Regressions with Asymptotically Collinear Regressor

Mynbaev, Kairat

Kazakh-British Technical University

August 2009

Regressions with Asymptotically Collinear Regressors

Kairat T. Mynbaev

International School of Economics

Kazakh-British Technical University

Tolebi 59, Almaty 050000, Kazakhstan

kairat_mynbayev@yahoo.com

Abstract

We investigate the asymptotic behavior of the OLS estimator for regressions with two slowly varying regressors. It is shown that the asymptotic distribution is normal one-dimensional and may belong to one of four types depending on the relative rates of growth of the regressors. The analysis establishes, in particular, a new link between slow variation and L_p-approximability. A revised version of this paper has been published in Econometrics Journal (2011), volume 14, pp. 304–320.

Keywords Asymptotic distribution theory; Linear regression; Asymptotically collinear regressors

1 Introduction

Regressions with asymptotically collinear regressors arise in a number of applications, both in linear and nonlinear settings. The examples are the log-periodogram analysis of long memory (see Robinson (1995); Hurvich et al. (1998); Phillips (1999) and references therein), the study of growth convergence (Barro & Sala-i Martin, 2003), and nonlinear least squares estimation (Wu, 1981). Phillips (2007) has developed a powerful method to analyze such regressions. Using the theory of slowly varying functions (for the definition see (Bingham et al., 1987)) he has proved asymptotic normality of OLS estimators (with an appropriate standardization). He has also shown that the usual regression formulas for asymptotic standard errors are valid. The limit distribution of the regression coefficients has been shown to be one-dimensional.

In the paper just cited, Phillips has considered a variety of situations, from simple regression to nonlinear regression. In case of simple regression and a polynomial regression in a slowly varying function his treatment is complete. However, in case of two different slowly varying regressors, as in

$$y_s = \beta_0 + \beta_1 L_1(s) + \beta_2 L_2(s) + u_s, \tag{1}$$

Phillips limited himself to a heuristic argument. The purpose of this paper is to provide a rigorous result for (1).

Following Phillips (2007) let us consider slowly varying functions L with Karamata representation

$$L(x) = c \exp\left(\int_a^x \varepsilon(t)dt/t\right) \tag{2}$$

and call the function ε in this representation an *ε-function of L*. We say that two models of form (1) with pairs of SV functions (L_1, L_2) and $(\tilde{L}_1, \tilde{L}_2)$ are of *different (asymptotic) types* if their asymptotic distributions contain functions of sample size n with different asymptotic behavior as $n \to \infty$. (Phillips, 2007, Theorem 5.1) suggests that there are two types of model (1): one kind of asymptotics is true when the ε-functions of L_1, L_2 satisfy $\varepsilon_2(n) = o(\varepsilon_1(n))$ and another holds when $\varepsilon_1(n) = o(\varepsilon_2(n))$. Our

Table 1: Basic SV functions ($l_1(x) = \log x,\ l_2(x) = \log(\log x)$)

L	ε	η	μ	δ
$L_1 = l_1$	$1/l_1$	$-1/l_1$	0	0
$L_2 = l_2$	$1/(l_1 l_2)$	$-(1+l_2)/(l_1 l_2)$	$-1/(2l_1)$	$-1/(2l_1^2)$
$L_3 = 1/l_1$	$-1/l_1$	$-1/l_1$	$-1/l_1$	$1/l_1^3$
$L_4 = 1/l_2$	$-1/(l_1 l_2)$	$-(1+l_2)/(l_1 l_2)$	$-\frac{2+l_2}{2l_1 l_2}$	$\frac{2+l_2}{2l_1^2 l_2^3}$

classification theorem below shows that the number of different types is at least four (this number depends on notational conventions) and is determined by such fine characteristics of the regressors as ε-functions of their ε-functions. In all cases we prove a Phillips type result that the limit distribution is normal and one-dimensional.

The most unexpected outcome is that the asymptotic distribution depends on the true β in a discontinuous fashion: the asymptotic variances along the axes $\beta_1 = 0$ and $\beta_2 = 0$ differ from those for $\beta_1 \neq 0$ and $\beta_2 \neq 0$. The method in principle is applicable to regressions with more than two different slowly varying regressors. However, we are not sure that such generalizations are required for empirical work.

In Section 2 we state the main results. All proofs are given in the Appendix.

2 Main results

2.1 Slowly varying functions

Here the properties of slowly varying functions are reviewed to the extent required later.

The name *slowly varying* and its abbreviation SV will be used for a positive measurable function on $[A, \infty)$, where $A > 0$, satisfying the condition $\lim_{x\to\infty} L(rx)/L(x) = 1$ for any $r > 0$. Functions with representation (2), where ε is continuous and $\lim_{x\to\infty} \varepsilon(x) = 0$, constitute a special case of SV functions. In (2) the constant c is allowed to be negative and the ε-function of L can be found as $\varepsilon(x) = xL'(x)/L(x)$.

In most of the present theory, the ε-function is also assumed to be of form (2), that is $\varepsilon(x) = c\exp\left(\int_a^x \eta(t)dt/t\right)$ where $\lim_{x\to\infty} \eta(x) = 0$ and η is continuous. Denote $\mu(x) = (\varepsilon(x) + \eta(x))/2,\ \delta(x) = L(x)\varepsilon(x)\mu(x)$. These functions will be called μ- and δ-*functions of* L, respectively. When the argument of ε, μ, δ is the sample size n, that argument will be usually suppressed. Table 1 contains a summary of practically important cases.

We observe that for l_1 both μ and δ are identically zero. In all other cases ε, μ and δ are nonzero for all large x. Our analysis shows that the asymptotic theory of model (1) depends on the asymptotic behavior of the ratios δ_1/δ_2 and $\varepsilon_1/\varepsilon_2$. In the next assumption we impose on these ratios conditions general enough to include all possible pairs of functions from Table 1.

Assumption ε-δ.

(a) In the pair (L_1, L_2) only one function is allowed to be $\log x$ (and have a vanishing δ-function). By changing the notation, if necessary, one can assume that if one of L_1, L_2 is $\log x$, then it is always L_1.

(b) If neither δ_1 nor δ_2 vanishes, we require either $\delta_1/\delta_2 \to 0$ or $\delta_2/\delta_1 \to 0$ to be true at infinity.

(c) None of ε_1 and ε_2 vanishes. Besides, either $\varepsilon_1/\varepsilon_2$ tends to a constant $\kappa \in \mathbb{R}$ or $\varepsilon_2/\varepsilon_1 \to 0$.

This assumption will be used everywhere without explicitly mentioning. Regarding condition (b), we note that for the functions from Table 1 convergence of δ_1/δ_2 to a constant different from 0 does not occur. Such convergence would require higher-order expansions for its analysis which will not be considered here. The case when both δ_1 and δ_2 vanish but L_1, L_2 are not $\log x$ would also require higher-order expansions. When $L_1(x) = \log x$, model (1) is called *semi-reduced*. When both δ_1 and δ_2 are nonzero, model (1) is called *non-reduced*.

Definition. We write $L = K(\varepsilon, \phi_\varepsilon, \theta_\varepsilon)$ if

(a) L is continuous on $[0, \infty)$ and has Karamata representation (2) for some $a > 0$, where ε is SV, continuous and $\varepsilon(x) \to 0$ as $x \to \infty$.

(b) There exist a constant $c > 0$ and a function ϕ_ε on $[0, \infty)$ with properties:

  (i) ϕ_ε is positive, nondecreasing on $[0, \infty)$, $\phi_\varepsilon(x) \to \infty$ as $x \to \infty$,

  (ii) there exist positive numbers $\theta_\varepsilon, X_\varepsilon$ such that $x^{-\theta_\varepsilon}\phi_\varepsilon(x)$ is nonincreasing on $[X_\varepsilon, \infty)$, and

  (iii) for all $x \geq c$

$$\frac{1}{c\phi_\varepsilon(x)} \leq |\varepsilon(x)| \leq \frac{c}{\phi_\varepsilon(x)}. \tag{3}$$

The right inequality in condition (3) means that L is slowly varying with remainder ϕ_ε (see Aljančić et al. (1955)). All practical examples from Table 1 satisfy the above definition with $\phi_\varepsilon(x) = 1/|\varepsilon(x)|$ and the number $\theta_\varepsilon > 0$ can be chosen as close to 0 as desired. This follows from the next property of SV functions: if L is SV, then for any $\theta > 0$, $x^\theta L(x) \to \infty$ and $x^{-\theta}L(x) \to 0$ as $x \to \infty$. (Phillips, 2007, Assumption SSV) does not have the (b) part while it seems to be essential for the most important statements.

2.2 Transformation of the regressor space

Assuming that $\mu(n) \neq 0$ denote

$$H(t,n) = \left[\frac{L(t)-L(n)}{L(n)\varepsilon(n)} - \log\left(\frac{t}{n}\right)\right]\frac{1}{\mu(n)},\ 0 < t \le n,$$

and let us call this function an *H-function* of L. Under certain conditions (Phillips, 2007)

$$H(rn,n) = \log^2 r + o(1) \text{ uniformly in } r \in [a,b], \tag{4}$$

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

(Phillips, 2007, p.573) suggested to transform the regressor space as follows. (4) implies $L_j(rn)/L_j(n) - 1 = \varepsilon_j(n)\log r + \varepsilon_j(n)\mu_j(n)\log^2 r[1+o(1)]$. Using this expansion and suppressing the argument n in $L_j, \varepsilon_j, \mu_j$ (1) can be rewritten as:

$$\begin{aligned} y_s &= \beta_0 + \beta_1 L_1 + \beta_2 L_2 \\ &\quad + \beta_1 L_1 \varepsilon_1 \log\frac{s}{n} + \beta_1 L_1 \varepsilon_1 \mu_1 \log^2\frac{s}{n}[1+o(1)] \\ &\quad + \beta_2 L_2 \varepsilon_2 \log\frac{s}{n} + \beta_2 L_2 \varepsilon_2 \mu_2 \log^2\frac{s}{n}[1+o(1)] + u_s. \end{aligned} \tag{5}$$

Dropping here $o(1)$ produces an approximation to (5):

$$\begin{aligned} y_s &= \beta_0 + \beta_1 L_1 + \beta_2 L_2 + (\beta_1 L_1 \varepsilon_1 + \beta_2 L_2 \varepsilon_2)\log\frac{s}{n} \\ &\quad + (\beta_1\delta_1 + \beta_2\delta_2)\log^2\frac{s}{n} + u_s. \end{aligned} \tag{6}$$

Denoting

$$\beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix},\ \gamma_n = \begin{pmatrix} \gamma_{n0} \\ \gamma_{n1} \\ \gamma_{n2} \end{pmatrix},\ A_n = \begin{pmatrix} 1 & L_1 & L_2 \\ 0 & L_1\varepsilon_1 & L_2\varepsilon_2 \\ 0 & \delta_1 & \delta_2 \end{pmatrix} \tag{7}$$

we obtain

$$y_s = \gamma_{n0} + \gamma_{n1}\log\frac{s}{n} + \gamma_{n2}\log^2\frac{s}{n} + u_s,\ \gamma_n = A_n\beta. \tag{8}$$

In (8) the regressors are not asymptotically collinear and therefore the asymptotic distribution of the OLS estimator $\hat{\gamma}_n$ is good (normal and non-degenerate). The asymptotic distribution of $\hat{\beta}$ is extracted from $\hat{\beta} = A_n^{-1}\hat{\gamma}_n$. We call the γ_i's γ*ood coefficients* and β_i's β*ad coefficients*. The matrix A_n is called a *transition matrix*.

The problem with this transformation is that it is impossible to prove that (6) approximates (5). Therefore (Phillips, 2007, Theorem 5.1) is true for (6) and not for the original regression. Now we describe a modification of this approach that allows us to avoid dropping any terms. Consider two cases.

Table 2: Transition matrix summary

Case	Subcase	Coefficients
Non-reduced model ($\delta_1 \neq 0,\ \delta_2 \neq 0$)	($\delta_1/\delta_2 \to 0$, $\beta_2 = 0$) or ($\delta_2/\delta_1 \to 0$, $\beta_1 \neq 0$)	$a_{32} = \delta_1,\ a_{33} = 0$
	($\delta_1/\delta_2 \to 0$, $\beta_2 \neq 0$) or ($\delta_2/\delta_1 \to 0$, $\beta_1 = 0$)	$a_{32} = 0,\ a_{33} = \delta_2$
Semi-reduced model ($L_1(x) = \log x,\ \delta_2 \neq 0$)		$a_{32} = 0,\ a_{33} = \delta_2$

Non-reduced model. Let δ_1, δ_2 be nonzero. Then we write

$$\begin{aligned} L_j(s) &= L_j(n) + L_j(n)\varepsilon_j(n)\log\frac{s}{n} \\ &\quad + L_j(n)\varepsilon_j(n)\left(\frac{L_j(s) - L_j(n)}{L_j(n)\varepsilon_j(n)} - \log\frac{s}{n}\right) \\ &= L_j(n) + L_j(n)\varepsilon_j(n)\log\frac{s}{n} + \delta_j(n)H_j(s,n) \end{aligned} \tag{9}$$

where H_j is the H-function of L_j. H_j is not equal to $\log^2\frac{s}{n}$ but it is close to it in some sense (see Appendix). Substitution of (9) in (1) yields

$$y_s = \gamma_{n0} + \gamma_{n1}\log\frac{s}{n} + \Delta_n + u_s \tag{10}$$

where

$$\begin{aligned} \gamma_{n0} &= \beta_0 + \beta_1 L_1(n) + \beta_2 L_2(n), \\ \gamma_{n1} &= \beta_1 L_1(n)\varepsilon_1(n) + \beta_2 L_2(n)\varepsilon_2(n), \end{aligned} \tag{11}$$

and

$$\Delta_n = \beta_1\delta_1(n)H_1(s,n) + \beta_2\delta_2(n)H_2(s,n). \tag{12}$$

Semi-reduced model. In this case by definition $L_1(s) = \log s$, $\delta_1 = 0$, $\delta_2 \neq 0$. We can still apply (9) to L_2. For L_1 we use simply $L_1(s) = L_1(n) + (L_1(s) - L_1(n)) = L_1(n) + \log\frac{s}{n}$. Since $L_1\varepsilon_1 \equiv 1$, (11) is true and (12) formally holds with $\delta_1 = 0$.

From (11) we see that the first two rows of the transition matrix are the same as in (7). By (12) $a_{31} = 0$. The analysis in Appendix shows that the other two elements of the last row of A_n are as described in Table 2. The dependence of the transition matrix on the true β is not continuous. In all cases the conditions $a_{32}(n) = 0$ and $a_{33}(n) = 0$ are mutually exclusive and therefore the transition matrix is triangular, unlike (7).

2.3 Convergence statements

It is clear from the previous subsection that the asymptotic distribution of $\hat{\gamma}_n$ should be derived first and that of $\hat{\beta}$ next. In principle, convergence of the γood coefficients is described by (Phillips, 2007, Theorem 4.1), where they are denoted α_n. However, Theorem 4.1 depends on Phillips' Lemma 7.4, the proof of which is incomplete, and Lemma 2.1(iii), which can be proved under more general assumptions on the linear process u_t. Therefore we provide an independent proof.

Assumption 1 (on the regressors)

(a) In the non-reduced case we assume that $L_i = K(\varepsilon_i, \phi_{\varepsilon_i}, \theta_{\varepsilon_i})$, $\varepsilon_i = K(\eta_i, \phi_{\eta_i}, \theta_{\eta_i})$ and η_i is slowly varying for $i = 1, 2$. Further, we suppose that the μ-functions of L_i are different from 0 in some neighborhood of infinity and satisfy

$$\frac{1}{c} \max\{|\varepsilon_i(x)|, |\eta_i(x)|\} \leq |\mu_i(x)| \leq \max\{|\varepsilon_i(x)|, |\eta_i(x)|\} \tag{13}$$

with some constant $c > 0$. Finally, $\max\{2\theta_{\varepsilon_i}, \theta_{\eta_i}\} < 1/2$ for $i = 1, 2$.

(b) In the semi-reduced case $L_1(x) = \log x$ and L_2 satisfies part (a).

The next assumption is less restrictive than the corresponding condition by Phillips.

Assumption 2 (on the linear process) For all $t > 0$, u_t has representation $u_t = \sum_{j=-\infty}^{\infty} c_j e_{t-j}$, where

(a) the numbers c_j satisfy $\sum_{j=-\infty}^{\infty} |c_j| < \infty$, $\sum_{j=-\infty}^{\infty} c_j \neq 0$ and

(b) the sequence of random variables $\{e_j\}$ is a martingale difference sequence (e_t is $\mathcal{F}_t$-measurable and $(e_t|\mathcal{F}_{t-1}) = 0$) such that $E(e_t^2|\mathcal{F}_{t-1}) = \sigma_e^2$ (a constant) for all t and e_t^2 are uniformly integrable. Here $\{\mathcal{F}_t\}$ is an increasing sequence of σ-fields.

Theorem 2.1. *Let Assumptions 1 and 2 hold. Then*

$$\sqrt{n}(\hat{\gamma}_n - \gamma_n) \xrightarrow{d} N(0, \sigma^2 G^{-1}). \tag{14}$$

Henceforth we denote $\sigma^2 = \left(\sigma_e \sum_{j=-\infty}^{\infty} c_j\right)^2$ *and* G *is the Gram matrix of the system* $f_j(x) = \log^{j-1} x$, $j = 1, 2, 3,$ *that is, the element* g_{ij} *of* G *equals* $g_{ij} = \int_0^1 f_i(x) f_j(x) dx$.

To describe the behavior of the βad coefficients denote

$$\varepsilon_{\min} = \begin{cases} \varepsilon_1 & \text{if } \varepsilon_1/\varepsilon_2 \to \kappa \in \mathbb{R}; \\ \varepsilon_2 & \text{if } \varepsilon_2/\varepsilon_1 \to 0; \end{cases}$$

$$B_n = \sqrt{n} \begin{pmatrix} \varepsilon_{\min}(\hat{\beta}_0 - \beta_0) \\ L_1\varepsilon_1(\hat{\beta}_1 - \beta_1) \\ L_2\varepsilon_2(\hat{\beta}_2 - \beta_2) \end{pmatrix}, \quad f(\kappa) = \begin{pmatrix} \kappa - 1 \\ 1 \\ -1 \end{pmatrix}, \quad g = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}.$$

Table 3: Type-wise OLS asymptotics

Case	Subcase	$\varepsilon_1/\varepsilon_2 \to \kappa \in \mathbb{R}$	$\varepsilon_2/\varepsilon_1 \to 0$
Non-reduced model ($\delta_1 \neq 0,\ \delta_2 \neq 0$)	($\delta_1/\delta_2 \to 0,\ \beta_2 = 0$) or ($\delta_2/\delta_1 \to 0,\ \beta_1 \neq 0$)	$\mu_1 B_n \xrightarrow{d} f(\kappa)\Gamma$	$\mu_1 B_n \xrightarrow{d} g\Gamma$
	($\delta_1/\delta_2 \to 0,\ \beta_2 \neq 0$) or ($\delta_2/\delta_1 \to 0,\ \beta_1 = 0$)	$\mu_2 B_n \xrightarrow{d} f(\kappa)\Gamma$	$\mu_2 B_n \xrightarrow{d} g\Gamma$
Semi-reduced model ($L_1(x) = \log x,\ \delta_2 \neq 0$)		$\mu_2 B_n \xrightarrow{d} f(\kappa)\Gamma$	$\mu_2 B_n \xrightarrow{d} g\Gamma$

Theorem 2.2 (Classification theorem)**.** *Let Assumptions 1 and 2 hold and let* $\Gamma \sim N(0, \sigma^2/4)$. *Then the relation between the* β*ad coefficients* (*contained in* B_n) *and* γ*ood coefficients* (*represented by* Γ) *is presented in Table 3.*

Since (Phillips, 2007, Theorem 5.1) is actually about regression with a quadratic form in $\log(s/n)$, no wonder its predictions are different from those in Table 3. In particular, the classification theorem captures a new effect that the asymptotic variance depends on the true β. The scaling factor $\varepsilon_{\min}$ is the same as in the Phillips theorem. The case of more than two different SV regressors should present an even larger number of different asymptotic types. (Phillips, 2007, Theorem 5.2) does not cover all possibilities.

Example. The following example from (Phillips, 2007) has iterated logarithmic growth, a trend decay component, and a constant regressor:

$$y_s = \beta_0 + \beta_1/\log s + \beta_2 \log(\log s) + u_s.$$

Such a model is relevant in empirical research where one wants to capture simultaneously two different opposing trends in the data. Here $L_1(s) = 1/\log s$, $L_2(s) = \log(\log s)$.

From Table 1 $\varepsilon_1 = -1/l_1$, $\mu_1 = -1/l_1$, $\delta_1(n) = l_1^{-3}$, $\varepsilon_2 = 1/(l_1 l_2)$, $\mu_2 = -1/(2l_1)$, $\delta_2(n) = -1/(2l_1^2)$. Since $\delta_1/\delta_2 = -2/l_1 \to 0$ and $\varepsilon_2/\varepsilon_1 = -1/l_2 \to 0$, we have $\varepsilon_{\min} = \varepsilon_2$ and by Table 3

$$\frac{\sqrt{n}}{\log^2 n} \begin{pmatrix} \frac{1}{\log(\log n)}(\hat{\beta}_0 - \beta_0) \\ -\frac{1}{\log n}(\hat{\beta}_1 - \beta_1) \\ \hat{\beta}_2 - \beta_2 \end{pmatrix} \xrightarrow{d} \begin{cases} 2g\Gamma & \text{if } \beta_2 \neq 0; \\ g\Gamma & \text{if } \beta_2 = 0. \end{cases}$$

The formula from (Phillips, 2007, pp.575-576), after correction of two typos, gives the same asymptotics

$$\frac{\sqrt{n}}{\log^2 n} \begin{pmatrix} \frac{1}{\log(\log n)}(\hat{\beta}_0 - \beta_0) \\ -\frac{1}{\log n}(\hat{\beta}_1 - \beta_1) \\ \hat{\beta}_2 - \beta_2 \end{pmatrix} \xrightarrow{d} g\Gamma,$$

regardless of β_2. The comments by Phillips apply: The coefficient of the growth term converges fastest but at less that an $\sqrt{n}$ rate. The intercept converges next fastest, and finally the coefficient of the evaporating trend. All of these outcomes relate to the strength of the signal from the respective regressor.

Final remarks. The statement of (Phillips, 2007, Theorem 3.1) regarding simple regression $y_s = \alpha + \beta L(s) + u_s$ is true if $L = K(\varepsilon, \phi_\varepsilon, \theta_\varepsilon)$ with $2\theta_\varepsilon < 1$ and u_t satisfies Assumption 2. The statements of (Phillips, 2007, Theorems 4.2, 4.3) regarding $y_s = \sum_{j=0}^{p} \beta_j L^j(s) + u_s$ (a polynomial regression in $L(s)$) are true if $L = K(\varepsilon, \phi_\varepsilon, \theta_\varepsilon)$ with $2\theta_\varepsilon p < 1$ and u_t satisfies Assumption 2. The proofs are obtained by using the central limit results contained in this paper or (Mynbaev, 2009), as appropriate.

A Proofs of Results

A.1 Bounds for first- and second-order regular variation

Lemma A.1. *If $L = K(\varepsilon, \phi_\varepsilon, \theta_\varepsilon)$, then for any $b > \theta_\varepsilon$ there exist numbers $M_b > 0$ and $a_b \geq \max\{a, c\}$ such that $|L(\lambda x)/L(x) - 1| \leq M_b \lambda^{-b}/\phi_\varepsilon(x)$ for all $x \geq a_b$ and $a_b/x \leq \lambda \leq 1$.*

This lemma is a special case of (Seneta, 1985, Lemma A.1.1). For the proof see also (Mynbaev, 2009). Since in practical cases the number θ_ε can be arbitrarily close to 0, the number $b > \theta_\varepsilon$ can also be as close to 0 as desired. Denote $G(t, n) = (L(t) - L(n))/(L(n)\varepsilon(n))$.

Lemma A.2. *If $L = K(\varepsilon, \phi_\varepsilon, \theta_\varepsilon)$ and $\varepsilon = K(\eta, \phi_\eta, \theta_\eta)$, then for any $b > \max\{2\theta_\varepsilon, \theta_\eta\}$ there exist constants $M_b > 0$ and $a_b \geq \max\{a, c\}$ such that*

$$|G(\lambda x, x) - \log \lambda| \leq M_b \lambda^{-b} \left(\frac{1}{\phi_\varepsilon(x)} + \frac{1}{\phi_\eta(x)} \right) \text{ for } x \geq a_b \text{ and } \frac{a_b}{x} \leq \lambda \leq 1.$$

Proof. Denote $r(\lambda, x) = L(\lambda x)/L(x)$, $U(\lambda, x) = \log r(\lambda, x)$. Let $x \geq c$ and $c/x \leq \lambda \leq 1$, where c is the constant from (3). Since $\lambda x \leq x$, (2) implies

$$U(\lambda, x) = -\int_{\lambda x}^{x} \varepsilon(t) \frac{dt}{t}. \tag{15}$$

Using the right inequality from (3) and the fact that ϕ_ε is nondecreasing we get

$$|U(\lambda, x)| \leq \int_{\lambda x}^{x} |\varepsilon(t)| \frac{dt}{t} \leq c \int_{\lambda x}^{x} \frac{1}{\phi_\varepsilon(t)} \frac{dt}{t} \leq \frac{c}{\phi_\varepsilon(\lambda x)} \int_{\lambda x}^{x} \frac{dt}{t} = -\frac{c \log \lambda}{\phi_\varepsilon(\lambda x)}. \tag{16}$$

Fix some $b_\varepsilon > \theta_\varepsilon$. Using monotonicity of ϕ_ε and the fact that it increases to ∞ at ∞, from $a_b \leq x\lambda$ we have $c/\phi_\varepsilon(\lambda x) \leq c/\phi_\varepsilon(a_b) < (b_\varepsilon - \theta_\varepsilon)/2$ for a sufficiently large $a_b > 0$. Then by (16)

$$|U(\lambda, x)| \leq -\frac{b_\varepsilon - \theta_\varepsilon}{2} \log \lambda. \tag{17}$$

On the other hand, by part (b) of the definition of the class $K(\varepsilon, \phi_\varepsilon, \theta_\varepsilon)$ the inequality $X_\varepsilon \le \lambda x \le x$ implies $(\lambda x)^{-\theta_\varepsilon}\phi_\varepsilon(\lambda x) \ge x^{-\theta_\varepsilon}\phi_\varepsilon(x)$ and $1/\phi_\varepsilon(\lambda x) \le \lambda^{-\theta_\varepsilon}/\phi_\varepsilon(x)$. Hence, from (16)

$$|U(\lambda, x)| \le -c\lambda^{-\theta_\varepsilon}(\log \lambda)/\phi_\varepsilon(x). \tag{18}$$

Now consider

$$r(\lambda, x) - 1 - \varepsilon(x)\log\lambda = e^{U(\lambda,x)} - 1 - U(\lambda, x) + U(\lambda, x) - \varepsilon(x)\log\lambda. \tag{19}$$

By Lemma A.1 applied to ε

$$|\varepsilon(\lambda x)/\varepsilon(x) - 1| \le c_1\lambda^{-b_\eta}/\phi_\eta(x) \text{ for all } x \ge a_b \text{ and } a_b/x \le \lambda \le 1 \tag{20}$$

where b_η is an arbitrary number $> \theta_\eta$ and c_1 depends on b_η. From (15) we have

$$\begin{aligned} |U(\lambda, x) - \varepsilon(x)\log\lambda| &= \left| -\int_{\lambda x}^{x} \varepsilon(t)\frac{dt}{t} + \varepsilon(x)\int_{\lambda x}^{x}\frac{dt}{t}\right| \\ = \left|\varepsilon(x)\int_{\lambda x}^{x}\left(\frac{\varepsilon(t)}{\varepsilon(x)} - 1\right)\frac{dt}{t}\right| &\le |\varepsilon(x)|\int_{\lambda}^{1}\left|\frac{\varepsilon(sx)}{\varepsilon(x)} - 1\right|\frac{ds}{s}. \end{aligned}$$

The conditions $a_b \le \lambda x$ and $\lambda \le s \le 1$ imply $a_b \le sx \le x$, so we can use (20) to get

$$\begin{aligned} |U(\lambda, x) - \varepsilon(x)\log\lambda| &\le \frac{c_1|\varepsilon(x)|}{\phi_\eta(x)}\int_{\lambda}^{1} s^{-b_\eta - 1}ds = \frac{c_2|\varepsilon(x)|}{\phi_\eta(x)}(\lambda^{-b_\eta} - 1) \\ &\le \frac{c_2|\varepsilon(x)|}{\phi_\eta(x)}\lambda^{-b_\eta} \text{ for } x \ge a_b \text{ and } \frac{a_b}{x} \le \lambda \le 1. \end{aligned} \tag{21}$$

Applying bounds (17) and (18) and an elementary inequality $|e^x - 1 - x| \le x^2 e^{|x|}$ we obtain

$$|e^{U(\lambda,x)} - 1 - U(\lambda, x)| \le U^2(\lambda, x)e^{|U(\lambda,x)|} \le c_3\frac{\lambda^{-2\theta_\varepsilon}\log^2\lambda}{\phi_\varepsilon^2(x)}\lambda^{-\frac{1}{2}(b_\varepsilon - \theta_\varepsilon)} \tag{22}$$

where $b_\varepsilon > \theta_\varepsilon$. Combining (19), (21) and (22) gives

$$\left|\frac{L(\lambda x)}{L(x)} - 1 - \varepsilon(x)\log\lambda\right| \le c_4\left[\frac{|\varepsilon(x)|}{\phi_\eta(x)}\lambda^{-b_\eta} + \frac{\log^2\lambda}{\phi_\varepsilon^2(x)}\lambda^{-\frac{1}{2}b_\varepsilon - \frac{3}{2}\theta_\varepsilon}\right]. \tag{23}$$

On the interval $(0, 1]$ the function $\log^2\lambda$ can be dominated by $c(\delta)\lambda^{-\delta}$ with any $\delta > 0$. Since the number $b_\varepsilon > \theta_\varepsilon$ is arbitrarily close to θ_ε, the number $a_\varepsilon \equiv \frac{1}{2}b_\varepsilon + \frac{3}{2}\theta_\varepsilon + \delta$ is larger than, and arbitrarily close to, $2\theta_\varepsilon$. Hence, the left inequality in (3) and (23) imply

$$|G(\lambda x, x) - \log\lambda| \le c_5\left(\frac{\lambda^{-b_\eta}}{\phi_\eta(x)} + \frac{\lambda^{-a_\varepsilon}}{\phi_\varepsilon^2(x)|\varepsilon(x)|}\right) \le c_6\left(\frac{1}{\phi_\eta(x)} + \frac{1}{\phi_\varepsilon(x)}\right)\lambda^{-b}.$$

Taking $b > \max\{2\theta_\varepsilon, \theta_\eta\}$ and putting $b_\eta = a_\varepsilon = b$ we satisfy both $b_\eta > \theta_\eta$ and $a_\varepsilon > 2\theta_\varepsilon$. The constant c_6 depends on b. $\square$

A.2 L_p-approximability

Here we prove a stronger version of the second-order regular variation (4) that uses the notion of L_p-approximability from (Mynbaev, 2001). Let $\{w_n\}$ be a sequence of vectors such that $w_n \in \mathbb{R}^n$ for each n and denote $\|f\|_p = \left(\int_0^1 |f(x)|^p dx\right)^{1/p}$, where $p < \infty$. Let Δ_{np} denote an *interpolation operator* defined by $\Delta_{np} w_n = n^{1/p} \sum_{t=1}^{n} w_{nt} 1_{i_t}$. Here w_{nt} are the coordinates of w_n; the intervals $i_t = [(t-1)/n, t/n),\ t = 1, ..., n$, form a partition of $[0, 1)$ and 1_A is an indicator of a set A, that is $1_A = 1$ on A and $1_A = 0$ outside A. We say that $\{w_n\}$ is *L_p-approximable* if there exists a function W on $[0, 1]$ such that $\|W\|_p < \infty$ and $\|\Delta_{np} w_n - W\|_p \to 0$. In this case we also say that $\{w_n\}$ is *L_p-close to W*.

Lemma A.3. *Suppose that $L = K(\varepsilon, \phi_\varepsilon, \theta_\varepsilon)$, $\varepsilon = K(\eta, \phi_\eta, \theta_\eta)$ and η is slowly varying. Assume, further, that the μ-function of L is different from zero for all large x and satisfies the condition of type* (13) *with some constant $c > 0$. For $p \in [1, \infty)$ define a vector $w_n \in \mathbb{R}^n$ by $w_{nt} = n^{-1/p} H(t, n)$, $t = 1, ..., n$. If $\max\{2\theta_\varepsilon, \theta_\eta\} < 1/p$, then $\{w_n\}$ is L_p-close to $f(x) = \log^2 x$.*

Proof. The definitions of w_n and Δ_{np} give $\Delta_{np} w_n = \sum_{t=1}^{n} H(t, n) 1_{i_t}$. This is equivalent to n equations $(\Delta_{np} w_n)(u) = H(t, n)$ for $u \in i_t,\ t = 1, ..., n$. 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 $t = [nu + 1]$. Hence, the above n equations take a compact form $(\Delta_{np} w_n)(u) = H([nu + 1], n),\ 0 \le u < 1$.

To reflect dependence on the domain of integration, denote $\|f\|_{p,(a,b)} = \left(\int_a^b |f(x)|^p dx\right)^{1/p}$. Let $0 < \delta \le 1/2$ and with the number a_b from Lemma A.2 put $n_1 \equiv a_b/\delta$. For $n > n_1$ the interval $(a_b/n, \delta)$ is not empty and by the triangle inequality

$$\begin{aligned} \|\Delta_{np} w_n - f\|_{p,(0,1)} \le\ & \|\Delta_{np} w_n - f\|_{p,(\delta,1)} + \|f\|_{p,(0,\delta)} \\ & + \|\Delta_{np} w_n\|_{p,(0,a_b/n)} + \|\Delta_{np} w_n\|_{p,(a_b/n,\delta)}. \end{aligned} \tag{24}$$

Since $|f|^p$ is integrable on $(0, 1)$, we have $\|f\|_{p,(0,\delta)} \to 0$ as $\delta \to 0$. For the other three terms at the right of (24) we consider three cases.

Case $\delta \le u < 1$. Under the conditions of this lemma (Phillips, 2007, Equation (63)) is true and implies (4). Therefore

$$H(rn, n) = [1 + o(1)] \log^2 r \text{ uniformly in } r \in \left[\delta, 1 + \frac{1}{2a_b}\right]. \tag{25}$$

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

$$\delta \le u < r \le u + 1/n < 1 + 1/n_1 \le 1 + 1/(2a_b). \tag{26}$$

This leads to $r = u + o(1)$ and $r \in [\delta, 1 + 1/(2a_b)]$. From these equations and (25) we see that $H([nu +$

$1],n) - \log^2 u = o(1)$ uniformly in $u \in [\delta, 1)$ which allows us to conclude that $\|\Delta_{np} w_n - f\|_{p,(\delta,1)} \to 0,\ n \to \infty$.

Case $a_b/n \le u < \delta$. Let $n > n_2 \equiv \max\{n_1, 2\}$. Then (26) and the conditions $u \in [a_b/n, \delta)$, $n > n_2$ imply $a_b/n \le u < r \le u + 1/n < \delta + 1/n_2 \le 1$. This means we can successively apply Lemma A.2, (3), condition (13) and (26) to get

$$
\begin{aligned}
|H([nu+1],n)| &= \left|\frac{G(rn,n) - \log r}{\mu(n)}\right| \le \frac{M_b r^{-b}}{|\mu(n)|}\left(\frac{1}{\phi_\eta(n)} + \frac{1}{\phi_\varepsilon(n)}\right) \\
&\le \frac{c_1 r^{-b}}{|\mu(n)|}\max\{|\varepsilon(n)|, |\eta(n)|\} \le c_2 r^{-b} \le c_2 u^{-b} \text{ for } u \in \left[\frac{a_b}{n}, \delta\right]
\end{aligned}
$$

where $b > \max\{2\theta_\varepsilon, \theta_\eta\}$. Hence, $\int_{a_b/n}^{\delta} |\Delta_{np} w_n|^p du \le c_3 \delta^{1-pb}$. Here the right-hand side tends to zero if $b < 1/p$. This is possible because of $\max\{2\theta_\varepsilon, \theta_\eta\} < 1/p$.

Case $0 < u < a_b/n$. By monotonicity the inequality $[nu+1]/n > u$ implies $|\log([nu+1]/n)| \le |\log u|$. On the other hand, $[nu+1] \le nu + 1 < a_b + 1$ and $L([nu+1]) \le c$ by continuity of L. Hence, $|G([nu+1],n)| \le (c/L(n) + 1)/|\varepsilon(n)|$ and

$$
\begin{aligned}
|H([nu+1],n)| &\le \left|\frac{G([nu+1],n)}{\mu(n)}\right| + \left|\frac{\log([nu+1]/n)}{\mu(n)}\right| \\
&\le \left|\frac{c + L(n)}{L(n)\varepsilon(n)\mu(n)}\right| + \left|\frac{\log u}{\mu(n)}\right|.
\end{aligned}
$$

All functions of n here are slowly varying and $|\log u|$ can be dominated by cu^{-a} with $0 < a < 1/p$. Therefore

$$
\|\Delta_{np} w_n\|_{p,(0,a_b/n)} \le \left|\frac{c + L(n)}{L(n)\varepsilon(n)\mu(n)}\right| \left(\frac{a_b}{n}\right)^{1/p} + \frac{c}{|\mu(n)|}\left(\frac{a_b}{n}\right)^{1/p-a} \to 0.
$$

This tends to zero as $n \to \infty$ because (a) sums, products and real powers of SV functions are SV and (b) for any $a > 0$ and any SV function L, the product $n^{-a}L(n)$ tends to zero as $n \to \infty$. □

Lemma A.4. *Under Assumptions 1 and 2 the last row of the transition matrix is described by Table 2.*

Proof. We shall need the following *linearity property* of L_p-approximable sequences: if $\{w_n\}$ is L_p-close to W, $\{v_n\}$ is L_p-close to V, $\{a_n\}$ and $\{b_n\}$ are numerical sequences converging to a and b, respectively, then $\{a_n v_n + b_n w_n\}$ is L_p-close to $aV + bW$. This follows from

$$
\begin{aligned}
&\|\Delta_{np}(a_n v_n + b_n w_n) - (aV + bW)\|_p \le |a_n - a|\,\|\Delta_{np} v_n\|_p \\
&+ |b_n - b|\,\|\Delta_{np} w_n\|_p + |a|\,\|\Delta_{np} v_n - V\|_p + |b|\,\|\Delta_{np} w_n - W\|_p \to 0
\end{aligned}
$$

where, by L_p-approximability, $\|\Delta_{np} v_n\|_p$ and $\|\Delta_{np} w_n\|_p$ are bounded.

Non-reduced model. The functions L_1 and L_2 satisfy part (a) of Assumption 1. By Lemma A.3 the sequences w_n^1, w_n^2 with components $w_{nt}^i = n^{-1/2} H_i(t,n)$, $t = 1, ..., n$, $i = 1, 2$, are L_2-close to $\log^2 x$. By linearity then $\{a_n w_n^1 + b_n w_n^2\}$ is L_2-close to $(a+b)\log^2 x$ whenever $a_n \to a,\ b_n \to b$.

Subcase $\delta_1/\delta_2 \to 0$. **(a)** Let $\beta_2 = 0$. By (12) $\Delta_n = \beta_1\delta_1 H_1(s,n)$, and we put

$$\gamma_{n2} = \beta_1\delta_1,\ \widetilde{H}(s,n) = H_1(s,n). \tag{27}$$

With this definition γ_{n2} is linear in β_i and

$$\Delta_n = \gamma_{n2}\widetilde{H}(s,n) \text{ and } \left\{n^{-1/2}\widetilde{H}(s,n)\right\} \text{ is } L_2\text{-close to } \log^2 x. \tag{28}$$

Definition (27) gives the corresponding cell in Table 2 ($a_{32} = \delta_1,\ a_{33} = 0$). (28) is the leading idea in this and subsequent definitions: with (28), (10) becomes

$$y_s = \gamma_{n0} + \gamma_{n1}\log\frac{s}{n} + \gamma_{n2}\widetilde{H}(s,n) + u_s \tag{29}$$

which is a realization of Phillips' idea (8).

(b) Let $\beta_2 \neq 0$. By (12) $\Delta_n = \beta_2\delta_2(H_1(s,n)\beta_1\delta_1/\beta_2\delta_2 + H_2(s,n))$. This suggests defining $\gamma_{n2} = \beta_2\delta_2$, $\widetilde{H}(s,n) = H_1(s,n)\beta_1\delta_1/\beta_2\delta_2 + H_2(s,n)$ which gives (28) and the corresponding definition in Table 2.

Subcase $\delta_2/\delta_1 \to 0$. **(a)** Let $\beta_1 = 0$. The choice $\gamma_{n2} = \beta_2\delta_2,\ \widetilde{H}(s,n) = H_2(s,n)$ obviously satisfies (28) and gives $a_{32} = 0,\ a_{33} = \delta_2$.

(b) If $\beta_1 \neq 0$ we define $\gamma_{n2} = \beta_1\delta_1,\ \widetilde{H}(s,n) = H_1(s,n) + H_2(s,n)\beta_2\delta_2/\beta_1\delta_1$ to satisfy (28) and $a_{32} = \delta_1,\ a_{33} = 0$.

In case of the semi-reduced model we have $\Delta_n = \beta_2\delta_2 H_2(s,n)$ and the choice is obvious: $\gamma_{n2} = \beta_2\delta_2,\ \widetilde{H}(s,n) = H_2(s,n)$. □

A.3 Proof of Theorem 2.1

Non-reduced model. Denote

$$X_n = \begin{pmatrix} 1 & \log(1/n) & \widetilde{H}(1,n) \\ ... & ... & ... \\ 1 & \log(n/n) & \widetilde{H}(n,n) \end{pmatrix}$$

the matrix of regressors in (29). We know from (28) that the third column of $W_n \equiv n^{-1/2}X_n$ is L_2-close to f_3. The first column of this matrix, $w_n = n^{-1/2}(1,...,1)'$, is L_2-close to f_1 because $\Delta_{n2}w_n$ is identically 1 on $(0,1)$. Letting $p = 2$, $k = 1$ in (Mynbaev, 2009, Theorem 3) we see that the second column, $w_n = n^{-1/2}(\log(1/n),...,\log(n/n))'$ is L_2-close to f_2. By (Mynbaev, 2001, Theorems 3.1(b) and 4.1(D)) we have $W_n'u^{(n)} \xrightarrow{d} N(0,\sigma^2 G),\ W_n'W_n \to G$ where $u^{(n)} = (u_1,...,u_n)'$. Now (14) follows from $\sqrt{n}(\hat{\gamma}_n - \gamma_n) = (W_n'W_n)^{-1}W_n'u^{(n)}$.

Semi-reduced model. In this case the situation is simpler because the first and second columns of X_n are the same, whereas $\widetilde{H}(s,n) = H_2(s,n)$. □

A.4 Proof of Theorem 2.2

We need the following well-known fact. Let A_n be a nonsingular matrix. If the parameter vector β in the linear model $y = X\beta + u$ has been transformed as $\gamma_n = A_n\beta$, to obtain $y = XA_n^{-1}\gamma_n + u$, then

$$\hat{\gamma}_n - \gamma_n = A_n(\hat{\beta} - \beta). \tag{30}$$

It turns out that only $\hat{\gamma}_{n2}$ affects the limit distribution of $\hat{\beta}$. By (Phillips, 2007, Lemma 7.8(iii)) the element g^{33} in the lower right corner of G^{-1} equals 1/4. Therefore Theorem 2.1 implies

$$\sqrt{n}(\hat{\gamma}_{n2} - \gamma_{n2}) \xrightarrow{d} N(0, \sigma^2/4). \tag{31}$$

Case $a_{33}(n) = 0$. By (11) and Table 2 the system $A_n\beta = \gamma_n$ takes the form

$$\begin{cases} \beta_0 + L_1\beta_1 + L_2\beta_2 = \gamma_{n0}, \\ L_1\varepsilon_1\beta_1 + L_2\varepsilon_2\beta_2 = \gamma_{n1}, \\ L_1\varepsilon_1\mu_1\beta_1 = \gamma_{n2}. \end{cases} \tag{32}$$

It is easy to check that

$$A_n^{-1} = \begin{pmatrix} 1 & -\frac{1}{\varepsilon_2} & \frac{1}{\mu_1}\left(\frac{1}{\varepsilon_2} - \frac{1}{\varepsilon_1}\right) \\ 0 & 0 & \frac{1}{L_1\varepsilon_1\mu_1} \\ 0 & \frac{1}{L_2\varepsilon_2} & -\frac{1}{L_2\varepsilon_2\mu_1} \end{pmatrix}. \tag{33}$$

Subcase $\varepsilon_1/\varepsilon_2 \to \kappa \in \mathbb{R}$. Note that

$$\mathrm{diag}[\varepsilon_1\mu_1, L_1\varepsilon_1\mu_1, L_2\varepsilon_2\mu_1]A_n^{-1} = \begin{pmatrix} \mu_1\varepsilon_1 & -\mu_1\frac{\varepsilon_1}{\varepsilon_2} & \frac{\varepsilon_1}{\varepsilon_2} - 1 \\ 0 & 0 & 1 \\ 0 & \mu_1 & -1 \end{pmatrix}. \tag{34}$$

Denoting

$$B_n^{(i)} = \sqrt{n} \begin{pmatrix} \varepsilon_i(\hat{\beta}_0 - \beta_0) \\ L_1\varepsilon_1(\hat{\beta}_1 - \beta_1) \\ L_2\varepsilon_2(\hat{\beta}_2 - \beta_2) \end{pmatrix}, \; i = 1, 2,$$

from (30), (33), (34) we have

$$\begin{aligned} \mu_1 B_n &= \mu_1 B_n^{(1)} = \sqrt{n}\mathrm{diag}[\varepsilon_1\mu_1, L_1\varepsilon_1\mu_1, L_2\varepsilon_2\mu_1](\hat{\beta} - \beta) \\ &= \begin{pmatrix} \mu_1\varepsilon_1 & -\mu_1\frac{\varepsilon_1}{\varepsilon_2} & \frac{\varepsilon_1}{\varepsilon_2} - 1 \\ 0 & 0 & 1 \\ 0 & \mu_1 & -1 \end{pmatrix} \sqrt{n}(\hat{\gamma}_n - \gamma_n). \end{aligned}$$

Now take into account that ε_1, η_1 and μ_1 vanish at infinity by the Karamata theorem, that $\varepsilon_1/\varepsilon_2 \to \kappa$ by assumption and that $\sqrt{n}(\hat{\gamma}_n - \gamma_n)$ converges in distribution by Theorem 2.1. Then the preceding equation and (31) imply

$$\mu_1 B_n = f(\kappa)\sqrt{n}(\hat{\gamma}_{n2} - \gamma_{n2}) + o_p(1) \xrightarrow{d} f(\kappa)\Gamma. \tag{35}$$

In the other cases the argument is similar, and we indicate only the relevant analogs of (33), (34) and (35).

Subcase $\varepsilon_2/\varepsilon_1 \to 0$. The equation $\mu_1 B_n = g\sqrt{n}(\hat{\gamma}_{n2} - \gamma_{n2}) + o_p(1) \xrightarrow{d} g\Gamma$ is obtained using

$$\mathrm{diag}[\varepsilon_2\mu_1, L_1\varepsilon_1\mu_1, L_2\varepsilon_2\mu_1]A_n^{-1} = \begin{pmatrix} \mu_1\varepsilon_2 & -\mu_1 & 1-\frac{\varepsilon_2}{\varepsilon_1} \\ 0 & 0 & 1 \\ 0 & \mu_1 & -1 \end{pmatrix}.$$

Case $a_{32}(n) = 0$. The first two equations in (32) do not change, and instead of the third one we have $L_2\varepsilon_2\mu_2\beta_2 = \gamma_{n2}$. Therefore

$$A_n^{-1} = \begin{pmatrix} 1 & -\frac{1}{\varepsilon_1} & \frac{1}{\mu_2}\left(\frac{1}{\varepsilon_1} - \frac{1}{\varepsilon_2}\right) \\ 0 & \frac{1}{L_1\varepsilon_1} & -\frac{1}{L_1\varepsilon_1\mu_2} \\ 0 & 0 & \frac{1}{L_2\varepsilon_2\mu_2} \end{pmatrix}.$$

Subcase $\varepsilon_1/\varepsilon_2 \to \kappa$. The relation $\mu_2 B_n \xrightarrow{d} -f(\kappa)\Gamma$ follows from

$$\mathrm{diag}[\varepsilon_1\mu_2, L_1\varepsilon_1\mu_2, L_2\varepsilon_2\mu_2]A_n^{-1} = \begin{pmatrix} \varepsilon_1\mu_2 & -\mu_2 & 1-\frac{\varepsilon_1}{\varepsilon_2} \\ 0 & \mu_2 & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Subcase $\varepsilon_2/\varepsilon_1 \to 0$. To prove that $\mu_2 B_n \xrightarrow{d} -g\Gamma$ we apply

$$\mathrm{diag}[\varepsilon_2\mu_2, L_1\varepsilon_1\mu_2, L_2\varepsilon_2\mu_2]A_n^{-1} = \begin{pmatrix} \varepsilon_2\mu_2 & -\mu_2\frac{\varepsilon_2}{\varepsilon_1} & \frac{\varepsilon_2}{\varepsilon_1} - 1 \\ 0 & \mu_2 & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$

A laborious comparison of the equations obtained and Table 2 allows one to fill out Table 3. Note that, because of Assumption 1, in all cases the coefficient μ_i in front of B_n is nonzero. □

References

Aljančić, S., Bojanić, R. and Tomić, M. (1955) Deux théorèmes relatifs au comportement asymptotique des séries trigonométriques. Srpska Akad. Nauka. Zb. Rad. Mat. Inst. 43, 15–26.

Barro, R.J. and Sala-i Martin, X. (2003) Economic Growth. MIT Press. 2nd edition.

Bingham, N. H., Goldie, C. M., and Teugels, J. L. (1987) Regular variation. Encyclopedia of Mathematics and its Applications, vol. 27. Cambridge: Cambridge University Press.

Hurvich, C. M., Deo, R., and Brodsky, J. (1998) The mean squared error of Geweke and Porter-Hudak's estimator of the memory parameter of a long-memory time series. *J. Time Ser. Anal.* 19, 19–46.

Mynbaev, K. T. (2001) L_p-approximable sequences of vectors and limit distribution of quadratic forms of random variables. *Adv. Appl. Math.* 26, 302–329.

Mynbaev, K. T. (2009) Central limit theorems for weighted sums of linear processes: L_p-approximability versus Brownian motion. *Econometric Theory* 25, 748–763.

Phillips, P. C. B. (1999) Discrete Fourier transforms of fractional processes. Cowles Foundation discussion paper no. 1243, Yale University.

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

Robinson, P. M. (1995) Log-periodogram regression of time series with long range dependence. *Ann. Statist.* 23, 1048–1072.

Seneta, E. (1985) Pravilno menyayushchiesya funktsii. Moscow: Nauka. Translated from English by I. S. Shiganov, translation edited and with a preface by V. M. Zolotarev, with appendices by I. S. Shiganov and V. M. Zolotarev.

Wu, Chien-Fu. (1981) Asymptotic theory of nonlinear least squares estimation. *Ann. Statist.* 9, 501–513.