Econometric Theory, **25**, 2009, 748–763. Printed in the United States of America.
doi:10.1017/S0266466608090282

CENTRAL LIMIT THEOREMS FOR WEIGHTED SUMS OF LINEAR PROCESSES: L_P-APPROXIMABILITY VERSUS BROWNIAN MOTION

KAIRAT T. MYNBAEV
Kazakh-British Technical University

Standardized slowly varying regressors are shown to be L_p-approximable. This fact allows us to provide alternative proofs of asymptotic expansions of nonstochastic quantities and central limit results due to P.C.B. Phillips, under a less stringent assumption on linear processes. The recourse to stochastic calculus related to Brownian motion can be completely dispensed with.

1. INTRODUCTION

Regressions with asymptotically collinear regressors have surprisingly many applications, as the references in Phillips (2007) show. Using the theory of slowly varying (SV) functions, Phillips has developed a method to deal with such regressions. His method can be divided into three blocks.

I. In the basement of the construction there is a derivation of asymptotic expansions of various deterministic aggregates arising in regression analysis. One of the most important pieces of this block is the relation

$$\int_0^1 \left(\frac{L(rn)}{L(n)} - 1 \right)^k dr = (-1)^k k! \varepsilon^k(n)[1 + o(1)], \qquad n \to \infty, \tag{1}$$

where k is a natural number and L is an SV function with Karamata representation

$$L(x) = c \exp\left(\int_a^x \frac{\varepsilon(t)}{t} dt \right) \quad \text{for } x \geq a \tag{2}$$

for some $a > 0$, where $c > 0$ is a constant and ε is a continuous function such that $\varepsilon(x) \to 0$ as $x \to \infty$.

The paper has been significantly improved following comments and suggestions by three anonymous referees and an associate editor. I am especially grateful to one of the referees for a cryptic remark that my paper "opens the way to use the important results of Cremers and Kadelka (1986)." Address correspondence to Kairat T. Mynbaev, Zhandosov Str. 57, Apt. 49, 050035 Almaty, Kazakhstan; e-mail: kairat@fulbrightweb.org.

 © 2009 Cambridge University Press 0266-4666/09 $15.00

https://doi.org/10.1017/S0266466608090282 Published online by Cambridge University Press

II. Next, a series of central limit results is established. All of them treat convergence in distribution of weighted sums of linear processes

$$\sum_{t=1}^{n} w_{nt} u_t, \tag{3}$$

where $w_n = \{w_{nt} : t = 1, \ldots, n\}$ is a deterministic vector of weights and u_t is a linear process satisfying the condition

$$u_t = \sum_{j=0}^{\infty} c_j e_{t-j}, \qquad \sum_{j=0}^{\infty} j\,|c_j| < \infty, \qquad \sum_{j=0}^{\infty} c_j \neq 0, \tag{4}$$

$$e_t = iid(0, \sigma_e^2), \qquad \mu_{2p} = \mathrm{E}|e_t|^{2p} < \infty \quad \text{for some } p > 2.$$

III. For several types of regressions (linear and nonlinear, with one or more SV functions as regressors) estimation methods are developed. In all of them the regressors are asymptotically collinear, and handling asymptotically degenerate variance-covariance matrices is the main difficulty.

In this paper we concentrate on blocks I and II and relate them to the L_p-approximability notion suggested in Mynbaev (2001). It is shown that many technical results established by Phillips can be obtained in a shorter way and under less restrictive conditions, with one exception: the proof of (1) given by Phillips is incomplete, and we impose an additional regularity assumption on L. The largest improvement is achieved in central limit theorems (CLTs). To prove his CLTs, Phillips employs advanced properties of Brownian motion, such as uniform strong approximation of a partial sum process by Brownian motion and the local law of iterated logs for Brownian motion. All this heavy artillery becomes unnecessary and the resulting statements are more general if one uses the CLT from Mynbaev (2001). To explain our approach and describe the results, we need, first of all, the definition of L_p-approximability.

The basic idea is simple: to approximate sequences of deterministic vector weights $\{w_n : n = 1, 2, \ldots\}$ with functions of a continuous argument. For example, for a given sequence $\{w_n\}$, where w_n has n coordinates for each n, one can postulate existence of a continuous function W on $[0, 1]$ such that $\max_{t=1,\ldots,n} |w_{nt} - W(t/n)| \to 0$; see a similar assumption in Tanaka (1996, eqn. (5.141)). The continuity requirement can be a serious obstacle (in application to SV functions, too), and a way out is to use L_p spaces.

Let $p \in [1, \infty]$, $\|W\|_p = \left(\int_0^1 |W(x)|^p\,dx\right)^{1/p}$ if $p < \infty$ and $\|W\|_\infty = \operatorname{ess\,sup}_{x\in(0,1)} |W(x)|$. Denote L_p the space of measurable functions on $(0, 1)$ with a finite norm $\|W\|_p < \infty$. Let us partition the interval $[0, 1)$ into subintervals $i_t = [(t-1)/n, t/n)$, $t = 1, \ldots, n$. Because, in general, elements of L_p are not continuous, instead of values $W(t/n)$ we need to use averages over subintervals $n\int_{i_t} W(x)\,dx$. Those averages are scaled by $n^{-1/p}$ for convenience. In this way

https://doi.org/10.1017/S0266466608090282 Published online by Cambridge University Press

we obtain a *discretization operator* $\delta_{np} : L_p \to \mathbb{R}^n$ defined by

$$(\delta_{np} W)_t = n^{1-1/p} \int_{i_t} W(x)\,dx, \qquad t = 1, \ldots, n.$$

Let l_p be a discrete analogue of L_p with the norm $\|w\|_p = (\sum_t |w_t|^p)^{1/p}$, $p < \infty$, and $\|w\|_\infty = \max_t |w_t|$, $p = \infty$. Then we say that a sequence $\{w_n\}$, where $w_n \in \mathbb{R}^n$ for each n, is *L_p-approximable* if there exists a function $W \in L_p$ such that

$$\left\| w_n - \delta_{np} W \right\|_p \to 0, \qquad n \to \infty. \tag{5}$$

In this case we also say that $\{w_n\}$ is *L_p-close to W*. One should not confuse this notion with its namesake from Pötscher and Prucha (1991), which postulates approximation of stochastic processes with other, less complex, stochastic processes. The sequence $\left\{\delta_{np} W : n = 1, 2, \ldots\right\}$ is called *L_p-generated* (Moussatat, 1976).

Instead of the Phillips Assumption LP (see (4), we assume the following condition.

Assumption 1. $u_t = \sum_{j=-\infty}^{\infty} c_j e_{t-j}$, $\sum_{j=-\infty}^{\infty} \left|c_j\right| < \infty$, $\sum_{j=-\infty}^{\infty} c_j \neq 0$, with $e_t = iid(0, \sigma_e^2)$ and uniformly integrable e_t^2.

Now we can state Theorem 4.1 from Mynbaev (2001) in the one-dimensional case, which is sufficient for our purposes.

THEOREM 1. *Let u_t satisfy Assumption 1 and let the sequence of weights $\{w_n\}$ be L_2-close to $W \in L_2$. Denote $\sigma^2 = \left(\sigma_e \sum_{j=-\infty}^{\infty} c_j\right)^2$. Then*

$$\sum_{t=1}^{n} w_{nt} u_t \to_d N\left(0, \sigma^2 \int_0^1 W^2(u)\,du\right). \tag{6}$$

Remark 1. For $W \in C[0,1]$ (the space of continuous functions on $[0,1]$), according to Mynbaev (2001, Thm. 3.3(b)), w_{nt} here can be replaced by $(1/\sqrt{n})$ $W(\frac{t}{n})$. Even in this special case the resulting relation

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{n} W\left(\frac{t}{n}\right) u_t \to_d N\left(0, \sigma^2 \int_0^1 W^2(u)\,du\right)$$

is true under weaker conditions on the linear process than equation (9) of Phillips (2007).

Remark 2. A rather general CLT for sums (3) with one-sided long-memory processes $u_t = \sum_{j=0}^{\infty} c_j e_{t-j}$ has been established by Hannan (1979). However, in Hannan's result there is no link between sequences of weights and functions of a continuous argument, which seems to be important for applications. Likewise,

https://doi.org/10.1017/S0266466608090282 Published online by Cambridge University Press

linear CLTs that can be derived from nonlinear CLTs (de Jong and Davidson, 2000) do not have the flexibility associated with the L_p-approximability notion. See in Section 3 how the main result on L_p-approximability (Theorem 3) is applied in five different situations.

Remark 3. The summability condition $\sum_{j=0}^{\infty} |c_j| < \infty$ (and even weaker) by itself appeared earlier; see, for example, Anderson (1971, Thm. 7.7.8), Hannan (1970, Thm. 11, p. 221), Hall and Heyde (1980, Cor. 5.2), and Phillips and Solo (1992). However, the weights in these results are equal, $w_{nt} = n^{-1/2}$, $t = 1, \ldots, n$, and this takes us to the next remark.

Remark 4. Consider the convolution operator $T_n : \mathbb{R}^n \to l_p$ defined by

$$(T_n w)_j = \sum_{t=1}^{n} w_t c_{t-j}, \qquad j \in \mathbb{Z}, \qquad w \in \mathbb{R}^n.$$

Then with u_t specified in Assumption 1 the weighted sum (3) can be written

$$\sum_{t=1}^{n} w_{nt} u_t = \sum_{j=-\infty}^{\infty} e_j (T_n w_n)_j. \tag{7}$$

In words, the weighted sum (3) generates a linear operator T_n acting on L_p-approximable sequences $\{w_n\}$. The Beveridge–Nelson decomposition (Beveridge and Nelson, 1981) obtains from (7) by selecting equal weights $w_{n1} = \cdots = w_{nn} = 1$ and $c_j = 0$ for $j < 0$. Phillips and Solo (1992) derive all their results applying some variation of the Beveridge–Nelson decomposition, and this is why they have equal weights everywhere. On the other hand, Theorem 1 is based on the analysis of T_n in the whole class of L_p-approximable sequences, and there are as many different L_p-approximable sequences as there are functions in L_p. This explains the flexibility of the suggested approach. The essence of this paper is that all weight sequences arising in the Phillips framework are L_p-approximable, which provides the ground for application of Theorem 1. See Mynbaev and Ullah (2008) for an application to a spatial model.

The main statements and ideas are described in Section 2. All proofs are given in Section 3.

2. MAIN RESULTS

The number of different conditions in the theory of regressions with SV regressors may be daunting. For this reason Phillips consolidates related conditions and in many cases assumes more than is strictly necessary. When we use a statement from his paper, under a condition a little leaner than his, it means that his proof holds without serious changes under our condition.

A positive measurable function L on $[a, \infty)$, $a > 0$, is called *slowly varying* (SV) if

https://doi.org/10.1017/S0266466608090282 Published online by Cambridge University Press

$$\lim_{x\to\infty}\frac{L(\lambda x)}{L(x)}=1 \qquad \text{for any } \lambda>0. \tag{8}$$

By the Karamata theorem (Seneta, 1985, p. 10) this condition is equivalent to existence of functions δ,ε such that

$$L(x)=\exp\left(\delta(x)+\int_a^x \varepsilon(t)\frac{dt}{t}\right), \tag{9}$$

where δ is bounded and measurable and the limit $c=\lim_{x\to\infty}\delta(x)$ exists, ε is continuous, and $\lim_{x\to\infty}\varepsilon(x)=0$. Phillips argues that, for the purposes of asymptotic theory, not much is lost if one assumes $\delta=const$. Formally, this simplification is justified when δ is continuously differentiable and $t\delta'(t)\to 0$ as $t\to\infty$ because then (9) can be rewritten as

$$L(x)=\exp\left(\delta(a)+\int_a^x (t\delta'(t)+\varepsilon(t))\frac{dt}{t}\right).$$

The term *slowly varying* and its abbreviation SV will be used only in the general situation (9). In the special case (2) we write $L=K(\varepsilon)$ and call ε the *ε-function* of L. The functions ε and L can be extended continuously to $[0,a]$, and L can be redefined to be positive and continuous everywhere.

For some expansions we need to assume that ε also has a representation of type (2),

$$\varepsilon(x)=c_\varepsilon\exp\left(\int_a^x \eta(t)\frac{dt}{t}\right), \qquad x\ge a,$$

$$\eta \text{ is continuous} \quad \text{and} \quad \eta(x)\to 0 \quad \text{as } x\to\infty,$$

with a possibly negative constant c_ε. In such cases we also write $\varepsilon=K(\eta)$. The functions ε and η give rise to a *μ-function* of L defined by $\mu=\frac{1}{2}(\varepsilon+\eta)$. There are also a *$G$-function* and *$H$-function* defined by

$$G(t,n)=\frac{L(t)-L(n)}{L(n)\varepsilon(n)}, \qquad H(t,n)=\frac{G(t,n)-\log(t/n)}{\mu(n)}.$$

Their importance has been revealed by Phillips (2007, eqn. (60) and Lem. 7.5):

(a) If $L=K(\varepsilon)$ and ε is slowly varying, then

$$G(rn,n)=[1+o(1)]\log r \quad \text{uniformly in } r\in[a,b], \tag{10}$$

for any $0<a<b<\infty$.

(b) If $L=K(\varepsilon)$, $\varepsilon=K(\eta)$, and η is SV, then

$$H(rn,n)=[1+o(1)]\log^2 r \quad \text{uniformly in } r\in[a,b], \tag{11}$$

for any $0<a<b<\infty$.

https://doi.org/10.1017/S0266466608090282 Published online by Cambridge University Press

Equation (10) means second-order regular variation, and (11) means third-order; see de Haan and Resnick (1996).

To prove (1), for r bounded away from zero Phillips uses (10). In the neighborhood of zero, his argument is insufficient. The reason is clear from the basic definition (8). When $x \to \infty$ in that definition, λx goes to infinity at the same rate. However, if in the ratio $\frac{L(\lambda)}{L(x)}$ λ and x go to infinity at different rates, simple examples show that anything may happen to the ratio and an additional condition is necessary. The same problem arises with the L_p-approximability notion. To overcome it, we appeal to the theory of SV functions with remainder due to Aljančić, Bojanić, and Tomić (1955). We actually use the interpretation of that theory contained in the appendix by I.S. Shiganov to Seneta (1985). As the Russian edition of Seneta's book is not easily accessible to the Western reader, the required facts from that edition are either stated or proved in full in Section 3.

Let us call a *remainder* a positive function ϕ on $[0, \infty)$ with the following properties:

(a) ϕ is nondecreasing and $\phi(x) \to \infty$ as $x \to \infty$ and
(b) there exist positive numbers θ, X such that $x^{-\theta}\phi(x)$ is nonincreasing on $[X, \infty)$.

A positive measurable function L on $[a, \infty)$ is called *slowly varying with remainder* ϕ if for any $\lambda > 0$

$$\frac{L(\lambda x)}{L(x)} = 1 + O\left(1/\phi(x)\right), \qquad x \to \infty. \tag{12}$$

If $L = K(\varepsilon)$ and L is SV with remainder ϕ, we write $L = K(\varepsilon, \phi)$. It is easy to show that for all examples of SV functions used by Phillips one can put $\phi(x) = 1/|\varepsilon(x)|$, and the number θ can be taken arbitrarily close to zero. To prove (1), in addition to our main statement on L_p-approximability, the next assumption is sufficient (here and subsequently the references in parentheses are to Phillips (2007).

Assumption 2 (Assumption SSV)**.** $L = K(\varepsilon, \phi)$ where ε is SV and the remainder ϕ satisfies the following condition: for some $c > 0$

$$\frac{1}{c\phi(x)} \le |\varepsilon(x)| \le \frac{c}{\phi(x)} \quad \text{for } x \ge c. \tag{13}$$

When this inequality is true, one can say that $|\varepsilon|$ is quasi-monotone in the neighborhood of ∞.

THEOREM 2 (Lemma 7.4). *If Assumption 2 holds, k is a natural number, and $\theta \in (0, 1/k)$, then (1) is true.*

Theorem 2 will not be used in what follows, but we prove it to show how Phillips's proof can be fixed. A couple of other statements in the Phillips paper

https://doi.org/10.1017/S0266466608090282 Published online by Cambridge University Press

have the same problem, either directly or through (1). We refer to (10) as a *pointwise version* of second-order regular variation and to (1) as an *integral version*. An L_p-approximability version is next.

THEOREM 3. *For $p \in [1,\infty)$ and integer $k \geq 0$ define a vector $w_n \in \mathbb{R}^n$ by $w_{nt} = n^{-1/p} G^k(t,n)$, $t = 1,\dots,n$. If L satisfies Assumption 2 and $p\theta k < 1$, then $\{w_n\}$ is L_p-close to $W_k(x) = \log^k x$.*

The advantage of using L_p-approximability is that several statements of the Phillips framework can be derived from it at almost no cost.

COROLLARY 1. *Suppose Assumption 2 holds. Then the next three relations are true.*

(i) (p. 595, line 4 from bottom) If $\theta k < 1$, then $\lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^n G^k(t,n) = (-1)^k k!$

(ii) (a weaker version of (14)) If $\theta < 1$, then $\frac{1}{n} \sum_{t=1}^n L^k(t) = L^k(n) - kL^k(n)\, \varepsilon(n)\, [1+o(1)]$.

(iii) (p. 564, line 2 from bottom) If $2\theta < 1$, then $\frac{1}{n} \sum_{t=1}^n (L(t) - \bar{L})^2 = L^2(n)\, \varepsilon^2(n)\, [1+o(1)]$ where $\bar{L} = \frac{1}{n} \sum_{t=1}^n L(t)$.

The sequences of weights arising in the Phillips analysis of regressions are $x_n = (L(1),\dots,L(n))$, $y_n = (L(1) - \bar{L},\dots,L(n) - \bar{L})$ (see Phillips's eqn. (12)), and $z_n = ((L(1) - L(n))^k,\dots,(L(n) - L(n))^k)$ (see the definition of K_{nj} on p. 570). From Corollary 1 we see that $\|x_n\|_2 = \sqrt{n} L(n)[1+o(1)]$, $\|y_n\|_2 = \sqrt{n} L(n)|\varepsilon(n)|\, [1+o(1)]$, and $\|z_n\|_2 = \sqrt{n} (L(n)|\varepsilon(n)|)^k \sqrt{(2k)!}[1+o(1)]$. Hence, for standardizing those sequences, instead of $\|\cdot\|_2$-norms, one can use $\sqrt{n} L(n)$ for x_n, $\sqrt{n} L(n)\, \varepsilon(n)$ for y_n (the sign does not matter because of the symmetry of normal distributions), and $\sqrt{n} (L(n)\varepsilon(n))^k$ for z_n. This explains the Phillips choice of weights in the central limit results for linear regression, considered next.

COROLLARY 2 (Lemma 2.1). *Suppose that Assumptions 1 and 2 hold. Then the following statements are true.*

(i) If $2\theta < 1$, then $1/(\sqrt{n} L(n)) \sum_{t=1}^n L(t) u_t \to_d N(0,\sigma^2)$.

(ii) If $2\theta < 1$, then $1/(\sqrt{n} L(n)\varepsilon(n)) \sum_{t=1}^n (L(t) - \bar{L}) u_t \to_d N(0,\sigma^2)$.

(iii) If $2\theta k < 1$, then $\frac{1}{\sqrt{n}} \sum_{t=1}^n G^k(t,n) u_t \to_d N(0,\sigma^2 (2k)!)$.

Remark 5. In the case $\sum_{j=-\infty}^{\infty} c_j = 0$ everywhere in Corollary 2 convergence in distribution can be replaced by convergence in probability to zero. This qualification will not be proved here.

Now we turn to the central limit result for nonlinear regression (eqn. (53)):

$$\frac{1}{\sqrt{n} L(n)} \sum_{t=1}^n f\left(\frac{t}{n}\right) L(t) u_t \to_d N\left(0, \sigma^2 \int_0^1 f^2(r,\theta_0)\, dr\right) \tag{14}$$

https://doi.org/10.1017/S0266466608090282 Published online by Cambridge University Press

(the middle term in the Phillips equation is omitted as it is in terms of Brownian motion, whereas we do not use it). In the following theorem we require just continuity of f instead of twice differentiability assumed by Phillips.

THEOREM 4 (Lemma 6.1). *Let f be a continuous function of arguments $(r, \theta) \in [0, 1] \times \Theta$, where Θ is a compact parameter space, and let θ_0 be an interior point of Θ. Taking a sequence of diagonal matrices whose elements diverge to ∞ as $n \to \infty$, define a shrinking neighborhood $N_n = \{\theta \in \Theta : \|C_n(\theta - \theta_0)\|_2 \leq 1\}$ of θ_0. If u_t satisfy Assumption 1, L satisfies Assumption 2, and $2\theta < 1$, then (14) holds uniformly over $\theta \in N_n$.*

Remark 6. Some arguments in the Phillips paper are heuristic. For example, it has not been proved that the ordinary least squares (OLS) estimators for his system (42) and the system obtained from (42) by omitting the $o(1)$ are asymptotically close to one another. His expansion (41) is not sufficient for this purpose. It is possible to make this step rigorous by proving an L_p-approximability version of (11) (under conditions more restrictive than Assumption 2). The result, not to be formulated or proved here, shows that the situation in the case of two different SV regressors is more complex than stated in Phillips (2007, Thm 5.1).

3. PROOFS

We apply only two nontrivial facts from the theory of SV functions with remainder. One of them is a modification of Karamata's representation theorem (Seneta, 1985, p. 100): L is an SV function with remainder ϕ if and only if it has representation (9) with the following bounds on the rates of convergence:

$$\delta(x) - c = O(1/\phi(x)), \qquad \varepsilon(x) = O(1/\phi(x)). \tag{15}$$

The proof is quite complex and will not be reproduced here. The second fact is the bound on $|L(\lambda x)/L(x) - 1|$ contained in the next lemma.

LEMMA 1 (Seneta, 1985, p. 102). *If L is SV with remainder ϕ, then for any $b > \theta$ (where θ is the number from the definition of the remainder) there exist numbers $M_b > 0$ and $B_b \geq a$ such that*

$$\left| \frac{L(\lambda x)}{L(x)} - 1 \right| \leq \frac{M_b \lambda^{-b}}{\phi(x)} \quad \text{for all } x \geq B_b \quad \text{and} \quad \frac{B_b}{x} \leq \lambda \leq 1 \tag{16}$$

(the larger x, the closer λ is allowed to be to 0).

Proof. In Seneta's book there is a typo: this lemma is stated with $b > 0$ instead of $b > \theta$. Let $x \geq a$ and $a/x \leq \lambda \leq 1$. Denote $U(\lambda, x) = \log(L(\lambda x)/L(x))$. An elementary bound $|e^x - 1| \leq |x| e^{|x|}$ shows that it suffices to estimate the right-hand side of

$$|L(\lambda x)/L(x) - 1| = |e^{U(\lambda,x)} - 1| \leq |U(\lambda, x)| e^{|U(\lambda,x)|}. \tag{17}$$

https://doi.org/10.1017/S0266466608090282 Published online by Cambridge University Press

The Karamata representation (9) implies

$$U(\lambda,x)=\log L(\lambda x)-\log L(x)=\delta(\lambda x)-\delta(x)-\int_{\lambda x}^{x}\varepsilon(t)\frac{dt}{t}, \tag{18}$$

where $\lambda x\le x$. By (15) there exists a constant $K>0$ such that

$$|\delta(x)-c|\le K/\phi(x),\qquad |\varepsilon(x)|\le K/\phi(x).$$

Because ϕ is nondecreasing, we get

$$\begin{aligned}|U(\lambda,x)|&\le|\delta(\lambda x)-c|+|c-\delta(x)|+\int_{\lambda x}^{x}|\varepsilon(t)|\frac{dt}{t}\\ &\le\frac{K}{\phi(\lambda x)}+\frac{K}{\phi(x)}+K\int_{\lambda x}^{x}\frac{1}{\phi(t)}\frac{dt}{t}\\ &\le\frac{2K}{\phi(\lambda x)}+\frac{K}{\phi(\lambda x)}\int_{\lambda x}^{x}\frac{dt}{t}=K(2-\log\lambda)/\phi(\lambda x).\end{aligned} \tag{19}$$

Fix some $b>\theta$. Using monotonicity of ϕ and the fact that it increases to ∞ at ∞, we can choose $B_b\ge a$ so large that $B_b\le x\lambda$ will imply $K/\phi(\lambda x)\le K/\phi(B_b)<(b-\theta)/2$. Then by (20)

$$|U(\lambda,x)|\le\frac{b-\theta}{2}(2-\log\lambda)=b-\theta-\frac{b-\theta}{2}\log\lambda. \tag{20}$$

On the other hand, by property (b) of the remainder ϕ the inequality $\lambda x\le x$ implies $(\lambda x)^{-\theta}\phi(\lambda x)\ge x^{-\theta}\phi(x)$ and $1/\phi(\lambda x)\le\lambda^{-\theta}/\phi(x)$. Hence, from (20)

$$|U(\lambda,x)|\le K\lambda^{-\theta}(2-\log\lambda)/\phi(x). \tag{21}$$

Using (21) for the first factor at the right of (17) and (20) for the second factor, we see that

$$|L(\lambda x)/L(x)-1|\le\frac{K\lambda^{-\theta}}{\phi(x)}(2-\log\lambda)e^{b-\theta}\lambda^{-(b-\theta)/2}. \tag{22}$$

It is easy to see that for any $a>0$ there is a constant $c>0$ such that $2-\log\lambda\le c\lambda^{-a}$ for $0<\lambda\le 1$. Putting here $a=(b-\theta)/2$ we can continue (22) as

$$|L(\lambda x)/L(x)-1|\le\frac{K}{\phi(x)}e^{b-\theta}c\lambda^{-\theta-(b-\theta)/2-(b-\theta)/2}=M_b\lambda^{-b}/\phi(x),$$

where $M_b=Kce^{b-\theta}$. ■

Remark 7. Because in practical cases the number θ can be arbitrarily close to zero, the number $b>\theta$ can also be as close to zero as desired.

https://doi.org/10.1017/S0266466608090282 Published online by Cambridge University Press

Proof of Theorem 2. Using the identity

$$\int_0^1 \log^k r\, dr = (-1)^k k! \tag{23}$$

we rewrite (1) as

$$\int_0^1 \left[G^k(rn,n) - \log^k r \right] dr \to 0. \tag{24}$$

Let $0 < \delta < 1$. Then (10) implies

$$\int_\delta^1 \left[G^k(rn,n) - \log^k r \right] dr \to 0, \qquad n \to \infty. \tag{25}$$

For $n > B_b/\delta$ the interval $(B_b/n, \delta)$ is not empty, and by Lemma 1

$$|G(rn,n)| \le \frac{M_b r^{-b}}{\phi(n)|\varepsilon(n)|} \quad \text{for } n \ge \frac{B_b}{\delta} \quad \text{and} \quad \frac{B_b}{n} \le r \le \delta.$$

Hence, using also the left inequality in (13), we have

$$\int_{B_b/n}^\delta |G^k(rn,n)|\, dr \le \int_0^\delta \left(\frac{M_b r^{-b}}{\phi(n)|\varepsilon(n)|} \right)^k dr \le \left(\frac{M_b}{c} \right)^k \int_0^\delta r^{-bk}\, dr = c_1 \delta^{1-bk}. \tag{26}$$

Here the right-hand side tends to zero if $b \in (0, 1/k)$. This condition can be satisfied because b is arbitrarily close to θ.

If $0 < r \le B_b/n$, then $rn \le B_b$ and $L(rn) \le c$ by the assumed continuity of L. As a result,

$$|G(rn,n)| \le \frac{1}{|\varepsilon(n)|} \left(\frac{c}{L(n)} + 1 \right), \tag{27}$$

where the function on the right is slowly varying and grows no faster than n^a, for any $a > 0$. It follows that

$$\int_0^{B_b/n} |G^k(rn,n)|\, dr \le \frac{1}{|\varepsilon(n)|^k} \left(\frac{c}{L(n)} + 1 \right)^k \frac{B_b}{n} \to 0, \qquad n \to \infty. \tag{28}$$

Expressions (25), (26), and (28), together with

$$\left| \int_0^\delta \log^k r\, dr \right| \to 0, \qquad \delta \to 0,$$

show that we can choose first a small δ and then a large n to make the left side of (24) as small as desired. ∎

https://doi.org/10.1017/S0266466608090282 Published online by Cambridge University Press

Proof of Theorem 3. Instead of the discretization operator δ_{np}, here it is more convenient to use the interpolation operator $\Delta_{np}: \mathbb{R}^n \to L_p$ defined by

$$\Delta_{np} w = n^{1/p} \sum_{t=1}^{n} w_t 1_{i_t},$$

where 1_{i_t} is the indicator of i_t. When $p < \infty$, definition (5) is equivalent to $\left\| \Delta_{np} w_n - W \right\|_p \to 0$ (Mynbaev, 2001, p. 313).

The case $k = 0$ is trivial, and so let $k > 0$. The equation

$$\Delta_{np} w_n = \sum_{t=1}^{n} G^k(t,n) 1_{i_t} \tag{29}$$

is equivalent to n equations $\left(\Delta_{np} w_n\right)(u) = G^k(t,n)$ for $u \in i_t$, $t = 1, \dots, n$. The condition $u \in i_t$ is equivalent to the condition that t is an integer satisfying $t \le nu + 1 < t + 1$, which, in turn, is equivalent to $t = [nu + 1]$ ($[a]$ stands for an integer part of a). Hence, (29) takes a compact form

$$\left(\Delta_{np} w_n\right)(u) = G^k([nu+1], n), \qquad 0 \le u < 1. \tag{30}$$

Let $0 < \delta \le 1/2$. We are going to apply Lemma 1. For $n > n_1 \equiv B_b/\delta$ the interval $(B_b/n, \delta)$ is not empty, and by the triangle inequality

$$\begin{aligned} \left\| \Delta_{np} w_n - W_k \right\|_p \le{}& \left\| \Delta_{np} w_n - W_k \right\|_{p,(\delta,1)} + \| W_k \|_{p,(0,\delta)} \\ &+ \left\| \Delta_{np} w_n \right\|_{p,(0,B_b/n)} + \left\| \Delta_{np} w_n \right\|_{p,(B_b/n,\delta)}. \end{aligned} \tag{31}$$

Here we denote $\| W \|_{p,(a,b)} = \left(\int_a^b |W(x)|^p \, dx \right)^{1/p}$. Obviously, $\| W_k \|_{p,(0,\delta)} \to 0$ as $\delta \to 0$. For the other three terms we consider three cases.

Case $\delta \le u < 1$. According to (10),

$$G^k(rn, n) = [1 + o(1)] \log^k r \quad \text{uniformly in } r \in \left[\delta, 1 + \frac{1}{2B_b} \right]. \tag{32}$$

Defining $r = [nu + 1]/n$, from the inequality $nu < [nu + 1] \le nu + 1$ we have

$$\delta \le u < r \le u + \frac{1}{n} < 1 + \frac{1}{n_1} \le 1 + \frac{1}{2B_b}, \tag{33}$$

so that

$$r = u + o(1) \quad \text{and} \quad r \in \left[\delta, 1 + \frac{1}{2B_b} \right]. \tag{34}$$

Equations (32) and (34) lead to $G^k([nu+1], n) - \log^k u = o(1)$ uniformly in $u \in (\delta, 1)$. This proves that

$$\left\| \Delta_{np} w_n - W_k \right\|_{p,(\delta,1)} \to 0, \qquad n \to \infty. \tag{35}$$

https://doi.org/10.1017/S0266466608090282 Published online by Cambridge University Press

Case $B_b/n \le u < \delta$. Let $n > n_2 \equiv \max\{n_1, 2\}$. Then (33) and the conditions $u \in [B_b/n, \delta)$, $n > n_2$ imply

$$\frac{B_b}{n} \le u < r \le u + \frac{1}{n} < \delta + \frac{1}{n_2} \le 1.$$

This means we can apply Lemma 1, the left inequality in (13), and (33) to get

$$|G^k([nu+1], n)| \le \left[\frac{M_b}{r^b \phi(n)|\varepsilon(n)|}\right]^k \le \left[\frac{M_b}{c}\right]^k u^{-bk} \quad \text{for } u \in \left[\frac{B_b}{n}, \delta\right].$$

Fixing $b \in (\theta, 1/(pk))$ we have with a new constant c_1 (independent of n and δ)

$$\int_{B_b/n}^{\delta} |\Delta_{np} w_n|^p \, du \le c_1 \int_0^{\delta} u^{-pbk} \, du = \frac{c_1}{1-pbk} \delta^{1-pbk}. \tag{36}$$

Case $0 < u < B_b/n$. In this case $[nu+1] \le nu+1 < B_b+1$ and $L([nu+1]) \le c$ by the assumed continuity of L. Hence, (27) is true, and

$$\left\| \Delta_{np} w_n \right\|_{p,(0,B_b/n)} \le \left[\left(\frac{c}{L(n)} + 1\right) \frac{1}{|\varepsilon(n)|}\right]^k \left(\frac{B_b}{n}\right)^{1/p}. \tag{37}$$

Here the expression on the right tends to zero because the expression in the brackets is a slowly varying function. From (35)–(37) we see that we can choose first a small δ and then a large n to make the left side of (31) as small as desired. ■

Proof of Corollary 1.

(i) Put $p = 1$ in Theorem 3. Equations (23) and (29) imply

$$\left| \frac{1}{n} \sum_{t=1}^{n} G^k(t, n) - (-1)^k k! \right| = \left| \int_0^1 \Delta_{n1} w_n \, du - \int_0^1 \log^k u \, du \right|$$

$$\le \| \Delta_{n1} w_n - W_k \|_1 \to 0.$$

(ii) Letting $k = 1$ in statement (i) we have

$$\frac{1}{n} \sum_{t=1}^{n} G(t, n) = -[1 + o(1)],$$

which can be rearranged as

$$\frac{1}{nL(n)} \sum_{t=1}^{n} L(t) = 1 - \varepsilon(n)[1 + o(1)]. \tag{38}$$

The ε-function in (2) is related to L by $\varepsilon(x) = xL'(x)/L(x)$. Therefore the ε-function of L^k equals $x(L^k(x))'/L^k(x) = k\varepsilon(x)$. Equation (12) implies $L^k(\lambda x)/L^k(x) = 1 + O(1/\phi(x))$. Thus, the condition $L = K(\varepsilon, \phi)$ implies $L^k = K(k\varepsilon, \phi)$, and L^k satisfies Assumption 2 with $k\varepsilon$ in place of ε. Application of (38) to L^k proves (ii).

https://doi.org/10.1017/S0266466608090282 Published online by Cambridge University Press

(iii) Simple algebra shows that

$$\frac{1}{n}\sum_{t=1}^{n}\left(L(t)-\bar{L}\right)^2=\frac{1}{n}\sum_{t=1}^{n}(L(t)-L(n))^2-\left[\frac{1}{n}\sum_{t=1}^{n}(L(t)-L(n))\right]^2,$$

and, therefore, by statement (i)

$$\frac{1}{nL^2(n)\varepsilon^2(n)}\sum_{t=1}^{n}\left(L(t)-\bar{L}\right)^2$$

$$=\frac{1}{n}\sum_{t=1}^{n}G^2(t,n)-\left[\frac{1}{n}\sum_{t=1}^{n}G(t,n)\right]^2\to 2!-1!=1. \qquad \blacksquare$$

Proof of Corollary 2. By Theorem 1, in all cases it is enough to establish L_2-approximability of the sequence of weights in question.

(i) Setting $p=2$, $k=1$ in Theorem 3 gives (see (30))

$$\int_0^1\left|\frac{L([nu+1])-L(n)}{L(n)\varepsilon(n)}-\log u\right|^2 du\to 0.$$

Multiply this relation by $\varepsilon^2(n)\to 0$ to obtain

$$\int_0^1\left|\frac{L([nu+1])}{L(n)}-1\right|^2 du\to 0.$$

In a way similar to the transition from (29) to (30) we have

$$\Delta_{n2}\left[\frac{1}{\sqrt{n}L(n)}(L(1),\dots,L(n))\right]=\frac{1}{L(n)}\sum_{t=1}^{n}L(t)1_{i_t}=\frac{L([nu+1])}{L(n)}.$$

The last two equations show that the sequence $w_n=1/(\sqrt{n}L(n))(L(1),\dots,L(n))$ is L_2-close to $W\equiv 1$.

(ii) From (38) we conclude that $\bar{L}=L(n)-\varepsilon(n)L(n)[1+o(1)]$ and that the sequence of weights in statement (ii) is

$$w_n=\frac{1}{\sqrt{n}L(n)\varepsilon(n)}(L(1)-\bar{L},\dots,L(n)-\bar{L})$$

$$=\frac{1}{\sqrt{n}}(G(1,n),\dots,G(n,n))+\frac{1+o(1)}{\sqrt{n}}(1,\dots,1).$$

By Theorem 3 with $p=2$, $k=0$ the second sequence on the right is L_2-close to $W_0\equiv 1$. The same theorem with $k=1$ shows that the first sequence is L_2-close to $W_1(x)=\log x$. Hence, w_n is L_2-close to $1+\log x$. The statement follows from the fact that $\int_0^1(1+\log u)^2\,du=1$.

Statement (iii) follows directly from Theorem 3 and (23). $\blacksquare$

https://doi.org/10.1017/S0266466608090282 Published online by Cambridge University Press

Proof of Theorem 4. Denote

$$w_n(\theta) = \left(f\left(\frac{1}{n}, \theta\right), \ldots, f\left(\frac{n}{n}, \theta\right)\right), \qquad w_n = \frac{1}{\sqrt{n} L(n)} (L(1), \ldots, L(n))$$

and let $w_n(\theta) w_n$ mean their element-wise product

$$w_n(\theta) w_n = (w_{n1}(\theta) w_{n1}, \ldots, w_{nn}(\theta) w_{nn}).$$

As n goes to infinity, the partition $\{i_t\}$ becomes finer and finer and the neighborhood N_n shrinks to θ_0. The weight w_n has been shown to be L_2-close to $W \equiv 1$. The idea is that $w_n(\theta)$ should be L_∞-close to $F(r) \equiv f(r, \theta_0)$ and the product $w_n(\theta) w_n$ should be L_2-close to $FW = F$.

Step 1. Let us prove that

$$\sup_{\theta \in N_n} \|z_n(\theta)\|_2 \to 0, \tag{39}$$

where $z_n(\theta) = w_n(\theta) w_n - \delta_{n2} F$. Because f is uniformly continuous on $[0, 1] \times \Theta$, for any $\varepsilon > 0$ there exists $\delta > 0$ such that $\left\|(r', \theta) - (r, \theta_0)\right\|_2 \le \delta$ implies $|f(r', \theta) - f(r, \theta_0)| \le \varepsilon$. Choose n_1 so large that

$$\sup_{r \in i_t} \left| r - \frac{t}{n} \right| \le \frac{\delta}{\sqrt{2}}, \qquad \sup_{\theta \in N_n} \|\theta - \theta_0\|_2 \le \frac{\delta}{\sqrt{2}} \quad \text{for } n \ge n_1.$$

Then

$$\sup_{r \in i_t,\ \theta \in N_n} \left\| \left(\frac{t}{n}, \theta\right) - (r, \theta_0) \right\|_2 = \sup_{r \in i_t,\ \theta \in N_n} \sqrt{\left(r - \frac{t}{n}\right)^2 + \|\theta - \theta_0\|_2^2} \le \delta,$$

which implies

$$\sup_{r \in i_t,\ \theta \in N_n} \left| f\left(\frac{t}{n}, \theta\right) - f(r, \theta_0) \right| \le \varepsilon. \tag{40}$$

Notice also that

$$\begin{aligned} \left| \left[\delta_{n2} \left(f\left(\frac{t}{n}, \theta_0\right) - f(\cdot, \theta_0) \right) \right]_t \right| &\le \sqrt{n} \int_{i_t} \left| f\left(\frac{t}{n}, \theta_0\right) - f(r, \theta_0) \right| dr \\ &\le \varepsilon \sqrt{n} \int_{i_t} dr = \varepsilon |(\delta_{n2} W)_t|. \end{aligned}$$

Now we bound the tth coordinate of $z_n(\theta)$ using the last estimate and (40):

$$|z_{nt}(\theta)| = \left| f\left(\frac{t}{n}, \theta\right) w_{nt} - (\delta_{n2} f(\cdot, \theta_0))_t \right|$$

https://doi.org/10.1017/S0266466608090282 Published online by Cambridge University Press

$$\le \left| w_{nt} \left[f\left(\frac{t}{n},\theta\right) - f\left(\frac{t}{n},\theta_0\right)\right]\right| + \left|[w_{nt} - (\delta_{n2}W)_t]\, f\left(\frac{t}{n},\theta_0\right)\right|$$

$$+ \left|\left[\delta_{n2}\left(f\left(\frac{t}{n},\theta_0\right) - f(\cdot,\theta_0)\right)\right]_t\right|$$

$$\le \varepsilon|w_{nt}| + \max_{r\in[0,1]} |f(r,\theta_0)|\,|w_{nt} - (\delta_{n2}W)_t| + \varepsilon|(\delta_{n2}W)_t|. \tag{41}$$

Further, denoting $c = \max_{r\in[0,1]} |f(r,\theta_0)|$, we deduce from (41)

$$|z_{nt}(\theta)| \le (\varepsilon + c)|w_{nt} - (\delta_{n2}W)_t| + 2\varepsilon|(\delta_{n2}W)_t|.$$

Applying l_2-norms and using boundedness of δ_{n2} (Mynbaev, 2001, eqn. (2.2)) we get

$$\sup_{\theta\in N_n} \|z_n(\theta)\|_2 \le (\varepsilon + c)\,\|w_n - \delta_{n2}W\|_2 + 2\varepsilon\,\|\delta_{n2}W\|_2$$

$$\le (\varepsilon + c)\,\|w_n - \delta_{n2}W\|_2 + 2\varepsilon,$$

which proves (39).

Step 2. Note that, because $\{\delta_{n2}F\}$ is L_2-generated, equation (6) is true with $w_{nt} = (\delta_{n2}F)_t$ and $W(x) = f(x,\theta_0)$. Because

$$\frac{1}{\sqrt{n}L(n)} \sum_{t=1}^{n} f\left(\frac{t}{n},\theta\right) L(t)u_t = \sum_{t=1}^{n} z_{nt}u_t + \sum_{t=1}^{n} (\delta_{n2}F)_t u_t,$$

to prove (14) it suffices to show that

$$\sup_{\theta\in N_n} \mathrm{E}\left(\sum_{t=1}^{n} z_{nt}(\theta)u_t\right)^2 \to 0, \qquad n\to\infty.$$

Identity (7), orthogonality of the innovations e_t, and the bound $\|T_n w\|_2 \le \sqrt{3}\,\sum_j |c_j|\,\|w\|_2$ (Mynbaev, 2001, Thm. 2.2(a)) lead to

$$\mathrm{E}\left(\sum_{t=1}^{n} z_{nt}(\theta)u_t\right)^2 = \mathrm{E}\left[\sum_{i=-\infty}^{\infty} e_i (T_n z_n(\theta))_i\right]^2$$

$$= \sigma_e^2\,\|T_n z_n(\theta)\|_2^2 \le c\,\|z_n(\theta)\|_2^2 \to 0, \qquad n\to\infty,$$

uniformly over $\theta \in N_n$. ■

https://doi.org/10.1017/S0266466608090282 Published online by Cambridge University Press

REFERENCES

Aljančić, S., R. Bojanić, & M. Tomić (1955) Deux théorèmes relatifs au comportement asymptotique des séries trigonométriques. *Zbornik Radova Matematički Institut SANU* 43, 15–26.

Anderson, T.W. (1971) *The Statistical Analysis of Time Series*. Wiley.

Beveridge, S. & C.R. Nelson (1981) A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the "business cycle." *Journal of Monetary Economics* 7, 151–174.

Cremers, H. & D. Kadelka (1986) On weak convergence of integral functionals of stochastic processes with applications to processes taking paths in L_p^E. *Stochastic Processes and Their Applications* 21, 305–317.

de Haan, L. & S. Resnick (1996) Second-order regular variation and rates of convergence in extreme-value theory. *Annals of Probability* 24, 97–124.

de Jong, R.M. & J. Davidson (2000) The functional central limit theorem and weak convergence to stochastic integrals part, I: Weakly dependent processes. *Econometric Theory* 16, 621–642.

Hall, P. & C.C. Heyde (1980) *Martingale Limit Theory and Its Application*. Academic Press.

Hannan, E.J. (1970) *Multiple Time Series*. Wiley.

Hannan, E.J. (1979) The central limit theorem for time series regression. *Stochastic Processes and Their Applications* 9, 281–289.

Moussatat, M.W. (1976) On the asymptotic theory of statistical experiments and some of its applications. Ph.D. Dissertation, University of California, Berkeley.

Mynbaev, K.T. (2001) L_p-approximable sequences of vectors and limit distribution of quadratic forms of random variables. *Advances in Applied Mathematics* 26, 302–329.

Mynbaev, K.T. & A. Ullah (2008) Asymptotic distribution of the OLS estimator for a purely autoregressive spatial model. *Journal of Multivariate Analysis* 99, 245–277.

Phillips, P.C.B. (2007) Regression with slowly varying regressors and nonlinear trends. *Econometric Theory* 23, 557–614.

Phillips, P.C.B. & V. Solo (1992) Asymptotics for linear processes. *Annals of Statistics* 20, 971–1001.

Pötscher, B.M. & I.R. Prucha (1991) Basic structure of the asymptotic theory in dynamic nonlinear econometric models, part I: Consistency and approximation concepts. *Econometric Reviews* 10, 125–216. Available at http://gemini.econ.umd.edu/jrust/econ551/syb551/node5.html.

Seneta, E. (1985) *Pravil′no menyayushchiesya funktsii* (in Russian) [*Regularly varying functions*] with appendixes by I.S. Shiganov and V.M. Zolotarev. "Nauka," Moscow.

Tanaka, K. (1996) *Time Series Analysis*. Wiley.

https://doi.org/10.1017/S0266466608090282 Published online by Cambridge University Press