t, can be written as

$$y_t = \mu + \beta x_t + u_t, \tag{$*$}$$

where it is assumed that

(i) x_t is an integrated process,

$$x_t = x_{t-1} + e_t, \tag{$**$}$$

and

(ii) the disturbances (u_t, e_t) form a bivariate white-noise stationary process.

The classical way of removing the stochastic trend relies on differencing procedures. However, because economic variables are typically cointegrated, the differencing of the data is counter-productive, since it obscures the long term relationship between y_t and x_t.

There are two well-developed methods to test for cointegration: (1) the Engle-Granger-Phillips approach (see Engle and Granger (1987), Park and Phillips (1988), Phillips (1991)) amounts to testing for unit roots, for which the Dickey–Fuller and Durbin–Watson statistics can be employed; (2)the Johansen (1988) approach is based on a vector autoregressive representation of the time-series treating all variables as endogenous (see, for example, Watson (1997)). For a detailed introduction to the theory of cointegration, we refer to the reviews in Banerjee et al. (1993), Hargreaves (1994), Bhaskara Rao (1994), Hatanaka (1996) and Johansen (1996).

In the literature on cointegration, it is generally assumed that the disturbances (u_t, e_t) are in the domain of attraction of the Gaussian distribution. However, numerous empirical studies contradict the Gaussian assumption. Heavy-tailed and asymmetrically distributed samples are not infrequently observed in empirical economic time series, and these empirical facts cannot be explained with the usual Gaussian models. P.C.B. Phillips (see Phillips (1990), Phillips and Loretan (1991), Loretan and Phillips (1994)) addressed the issue of heavy-tailedness and asymmetry in econometric time-series in a rigorous fashion by introducing stable non-Gaussian (Paretian) variables for modeling the innovation processes in econometric and time series models (see also Chan and Tran (1989), Phillips (1995), Caner (1996, 1998), Kim, Mittnik and Rachev (1996), Mittnik, Rachev and Paolella (1997), Rachev, Kim and Mittnik (1997), Paulauskas and Rachev (1998), Mittnik and Rachev (1999)).

In this paper we extend the stable Paretian model in econometrics, developing the asymptotic theory for the cointegration model (*), (**) under the assumption that the bivariate innovation process (u_t, e_t) has heavy-tailed marginal distributions, specifically, we assume that (u_t, e_t) are in the normal domain of attraction of a bivariate infinitely divisible vector with stable components, having possibly different indexes of stability (i.e., different degrees of heavy-tailedness).

The paper is organized as follows. In Section 2, after a more detailed description of the innovation process (u_t, e_t), $t \geq 1$, we state and prove our main result (Theorem 1). We describe the limiting behavior of the joint 4-dimensional distribution of the estimators for β, μ, and the corresponding t-statistics t_β and t_μ. This general result extends some of the results of Park and Phillips (1988) which were derived under the assumption of the finite variance innovation process. In particular, Theorem 1 provides limiting expressions for all the statistics which are involved in the cointegration model driven by heavy-tailed dependent disturbances with different indexes of stability. In Section 3, we present numerical simulation results of the pre-limiting and limiting distributions of the test statistics derived in Section 2.

2. Statistical Inference with Heavy-tailed Variables

Consider the regression model

$$y_i = \mu + \beta x_i + u_i, \qquad i = 1, ..., n, \tag{1}$$

where sequence (x_i) is generated by a random-walk

$$x_i = x_{i-1} + e_i, \qquad i \geq 1. \tag{2}$$

The unknown parameters μ and β are to be estimated. We assume that the sequence of two-dimensional random variables (u_i, e_i), $i \geq 1$, is a sequence of i.i.d. random vectors in the domain of normal attraction (DAN) of some two-dimensional $\bar{\alpha}$-stable random vector, where $\bar{\alpha} = (\alpha_1, \alpha_2)$, $0 < \alpha_j \leq 2$, $j = 1, 2$. This means that the sums $\sum_{i=1}^{n}(u_i, e_i)$, properly normalized by diagonal matrices, converge in distribution to an infinitely divisible vector with each coordinate being stable. Note that those stable coordinates may have different indexes of stability (see (6) below). (We refer to Resnick and Greenwood (1979) for a description of the necessary and sufficient conditions characterizing DAN for $\bar{\alpha}$–stable random pairs and to Feller (1996) for a detailed analysis of multidimensional infinitely divisible laws). Furthermore, we assume that if the first moments of u_i and e_i exist, then $E(u_i) = E(e_i) = 0$. In Phillips and Durlauf (1986) and Park and Phillips (1988), the authors examined the model when the innovations (u_i, e_i), $i \geq 1$ are assumed to be normally distributed weakly dependent random vectors.

Denote

$$S_n = (S_{n1}, S_{n2}), \quad S_{n1} = n^{-1/\alpha_1} \sum_{i=1}^{n} u_i, \quad S_{n2} = n^{-1/\alpha_2} \sum_{i=1}^{n} e_i,$$

$$Z_n(t) = \big(Z_{n1}(t), Z_{n2}(t)\big), \quad Z_{nj}(t) = S_{[nt],j} \quad 0 \leq t \leq 1, \tag{3}$$

where $[a]$ stands for the integer part of a. Let $D = D[0,1]$ be the Skorokhod space of càdlàg functions defined on $[0,1]$ and equipped with the Skorokhod metric, setting D as a Polish space (see, for example, Billingsley (1968)). By $D_k = D\big([0,1], R^k\big)$ we denote the corresponding Skorokhod space of $\mathbf{R}^k$-valued càdlàg functions; and $D^k = D \times \cdots \times D$ will denote usual product of k topological spaces with the product topology.

In what follows, we use the following notation for norms in $\mathbf{R}^2$: $\|x\| = \big(x_1^2 + x_2^2\big)^{1/2}$ and $\|x\|_\infty = \max_{i=1,2} |x_i|$.

Next, let $\big\{\xi(t) = \big(\xi_1(t), \xi_2(t)\big), t \geq 0\big\}$ be a *Lévy process* with values in $\mathbf{R}^2$, i.e., a stochastically continuous bivariate process with independent and strictly stationary increments. Then, it is well-known (see, for example, Protter (1990) or Gikhman and Skorokhod

(1969)) that there exist a vector $a \in \mathbf{R}^2$, a symmetric non-negative defined matrix Γ, and a measure ν on $\mathbf{R}^2$ satisfying

$$\nu\big(\{0\}\big) = 0, \qquad \int_{R^2} \|x\|^2 \big(1 + \|x\|^2\big)^{-1} \nu(dx) < \infty,$$

such that for any $z \in \mathbf{R}^2$ the characteristic function of $\xi(t)$ has the following form:

$$E \exp\big\{i\big(z, \xi(t)\big)\big\} = \exp\big\{th(z)\big\}, \tag{4}$$

where (z, x) for $z, x \in \mathbf{R}^2$ denotes usual scalar product and

$$h(z) = i(z,a) - \frac{1}{2}(\Gamma z, z) + \int_{\|x\| \le 1} \big(e^{i(x,z)} - 1 - (x,z)\big)\nu(dx) + \int_{\|x\|>1} \big(e^{i(x,z)} - 1\big)\nu(dx). \tag{5}$$

In (5) ν is the so-called Lévy measure while the matrix Γ defines the Gaussian part of the distribution of ξ.

We shall start our analysis of (1) and (2) with some auxiliary results; we shall investigate the following limiting assertions: as $n \to \infty$,

$$S_n \xrightarrow{d} \xi(1) \qquad \text{in} \qquad \mathbf{R}^2 \tag{6}$$

$$Z_n(\cdot) \xrightarrow{d} \xi(\cdot) \qquad \text{in} \qquad D^2 \tag{7}$$

$$Z_n(\cdot) \xrightarrow{d} \xi(\cdot) \qquad \text{in} \qquad D_2, \tag{8}$$

where $\xrightarrow{d}$ stands for the weak convergence in the corresponding space. Note that the convergence in (6) is equivalent to the domain-of-attraction assumptions we made for innovation sequence (u_i, e_i), $i \ge 1$.

We exclude the Gaussian case, $\alpha_1 = \alpha_2 = 2$, in the following proposition, since this case is well-studied. In what follows, $\mathcal{B}(X)$ stands for the class of Borel sets of a metric space X.

Proposition 1. *(i) Case* $0 < \alpha_i < 2$, $i = 1, 2$. *Suppose that in* (5), $a = 0$ *and* $\Gamma = 0$. *Then,* (6)–(8) *are equivalent and each of them is equivalent to the existence of the following limit:*

$$\lim_{n\to\infty} nP\big\{\big(n^{-1/\alpha_1} u_j, n^{-1/\alpha_2} e_j\big) \in A\big\} = \nu(A), \tag{9}$$

for all $A \in \mathcal{B}(\mathbf{R}^2 \setminus \{0\})$ *such that* $\nu(\partial A) = 0$ *and* $\nu(A) < \infty$.

(ii) Case $0 < \alpha_1 < \alpha_2 = 2$: *Suppose that in* (5), $a = 0$, $\Gamma = \begin{bmatrix} 0 & 0 \\ 0 & \sigma^2 \end{bmatrix}$ *and let* ν *be a Lévy measure on the real line. Then,* (6)–(8) *are equivalent, and each of them is equivalent to the following two assertions: for any* $A \in (\mathcal{B}(\mathbf{R}) \setminus \{0\})$ *such that* $\nu(\partial A) = 0$, $\nu(A) < \infty$, *we have*

$$\lim_{n\to\infty} nP\{n^{-1/\alpha_1} u_1 \in A\} = \nu(A);$$

and for all $\varepsilon > 0$,

$$\lim_{n\to\infty} n\Big\{E\big[n^{-1}e_1^2\mathbf{1}(|e_1| < \varepsilon\sqrt{n})\big] - \big(E\big[n^{-1/2}e_1\mathbf{1}(|e_1| < \varepsilon\sqrt{n})\big]\big)^2\Big\} = \sigma^2.$$

Resnick and Greenwood (1979) showed the equivalence of (6) and (7). Paulauskas and Rachev (1998) stated (for the general d-dimensional case) that (6) implies (8). Note that, in general, (8) is a stronger relationship than (7).

The Lévy measure of the process ξ in case (i) can be described as follows. Define the mapping $\tau : \mathbf{R}^2 \to \mathbf{R}^2$, $\tau(x) = (\operatorname{sign} x_1|x_1|^{1/\alpha_1}, \operatorname{sign} x_2|x_2|^{1/\alpha_2})$, and let $\tilde\nu = \nu \circ \tau$. Then,

$$\tilde\nu\Big\{x : \|x\| > r, \frac{x}{\|x\|} \in B\Big\} = r^{-1}H(B),$$

where H is a finite measure on the unit sphere $S^2 = \{x \in \mathbf{R}^2 : \|x\| = 1\}$, and $B \in \mathcal{B}(S^2)$. If $\alpha_1 = \alpha_2 = \alpha < 2$, we obtain the well-known condition for a random vector to be in the α-stable DNA:

$$\nu\Big\{x : \|x\| > r, \frac{x}{\|x\|} \in B\Big\} = r^{-\alpha}H(B).$$

It is known (see Sharpe, 1969) that, in Case (ii) of Proposition 1, the first component ξ_1 of the Lévy process ξ and the second component,- the Brownian motion ξ_2,- are independent processes.

With these facts on the innovation process (u_t, e_t), $t \geq 1$, we have completed the preliminary stochastic analysis of the model (1), (2). Our next goal is to study the joint asymptotic distribution of the ordinary least squares (OLS) estimators of β and μ,

$$\widehat\beta_n := \frac{n\sum_{j=1}^n y_jx_j - \sum_{i=1}^n x_i \sum_{j=1}^n y_j}{n\sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} \quad \text{and} \quad \widehat\mu_n := \frac{\sum_{j=1}^n x_j^2 \sum_{i=1}^n y_i - \sum_{j=1}^n x_j \sum_{i=1}^n y_ix_i}{n\sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}, \tag{10}$$

and the corresponding t-statistics

$$t_{\widehat{\beta}} := \frac{\widehat{\beta}_n - \beta}{s_{\widehat{\beta}}} \qquad t_{\widehat{\mu}} := \frac{\widehat{\mu}_n - \mu}{s_{\widehat{\mu}}}. \tag{11}$$

In (11) we have set

$$s^2_{\widehat{\beta}} := \frac{\widehat{\sigma}^2_u}{\sum_{i=1}^n (x_i^2 - \bar{x}^2)}, \qquad s^2_{\widehat{\mu}} := \frac{\widehat{\sigma}^2_u n^{-1} \sum_{i=1}^n x_i^2}{\sum_{i=1}^n (x_i^2 - \bar{x}^2)},$$

and, $\bar{x} = n^{-1} \sum_{i=1}^n x_i$, $\widehat{\sigma}^2_u := n^{-1} \sum_{j=1}^n \widehat{u}_j^2$, $\widehat{u}_j := y_j - \widehat{\mu}_n - \widehat{\beta}_n x_j$.

Next, we introduce some notations related to the limiting distributions of the statistics defined in (10) and (11). Recall that $\xi(t) = \big(\xi_1(t), \xi_2(t)\big)$ is the limiting process in (8) and thus, $\xi_i(t)$, $i = 1, 2$, are α_i-stable Lévy processes, possibly dependent, if $\alpha_i < 2$. If $\alpha_1 = 2$, we assume that $Eu_1^2 = \sigma_1^2$ and then indeed ξ_1 is a Brownian motion. Similarly, if $\alpha_2 = 2$, then $Ee_1^2 = \sigma_2^2$ and ξ_2 is a Brownian motion. Further, by

$$\int_0^t \xi_i^-(s) d\xi_j(s), \qquad i, j = 1, 2,$$

we denote an *Itô* stochastic integral. (For a detailed treatment of the theory of stochastic integration we refer to the monographs Protter (1990), Elliott (1982) and Kopp (1984)). For $x \in D[0,1]$ and $0 < s \le 1$, $x^-(s)$ denotes, as usual, the left limit, $\lim_{u \uparrow s}(u)$. To simplify notation, we suppress the superscript in the stochastic integral and simply write $\int_0^t \xi_i(s)\, d\xi_j(s)$ or $\int_0^t \xi_i\, d\xi_j$, when there will be no ambiguity. Define next the so-called "square brackets" process (see, for example, Kopp (1984), p. 160):

$$\big[\xi_i, \xi_j\big]_t := \xi_i(t)\xi_j(t) - \int_0^t \xi_j\, d\xi_i - \int_0^t \xi_i\, d\xi_j, \qquad \big[\xi_i\big]_t := \big[\xi_i, \xi_i\big]_t.$$

Set $\gamma_1 := 1 - \frac{1}{\alpha_1} + \frac{1}{\alpha_2}$ and $\gamma_2 := 1 - \frac{1}{\alpha_1}$. Our first theorem deals with the asymptotic behavior of the joint distribution of the 4-dimensional vector

$$\big(n^{\gamma_1}(\widehat{\beta}_n - \beta), n^{\gamma_2}(\widehat{\mu}_n - \mu), t_{\widehat{\beta}}, t_{\widehat{\mu}}\big), \quad as \quad n \to \infty. \tag{12}$$

It turns out that the weak limit of the above sequence can be expressed as a rather complicated functional of the process ξ. To make the formulation of the result more concise, we define the following random quantities:

$$\begin{aligned}
Y_1 &:= \xi_1(1)\xi_2(1) - \int_0^1 \xi_1(u)\, d\xi_2(u), \qquad Y_2 := \xi_1(1),\\
Y_3 &:= \int_0^1 \xi_2(u)\, du, \qquad Y_4 := \int_0^1 \xi_2^2(u)\, du,\\
Y_5 &:= \left[\xi_1\right]_1 := \xi_1^2(1) - 2\int_0^1 \xi_1(u)\, d\xi_1(u), \quad \text{and} \quad Y_6 := Y_4 - Y_3^2.
\end{aligned}$$

Theorem 1. *Suppose that* $1 < \alpha_1 < 2$, *or* $\alpha_2 < \alpha_1(1-\alpha_1)^{-1}$, *if* $\alpha_1 \leq 1$. *Then, as* $n \to \infty$,

$$\left(n^{\gamma_1}(\widehat{\beta}_n - \beta), n^{\gamma_2}(\widehat{\mu}_n - \mu), t_{\widehat{\beta}}, t_{\widehat{\mu}}\right) \xrightarrow{d} (V_1, V_2, V_3, V_4), \tag{13}$$

where

$$V_1 = \frac{Y_1 - Y_2 Y_3}{Y_6}, \qquad V_2 = \frac{(Y_2 Y_4 - Y_1 Y_3)}{Y_6}, \tag{14}$$

$$V_3 = \frac{Y_1 - Y_2 Y_3}{\sqrt{Y_5 Y_6}}, \qquad V_4 = \frac{(Y_2 Y_4 - Y_1 Y_3)}{\sqrt{Y_4 Y_5 Y_6}}. \tag{15}$$

If $\alpha_1 = 2$, $0 < \alpha_2 < 2$, *then* V_3 *and* V_4 *in* (13) *admit the following representations:*

$$V_3 = \frac{\tilde{Y}_1 - \tilde{Y}_2 Y_3}{\sqrt{Y_6}} \quad \text{and} \quad V_4 = \frac{(\tilde{Y}_2 Y_4 - \tilde{Y}_1 Y_3)}{\sqrt{Y_4 Y_6}},$$

where $\tilde{Y}_1 := W(1)\xi_2(1) - \int_0^1 W(u)\, d\xi_2(u)$, $\tilde{Y}_2 = W(1)$, *and* $\{W(t), t \geq 0\}$ *is a standard Wiener process (Brownian motion).*

Remark 1. Paulauskas and Rachev (1998) studied the multivariate model (1), (2), but without the intercept and derived the limit distribution for OLS estimator $\widehat{\beta}_n$. Combining their result with our Theorem 1, it would not be difficult to state the multivariate version of Theorem 1.

Remark 2. As a consequence of Theorem 1 we can derive the marginal limiting relations for $\widehat{\beta}_n - \beta$, $\widehat{\mu}_n - \mu$, $t_{\widehat{\beta}}$ and $t_{\widehat{\mu}}$. In particular, the confidence intervals for β and μ can be constructed. Also as a corollary of Theorem 1, we can study the asymptotic joint distribution of the bivariate statistics $\left(n^{\gamma_1}(\widehat{\beta}_n - \beta), n^{\gamma_2}(\widehat{\mu}_n - \mu)\right)$ and $(t_{\widehat{\beta}}, t_{\widehat{\mu}})$. Note that the limiting pairs (V_1, V_2) and (V_3, V_4) have dependent components.

Remark 3. Some comments are appropriate on the restriction for the multiindex $\bar{\alpha}$ ensuring the consistency of estimators under consideration. The restriction $\alpha_1 > 1$ is rather natural, since if the "noise" in (1) has no finite mean then it is impossible by OLS to recover μ. It is more difficult to interpret the second relation which says that even in the case $0 < \alpha_1 \le 1$, where there is no consistency for $\widehat{\mu}_n$, we still can get the consistent estimator for b if $\alpha_2 < \alpha_1(1-\alpha_1)^{-1}$. This restriction on α_1 and α_2 remains even if we consider the model without intercept μ (see Theorem 2 bellow).

Proof of Theorem 1. From the definitions of $\widehat{\beta}_n, \widehat{\mu}_n, t_{\widehat{\beta}}$ and $t_{\widehat{\mu}}$ (see (10) and (11)), we obtain the following representations:

$$\widehat{\beta}_n - \beta = \frac{Y_{1n}}{Y_{3n}}, \quad \widehat{\mu}_n - \mu = \frac{Y_{2n}}{Y_{3n}}, \quad t_{\widehat{\beta}} = \frac{\widehat{\beta}_n - \beta}{s_{\widehat{\beta}}} = \frac{Y_{1n}}{Y_{3n} s_{\widehat{\beta}}}, \tag{16}$$

$$t_{\widehat{\mu}} = \frac{\widehat{\mu}_n - \mu}{s_{\widehat{\mu}}} = \frac{Y_{2n}}{Y_{3n} s_{\widehat{\mu}}},$$

where

$$\begin{aligned} Y_{1n} &:= n\sum_{i=1}^n u_i x_i - \sum_{i=1}^n x_i \sum_{i=1}^n u_i, \\ Y_{2n} &:= \sum_{i=1}^n x_i^2 \sum_{i=1}^n u_i - \sum_{i=1}^n x_i \sum_{i=1}^n u_i x_i, \\ Y_{3n} &:= n\sum_{i=1}^n x_i^2 - \Big(\sum_{i=1}^n x_i\Big)^2. \end{aligned} \tag{17}$$

Our next step is to derive the right-order normalization coefficients for $\widehat{\beta}_n - \beta$ and $\widehat{\mu} - \mu$, and at the same time to express all sums entering the expressions in (17) as functionals of the processes Z_{n1} and Z_{n2}, see (3). Because $x_k = x_0 + \sum_{m=1}^k e_m$, the right normalization for sum $\sum_{k=1}^n u_k x_k$ is given by $n^{-1/\alpha_1 - 1/\alpha_2}$. In fact, we can write

$$\begin{aligned} &n^{-1/\alpha_1-1/\alpha_2} \sum_{k=1}^n u_k x_k \\ =&n^{-1/\alpha_1-1/\alpha_2}\Big(\Big(\sum_{k=1}^n u_k\Big)\Big(\sum_{m=1}^n e_m\Big) - \sum_{m=1}^n e_m \sum_{k=1}^{m-1} u_k\Big) + o_p(1) \\ =&Z_{n1}(1)Z_{n2}(1) - \int_0^1 Z_{n1}(t)\, dZ_{n2}(t) + o_p(1). \end{aligned} \tag{18}$$

In a similar fashion we obtain

$$n^{-1-1/\alpha_2}\sum_{k=1}^{n} x_k^2 = \int_0^1 Z_{n2}^2(t)\,dt + o_p(1), \tag{19}$$

$$n^{-1-2/\alpha_2}\sum_{i=1}^{n} x_i = \int_0^1 Z_{n2}(t)\,dt + o_p(1), \tag{20}$$

$$n^{-1/\alpha_1}\sum_{i=1}^{n} u_i = Z_{n1}(1). \tag{21}$$

Therefore, the right normalization factors for Y_{1n}, Y_{2n}, Y_{3n} are $n^{-1-1/\alpha_1-1/\alpha_2}$, $n^{-1-1/\alpha_1-2/\alpha_2}$, and n^{-2-2/α_2}, respectively. Further more, the right normalization for $\widehat{\beta}_n - \beta$ is n^{γ_1}, and for $\widehat{\mu}_n - \mu$ is n^{γ_2}.

Consider next $s_{\widehat{\beta}}$ and $s_{\widehat{\mu}}$. We write

$$\begin{aligned}
\widehat{\sigma}_u^2 &= n^{-1}\sum_{i=1}^{n}\big(u_i - (\widehat{\mu}_n - \mu) - (\widehat{\beta}_n - \beta)x_i\big)^2 \\
&= n^{-1}\sum_{i=1}^{n} u_i^2 + (\widehat{\mu}_n - \mu)^2 + n^{-1}(\widehat{\beta}_n - \beta)^2\sum_{i=1}^{n} x_i^2 \\
&\quad - 2n^{-1}(\widehat{\mu}_n - \mu)\sum_{i=1}^{n} u_i - 2n^{-1}(\widehat{\beta}_n - \beta)\sum_{i=1}^{n} u_i x_i \\
&\quad + 2n^{-1}(\widehat{\mu}_n - \mu)(\widehat{\beta}_n - \beta)\sum_{i=1}^{n} x_i.
\end{aligned}$$

Using the information about normalization for all terms involved in the expression of $\widehat{\sigma}_u^2$, one can verify that the leading term is $n^{-1}\sum_{i=1}^n u_i^2$, and that the right normalization for this term is n^{1-2/α_1}. Therefore

$$\begin{aligned}
n^{1-2/\alpha_1}\widehat{\sigma}_u^2 &= n^{-2/\alpha_1}\sum_{i=1}^{n} u_i^2 + o_p(1) \\
&= Z_{n1}^2(1) - 2\int_0^1 Z_{n1}(t)\,dZ_{n1}(t) + o_p(1).
\end{aligned} \tag{22}$$

Since $\sum_{i=1}^n (x_i^2 - \bar{x}^2) = n^{-1}Y_{3n}$, we see that the right normalization for $s_{\widehat{\beta}}^2$ is $n^{2\gamma_1}$, and for $s_{\widehat{\mu}}^2$ is $n^{2\gamma_2}$. Thus, we have

$$n^{\gamma_1}s_{\widehat{\beta}} = \frac{n^{\gamma_1}\widehat{\sigma}_u}{(n_{-1}Y_{3n})^{1/2}} = \frac{\Big(n^{-2/\alpha_1}\sum_{i=1}^{n} u_i^2\Big)^{1/2}}{\big(n^{-2-2/\alpha_2}Y_{3n}\big)^{1/2}} + o_p(1),$$

$$
\begin{aligned}
n^{\gamma_2} s_{\widehat{\mu}} &= \frac{n^{\gamma_2} \widehat{\sigma}_u \Big(n^{-1} \sum_{i=1}^{n} x_i^2\Big)^{1/2}}{(n_{-1} Y_{3n})^{1/2}} \\
&= \frac{\Big(n^{-2/\alpha_1} \sum_{i=1}^{n} u_i^2\Big)^{1/2} \Big(n^{-1-2/\alpha_2} \sum_{i=1}^{n} x_i^2\Big)^{1/2}}{\big(n^{-2-2/\alpha_2} Y_{3n}\big)^{1/2}} + o_p(1). \qquad (23)
\end{aligned}
$$

These expressions show that $t_{\widehat{\beta}}$ and $t_{\widehat{\mu}}$ are properly normalized.

Combining (16)–(23), we have

$$
\big(n^{\gamma_1}(\widehat{\beta}_n - \beta), n^{\gamma_2}(\widehat{\mu}_n - \mu), t_{\widehat{\beta}}, t_{\widehat{\mu}}\big) = \big(V_{1n}, V_{2n}, V_{3n}, V_{4n}\big) + o_p(1), \qquad (24)
$$

where

$$
\begin{aligned}
V_{1n} &:= \frac{A_{1n} - A_{2n} A_{3n}}{A_{6n}}, \quad V_{2n} := \frac{A_{2n} A_{4n} - A_{1n} A_{3n}}{A_{6n}}, \\
V_{3n} &:= \frac{A_{1n} - A_{2n} A_{3n}}{\sqrt{A_{5n} A_{6n}}}, \quad V_{4n} := \frac{A_{2n} A_{4n} - A_{1n} A_{3n}}{\sqrt{A_{4n} A_{5n} A_{6n}}},
\end{aligned} \qquad (25)
$$

and

$$
\begin{aligned}
A_{1n} &:= Z_{n1}(1) Z_{n2}(1) - \int_0^1 Z_{n1}(u)\, dZ_{n2}(u), \\
A_{2n} &:= Z_{n1}(1), \quad A_{3n} = \int_0^1 Z_{n2}(u)\, d(u), \\
A_{4n} &:= \int_0^1 Z_{n2}^2(u)\, d(u), \\
A_{5n} &:= Z_{n1}^2(1) - 2 \int_0^1 Z_{n1}(u)\, dZ_{n1}(u), \text{ and} \\
A_{6n} &:= A_{4n} - A_{3n}^2.
\end{aligned}
$$

Although we expressed vector (12) as a function of Z_n plus a negligible part, this function is not continuous, due to the presence of stochastic integrals, and thus, we cannot immediately apply the continuous mapping theorem. The essential ingredient in the proof is the following proposition of Paulauskas and Rachev (1998).

Proposition 2. *Suppose that the sequence* $Z_n(t) = \big(Z_{n1}(t), Z_{n2}(t)\big)$, $0 \leq t \leq 1$, *is defined by* (3) *and that* (8) *holds. Then, as* $n \to \infty$,

$$
\begin{aligned}
&\Big(Z_n(t), \int_0^t Z_{n1}(s)\, dZ_{n1}(s), \int_0^t Z_{n1}(s)\, dZ_{n2}(s)\Big) \\
&\xrightarrow{d} \Big(\xi(t), \int_0^t \xi_1(s)\, d\xi_1(s), \int_0^t \xi_1(s)\, d\xi_2(s)\Big) \text{ in } D_4.
\end{aligned} \qquad (26)
$$

The proof of this proposition is based on results of Kurtz and Protter (1991) and Jakubowski et al. (1989) on convergence of stochastic integrals for semimartingales.

Having shown the relations (24) and (26), we can complete the proof of Theorem 1 by using the well-known fact that if $X_n \xrightarrow{d} X_0$ and $Y_n \xrightarrow{P} 0$, then $X_n + Y_n \xrightarrow{d} X_0$, and by applying the continuous mapping theorem. To this end let us define the mapping $\mathbf{f} : D_4 \to \mathbf{R}^4$, $\mathbf{f} = (f_1, f_2, f_3, f_4)$ with coordinates $f_1(\mathbf{x}) = g_1(\mathbf{x})\big(g_2(\mathbf{x})\big)^{-1}$, $f_2(\mathbf{x}) = g_3(\mathbf{x})\big(g_2(\mathbf{x})\big)^{-1}$, $f_3(\mathbf{x}) = g_1(\mathbf{x})\big(g_4(\mathbf{x})\big)^{-1}$, and $f_4(\mathbf{x}) = g_3(\mathbf{x})\big(g_5(\mathbf{x})\big)^{-1}$. Here, $\mathbf{x} = (x_1, x_2, x_3, x_4) \in D_4$ and functions $g_i \colon D_4 \to \mathbf{R}$, $i = 1, 2, 3, 4$ are defined as follows:

$$
\begin{aligned}
q_1(\mathbf{x}) :=&\ x_1(1)x_2(1) - x_4(1) - x_2(1)\int_0^1 x_2(u)\,du,\\
q_2(\mathbf{x}) :=&\ \int_0^1 x_2^2(u)\,du - \Big(\int_0^1 x_2(u)\,du\Big)^2,\\
q_3(\mathbf{x}) :=&\ x_2(1)\int_0^1 x_2^2(u)\,du - \Big(\int_0^1 x_2(u)\,du\Big)\big(x_1(1)x_2(1) - x_4(1)\big),\\
q_4(\mathbf{x}) :=&\ \big((x_1^2(1) - 2x_3(1)g_2(x)\big)^{1/2}, \quad \text{and}\\
q_5(\mathbf{x}) :=&\ q_4(x)\Big(\int_0^1 x_2^2(u)\,du\Big)^{1/2}.
\end{aligned}
$$

Let M_f denote the set of points of discontinuity of the mapping $\mathbf{f}$ and let m stand for the distribution of the limiting right-hand side vector in (26). Then we have

$$
\begin{aligned}
m(M_f) :=&\ m\Big(\Big\{\mathbf{x} \in D_4 \colon g_2(\mathbf{x}) = 0,\ \text{or}\ x_1^2(1) - 2x_3(1) = 0,\\
&\qquad \text{or} \int_0^1 x_2(u)\,du = 0\Big\}\Big)\\
\le&\ P\big\{\xi_2(t) = 0\big\} + P\big\{\big[\xi_1\big]_1 = 0\big\} = 0.
\end{aligned}
$$

Applying the continuous mapping theorem (see Billingsley, 1968) with the function $\mathbf{f}$ and making use of (26), we prove (13).

It remains to consider the case $\alpha_1 = 2$. In this case the limiting behavior of sum $\sum_{t=1}^n u_t^2$ is different from the case $\alpha_1 < 2$. We assume that $Eu_1^2 = \sigma^2 < \infty$, therefore, by the strong law of large numbers,

$$
n^{-1}\sum_{t=1}^n u_t^2 \to \sigma^2 \quad \text{a.s.} \tag{27}
$$

Now in (25), we replace the quantities V_{3n} and V_{4n} by

$$\tilde{V}_{3n} = \frac{A_{1n} - A_{2n}A_{3n}}{\sqrt{A_{1n}}}, \quad \tilde{V}_{4n} = \frac{A_{2n}A_{4n} - A_{1n}A_{3n}}{\sqrt{A_{4n}A_{6n}}}$$

and replace, accordingly, the mapping $\mathbf{f}$ into $\tilde{\mathbf{f}}$. Then, as before, we apply $\tilde{\mathbf{f}}$ to (26). The last step in the proof is to recall the fact that if $X_n \xrightarrow{d} X_0$ and $Y_n \xrightarrow{p} a$ (in our case this will be $A_{n5} = n^{-1} \sum_{i=1}^{n} u_i^2 \to \sigma^2$ a.s. by (27)), then $(X_n, Y_n) \xrightarrow{d} (X_0, a)$. Applying the continuous mapping theorem once more , now with the map h: $R^5 \to R^4$, $h(x_1, ..., x_5) = \left(x_1, x_2, x_3 x_5^{-1}, x_4 x_5^{-1}\right)$, we prove the theorem.

Consider next the model (1), (2) with $\mu = 0$. Then the OLS estimator for β is given by

$$\tilde{\beta}_n := \frac{\sum_{j=1}^{n} y_j x_j}{\sum_{j=1}^{n} x_j^2} \tag{28}$$

and the corresponding t-statistic is

$$t_{\tilde{\beta}} := \frac{\tilde{\beta}_n - \beta}{s_{\tilde{\beta}}}, \tag{29}$$

where $s_{\tilde{\beta}}^2 := \widehat{\sigma}_u^2 \Big/ \sum_{i=1}^{n} x_i^2$, and $\widehat{\sigma}_u^2 = n^{-1} \sum_{j=1}^{n} \widehat{u}_j^2$, $\widehat{u}_j = y_j - \tilde{\beta}_n x_j$. Using the same arguments as in the proof of Theorem 1 we obtain the following asymptotic results.

Theorem 2. *Suppose that* $\mu = 0$ *in the cointegration model* (1), (2), *with* $\alpha_2 < \alpha_1(1 - \alpha_1)^{-1}$, *if* $\alpha_1 \leq 1$. *Then, as* $n \to \infty$,

$$n^{1 - \frac{1}{\alpha_1} + \frac{1}{\alpha_2}} \left(\tilde{\beta}_n - \beta, t_{\tilde{\beta}}\right) \xrightarrow{d} (Z, V), \tag{30}$$

where $\tilde{\beta}_n$ *and* $t_{\tilde{\beta}}$ *are given by* (28) *and* (29), *and*

$$Z := \frac{\xi_1(1)\xi_2(1) - \int_0^1 \xi_1(s)\, d\xi_2(s)}{\int_0^1 \xi_2^2(s)\, ds} \tag{31}$$

$$V = \frac{\xi_1(1)\xi_2(1) - \int_0^1 \xi_1(s)\, d\xi_2(s)}{\left(\left[\xi_1\right]_1 \int_0^1 \xi_2^2(s)\, ds\right)^{1/2}}. \tag{32}$$

Here $\xi = \xi = (\xi_1, \xi_2)$ is the limiting process in (6). If $\alpha_1 = 2$ and $0 < \alpha_2 < 2$, then the limiting relationship (30) still holds with Z given by (31) and $\xi_1 = W$,

$$V = \frac{W(1)\xi_2(1) - \int_0^1 \xi_2(s)\,dW(s)}{\left(\int_0^1 \xi_2^2(s)\,ds\right)^{1/2}}, \tag{33}$$

where $W(t)$, $t \geq 0$ is a Brownian motion, and W and ξ_2 are independent.

3. Simulation of limit distribution

Because the limiting vector $(V_1, \ldots, V_4)$ in (13) has a rather complicated structure and there is no close-form analytical expression of its distribution, we use $(V_{1n}, \ldots, V_{4n})$ defined in (25) with sufficiently large n to simulate and analyze the distribution of $(V_1, \ldots, V_4)$. To do so, we generated values (u_{ji}, e_{ji}), $j = 1, 2, ..., n$, $i = 1, ..., m$, of a vector (u, e), where u and e are two independent stable random variables with exponents α_1 and α_2 and skewness parameters β_1 and β_2, respectively. Then for each $i = 1, 2, ..., m$ we evaluated the quantities V_{jn}, $j = 1, 2, 3, 4$ making use of (25). Note, that since Z_{nj} are step-functions, the quantities $A_{j,n}$, $j = 1, 2, ...6$ are expressed by means of various sums. We consider the empirical distribution of m 4-dimensional values as an approximation of the limiting distribution in (13). Starting with values $n = m = 500$, we then increased the sample sizes to $n = 1200$, $m = 800$ (and ,in some cases, to $n = 2000$). Sufficiently large values for n guarantee stability (stability here means that the prelimiting distribution is close to the limiting one) for the pair (V_3, V_4). Unfortunately, even $n = 2000$ does not show stability of the empirical distribution of the pair (V_1, V_2).

The simulation results are presented in a number of tables and graphs. We chose 9 parameters settings for α_1, α_2, β_1, and β_2. In addition, we generated the distribution of $(V_1, \ldots, V_4)$ for the Gaussian case, ($\alpha_1 = \alpha_2 = 2$), which we denote by number 10 in tables and graphs. For example, the notation "$V3$, case 10" stands for the coordinate V_3 in the case of Gaussian innovations, while "$V2$, case 3" denotes coordinate V_2 in the case of stable innovations with parameters given in case 3. The parameter values are given in Table 1.

cases / param	1	2	3	4	5	6	7	8	9	10
α_1	1.1	1.4	1.8	1.1	1.8	1.4	1.4	1.1	1.8	2
α_2	1.3	1.4	1.7	1.7	1.3	1.4	1.4	1.8	1.1	2
β_1	-0.50	0.0	0.25	-0.50	0.25	-0.5	0.5	0.9	-0.9	-
β_2	-0.25	0.0	0.50	-0.25	0.50	0.5	-0.5	-0.9	0.9	-
EV_3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
$VarV_3$	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
EV_4	-	0.0	0.0	-	0.0	0.4	-0.3	-	-0.1	0.0
$VarV_4$	-	1.0	1.0	-	1.0	1.3	1.1	-	1.0	1.0

Table 1

In Fig. 1 there are several examples of plotted histograms of marginal densities of random variables V_3 and V_4 in some cases. We have not included graphs of the marginal distributions of V_1 and V_2 since, as mentioned above, the stability achieved for the first pair of coordinates is not satisfactory. One can find more graphical material on simulation results (including graphics of two-dimensional densities of the pair (V_3, V_4)) in our technical report [30].

Based on simulation results , the following observations can be made. The marginal distributions of (V_1, V_2) are heavy-tailed, while the pair (V_3, V_4) has very light tails (this fact is due to self-normalizing effect) and is much more convenient for constructing confidence intervals. We conjecture that for the pair (V_3, V_4) lower-order moments exist, more over, $EV_j = 0$, $VarV_j = 1$, $j = 3, 4$ independently of values of parameters α_i and β_i, $i = 1, 2$. Although at present we are not able to prove this, simulation shows that: increasing n stabilizes the empirical mean and variance. Therefore, in Table 1 we provide values of EV_{jn} and $Var(V_{jn})$, $j = 3, 4$, rounded to one decimal point, as the theoretical mean and variance of V_3 and V_4. In the cases 1, 4, 8 bar sign is left instead of EV_4 and $VarV_4$, since there was no stability in calculations. It seems that instability in these cases is caused by the closeness of the parameter α_1 to the boundary value 1. We recall that for $\alpha_1 = 1$ there is no consistency of the estimator $\widehat{\mu}_n$, therefore in the case $\alpha_1 = 1,1$ convergence can be rather slow, especially if we take expected values or moments of higher order. Moreover, the centralized and normalized distributions $(Var(V_i))^{-1/2}(V_i - EV_i)$ $i = 3, 4$ are very similar to the standard normal distribution. In fact, the Chi-square criterion rejects the hypothesis about normality of these distributions with 90%-significance level only in three

cases for V_3 and in none for V_4. We have no theoretical explanation for this fact and intend to conduct further research on this. Therefore, to construct confidence intervals for the parameter β, we can use the probabilities $P\big\{|V_3| < x\big\}$ given in Table 2, and for the parameter μ we can use probabilities

$$P\Big\{\big|(Var(V_4))^{-1/2}(V_4 - EV_4)\big| < x\Big\}$$

given in Table 3 and using values EV_4 and $Var(V_4)$ from Table 1. Tables 2 and 3 support the conjecture that distributions of V_3 and V_4 are not sensitive to the changes of the parameters α's and β's, since all columns in these tables are rather similar and close to corresponding probabilities of standard normal law, which in both tables are given in the column with number N. That is, in this column, there are given probabilities $P\{|\eta| \le x\}$, where η is a standard normal random variable with mean zero and unit variance.

case x	1	2	3	4	5	6	7	8	9	10	N
1.5	0.8613	0.8588	0.8600	0.8738	0.8563	0.8588	0.8725	0.8588	0.8750	0.8563	0.8664
1.6	0.8888	0.8875	0.8788	0.8975	0.8838	0.8875	0.8988	0.8863	0.8913	0.8938	0.8904
1.7	0.9125	0.9125	0.9050	0.9250	0.9138	0.9175	0.9175	0.9100	0.9113	0.9113	0.9109
1.8	0.9363	0.9300	0.9263	0.9375	0.9363	0.9275	0.9438	0.9250	0.9225	0.9313	0.9281
1.9	0.9450	0.9463	0.9425	0.9500	0.9500	0.9475	0.9525	0.9425	0.9363	0.9475	0.9426
2.0	0.9563	0.9563	0.9625	0.9550	0.9638	0.9625	0.9650	0.9550	0.9575	0.9650	0.9545
2.1	0.9650	0.9663	0.9688	0.9675	0.9700	0.9638	0.9738	0.9650	0.9663	0.9738	0.9643
2.2	0.9700	0.9725	0.9738	0.9738	0.9788	0.9750	0.9800	0.9763	0.9725	0.9850	0.9722
2.3	0.9775	0.9763	0.9838	0.9775	0.9825	0.9838	0.9850	0.9788	0.9800	0.9875	0.9786
2.4	0.9838	0.9838	0.9900	0.9800	0.9863	0.9888	0.9888	0.9838	0.9863	0.9888	0.9836
2.5	0.9863	0.9863	0.9925	0.9825	0.9925	0.9938	0.9900	0.9888	0.9913	0.9913	0.9876
2.6	0.9875	0.9888	0.9950	0.9863	0.9950	0.9950	0.9913	0.9938	0.9938	0.9950	0.9907
2.7	0.9913	0.9913	0.9975	0.9950	0.9950	0.9963	0.9913	0.9963	0.9950	0.9988	0.9931
2.8	0.9938	0.9913	1.0000	0.9963	0.9963	0.9975	0.9925	0.9975	0.9975	1.0000	0.9949
2.9	0.9963	0.9950	1.0000	0.9963	0.9963	0.9975	0.9975	0.9975	0.9988	1.0000	0.9963
3.0	0.9975	0.9975	1.0000	0.9975	0.9988	0.9988	0.9988	0.9988	1.0000	1.0000	0.9973
3.1	0.9988	0.9988	1.0000	0.9988	0.9988	1.0000	0.9988	0.9988	1.0000	1.0000	0.9981
3.2	0.9988	0.9988	1.0000	1.0000	0.9988	1.0000	0.9988	0.9988	1.0000	1.0000	0.9986
3.3	0.9988	0.9988	1.0000	1.0000	1.0000	1.0000	1.0000	0.9988	1.0000	1.0000	0.9990
3.4	0.9988	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9988	1.0000	1.0000	0.9993
3.5	0.9988	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9988	1.0000	1.0000	0.9995

Table 2 $P\{|V_3| < x\}$

case x	1	2	3	4	5	6	7	8	9	10	N
1.5	0.8913	0.8650	0.8675	0.8988	0.8538	0.8738	0.8613	0.9063	0.8538	0.8663	0.8664
1.6	0.9063	0.8950	0.8963	0.9175	0.8850	0.9025	0.8938	0.9175	0.8850	0.8938	0.8904
1.7	0.9163	0.9125	0.9100	0.9363	0.9100	0.9238	0.9163	0.9238	0.9025	0.9088	0.9109
1.8	0.9313	0.9350	0.9263	0.9463	0.9238	0.9413	0.9325	0.9388	0.9250	0.9213	0.9281
1.9	0.9463	0.9475	0.9350	0.9550	0.9425	0.9525	0.9425	0.9500	0.9375	0.9325	0.9426
2.0	0.9588	0.9625	0.9500	0.9575	0.9675	0.9613	0.9525	0.9613	0.9513	0.9425	0.9545
2.1	0.9663	0.9700	0.9675	0.9650	0.9750	0.9738	0.9675	0.9663	0.9525	0.9563	0.9643
2.2	0.9700	0.9775	0.9738	0.9700	0.9838	0.9763	0.9738	0.9700	0.9663	0.9725	0.9722
2.3	0.9763	0.9838	0.9763	0.9700	0.9888	0.9800	0.9788	0.9713	0.9750	0.9775	0.9786
2.4	0.9763	0.9913	0.9863	0.9750	0.9913	0.9863	0.9825	0.9788	0.9825	0.9825	0.9836
2.5	0.9763	0.9938	0.9900	0.9750	0.9963	0.9900	0.9875	0.9800	0.9863	0.9888	0.9876
2.6	0.9850	0.9950	0.9913	0.9788	0.9975	0.9900	0.9900	0.9813	0.9925	0.9900	0.9907
2.7	0.9863	0.9950	0.9975	0.9800	0.9988	0.9925	0.9913	0.9850	0.9963	0.9950	0.9931
2.8	0.9900	0.9963	0.9988	0.9813	0.9988	0.9938	0.9938	0.9875	0.9975	0.9975	0.9949
2.9	0.9913	0.9988	0.9988	0.9838	0.9988	0.9938	0.9938	0.9888	0.9975	0.9988	0.9963
3.0	0.9938	0.9988	1.0000	0.9863	0.9988	0.9938	0.9950	0.9888	0.9988	0.9988	0.9973
3.1	0.9938	0.9988	1.0000	0.9875	0.9988	0.9963	0.9975	0.9900	0.9988	1.0000	0.9981
3.2	0.9938	0.9988	1.0000	0.9888	1.0000	0.9963	0.9988	0.9913	1.0000	1.0000	0.9986
3.3	0.9950	0.9988	1.0000	0.9888	1.0000	0.9975	0.9988	0.9913	1.0000	1.0000	0.9990
3.4	0.9950	1.0000	1.0000	0.9913	1.0000	1.0000	0.9988	0.9913	1.0000	1.0000	0.9993
3.5	0.9950	1.0000	1.0000	0.9925	1.0000	1.0000	0.9988	0.9925	1.0000	1.0000	0.9995

Table 3 $P\big\{(VarV_4)^{-1/2}|V_4 - EV_4| \le x\big\}$

4. Conclusion

We have extended Phillips' approach to econometric models with heavy-tailed innovations by developing asymptotic theory for cointegration models with innovations having infinitely divisible distributions. This allows us to consider models with innovations having any type of tail-behavior. Our main result provides the joint asymptotic distribution for all statistics involved in the cointegration model with drift and innovations with possibly different tail behavior. This is achieved by an extensive use of the modern theory for stochastic integration. We provide simulation studies for the limiting distributions. Based on our simulation results for marginal distributions of V_3 and V_4 , we conclude that one can construct satisfactory confidence intervals for the unknown parameters β and μ.

References

1. Banerjee A., Dolado J., Galbraith J. W. and Henry D. (1994) *Co-integration, Error Correction, and the Econometric Analysis of Non-Stationary Data,* Advanced Texts in Econometrics, Oxford University Press, Oxford.
2. Bhaskara Rao B. (edt.) (1994) *Cointegration for the Applied Economist*, Macmillan Press, London.
3. Billingsley P. (1968), *Convergence of Probability Measures*, New York: Wiley.
4. Caner M. (1996), *Tests for Cointegration With Infinite Variance Errors, Discussion Paper*, Department of Economics, Brown University.
5. Caner M. (1998), *Tests for Cointegration with infinite variance errors*, J.of Econometrics, **86**, 155-175.
6. Chan N. H. and Tran L.T. (1989), *On the first order autoregressive process with infinite variance*, Econometric Theory, **5**, 354–362.
7. Elliott R. J. (1982), *Stochastic Calculus and Applications*, Springer, New York.
8. Engle R. F. and Granger C. W. J. (1987), *Co-integration and error corrections: representation, estimation and testing*, Econometrica, **55**, 251–276.
9. Engle R. F. and Yoo S. B. (1987), *Forecasting and testing in Cointegrated Systems*, J. Econometrics, **35**, 143–159.
10. Feller W. (1971), *An Introduction to Probability Theory and Its Application*, **2**(2nd ed.), Wiley, New York.
11. Gikhman I. I. and Skorokhod A. V. (1969), *Introduction to the Theory of Random Processes*, V.B. Saunders Co., Philadelphia.
12. Granger C.W.J. (1981), *Some properties of time series data and their use in econometric model specification*, J. Econometrics, **16**, 121–130.
13. Hannan E. J. and Kanter M. (1977), *Autoregressive Processes With Infinite Variance*, J. of Applied Probability, **14**, 411–415.
14. Hargreaves C. P. (edt.) (1994), *Nonstationary Time Series Analysis and Cointegration, Advanced Texts in Econometrics*, Oxford University Press, Oxford.
15. Hatanaka M. (1996), *Time-Series-Based Econometrics. Unit Roots and Co-Integration*, Advanced Texts in Econometrics, Oxford University Press, Oxford.
16. Jakubowski A., Memin J., Pages G. (1989), *Convergence en loi des suites d'integrales stochastiques sur l'espace D^1 de Skorokhod*, Probab. Th. Rel. Fields, **81**, 111–137.
17. Johansen S. (1988), *Statistical analysis of cointegration vectors*, Journal of Economic Dynamics and Control, **12**, 231–254.
18. Johansen S. (1991), *Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models*, Econometrics, **59**(6), 1551–1580.
19. Johansen S. (1996), *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, Advanced Texts in Econometrics, Oxford University Press, Oxford.

20. Kim J.-R., Mittnik S. and Rachev S. T. (1996), *Detecting asymmetries in observed time series and unobserved disturbances*, Studies in Nonlinear Dynamics and Econometrics **1**, 131–138.
21. Kopp P. E. (1984), *Martingales and stochastic integrals*, Cambridge Univ. Press, Cambridge.
22. Kurtz T. G. and Protter P. (1991), *Weak limit theorems for stochastic integrals and stochastic differential equations*, Ann. Probability, **19**, 3, 1035–1070.
23. Lévy P. (1937), *Theorie de L'addition des Variables Aleatoires (2nd ed.)*, Gauthier–Villars, Paris.
24. Loretan, M. and Phillips P. C. B. (1994), *Testing the covariance stationarity of heavy-tailed time series*, Journal of Empirical Finance, **1**, 788–809.
25. Mittnik S. and Rachev S. T. (1999), *Asset and Option Pricing with Alternative Stable Models*, Wiley, forthcoming.
26. Mittnik S., Rachev S. T. and Paolella (1997), *Stable Paretian modeling in finance: some empirical and theoretical aspects*, A Practical Guide to Heavy Tails: Statistical Techniques for Analyzing Heavy Tailed Distributions, edt. R.Adler et al., Birkhauser, Boston, in press.
27. Park J.Y. and Phillips P. C. B. (1988), *Statistical inference in regressions with integrated processes: part I*, Econometric Theory, **4**, 468–497.
28. Park J.Y. (1992), *Canonical cointegrating regressions*,Econometrica, **60**, 11, 119-143.
29. Paulauskas V. and Rachev S. T. (1998), *Cointegrated processes with infinite variance innovations*, Ann. Appl. Probab. **8**, 3, 775-792.
30. Paulauskas V. and Rachev S. T. (1998), *Statistical Inference in Regression with Heavy-tailed Integrated Variables*, UCSB Technical report, **346**.
31. Phillips P. C. B. (1990), *Time series regression with a unit root and infinite-variance errors*, Econometric Theory, **6**, 44–62.
32. Phillips P. C. B. (1991), *Optimal inference in cointegrated systems,* , Econometrica , **59**, 283–306.
33. Phillips P. C. B. (1995), *Robust nonstationary regression* , Econometric Theory, **11**, 912-951.
34. Phillips P. C. B. (1996), *Econometric model determination*, Econometrica, **64**, 763–812.
35. Phillips P. C. B. and Loretan M. (1991), *On Durbin–Watson ratio under infinite-variance errors*, Journal of Econometrics, **47**, 85–114.
36. Phillips P. C. B. and Durlauf S. N. (1986), *Multiple time series regression with integrated processes*, Rev. Econom. Stud., **53**, 473–495.
37. Protter P. (1990), *Stochastic Integration and Differential Equations*, Springer, Berlin .

38. Rachev S. T., Kim J-R., Mittnik S. (1997), *Econometric modeling in the presence of heavy-tailed innovations: a survey of some recent advances*, Stochastic Models, **13**, 841–866.
39. Resnick S. and Greenwood P. (1979), *Bivariate Stable Characterization and Domains of Attraction*, J. Multivar. Anal., **9**, 206–221.
40. Sharpe M. (1969), *Operator-stable probability distributions on vector groups* , Trans. Amer. Math. Soc., **136**, 51–65.
41. Stock J. H. and Watson W. W. (1988), *Testing for common trends*, J. Amer. Statist. Assoc., **83**, 1097–1107.
42. Toda H. Y., Yamamoto T. *Statistical inference in vector autoregressions with possibly integrated processes*, J. Econometrics, **66**, 225–250.
43. Watson M. W. (1997), *Vector Autoregressions and Cointegration*, Handbook in Econometrics, Volume 4, Edited by R.F.Engle and D.L.McFadden, Chapter 47, Elsevier, to appear.