

Global COE Hi-Stat Discussion Paper Series 208

**Research Unit for Statistical
and Empirical Analysis in Social Sciences (Hi-Stat)**

**Asymptotic Efficiency of the OLS Estimator with
Singular Limiting Sample Moment Matrices**

Yoshimasa Uematsu

October 2011

Hi-Stat

Institute of Economic Research
Hitotsubashi University
2-1 Naka, Kunitatchi Tokyo, 186-8601 Japan
<http://gcoe.ier.hit-u.ac.jp>

Asymptotic Efficiency of the OLS Estimator with Singular Limiting Sample Moment Matrices

Yoshimasa Uematsu*

Graduate School of Economics, Hitotsubashi University

Abstract

This paper presents a time series model that has an asymptotically efficient ordinary least squares (OLS) estimator, irrespective of the singularity of its limiting sample moment matrices. In the literature on stationary time series analysis, Grenander and Rosenblatt's (1957) (G–R) classical result is used to judge the asymptotic efficiency of regression coefficients on deterministic regressors satisfying Grenander's condition. Without this condition, however, it is not obvious that the model is efficient. In this paper, we introduce such a model by proving the efficiency of the model with a slowly varying (SV) regressor under the same condition on error terms constrained in G–R. This kind of regressor is known to display asymptotic singularity in the sample moment matrices, as in Phillips (2007), such that Grenander's condition fails.

*The author thanks Eiji Kurozumi for constructive remarks on an earlier version of this paper and Katsumi Shimotsu for helpful suggestions in refining the current version. The author also appreciates the useful comments of workshop participants at Hitotsubashi University along with delegates at the 2011 Japanese Joint Statistical Meeting held at Kyushu University and gratefully acknowledges the financial support of a Japan Society for the Promotion of Science Research Fellowship for Young Scientists. Address correspondence to Yoshimasa Uematsu, Graduate School of Economics, Hitotsubashi University, 2-1 Naka, Kunitachi, Tokyo 186-8601, Japan; e-mail: ed111001@g.hit-u.ac.jp.

1 INTRODUCTION

Discussion of the asymptotic efficiency of regression coefficients on time series models dates from the middle of the 20th century. When the regressors are deterministic functions of time such that the disturbances may be serially autocorrelated, Grenander (1954), Rosenblatt (1956) and Grenander and Rosenblatt (1957) (G–R) have found the necessary and sufficient condition for ordinary least squares (OLS) estimators to be asymptotically efficient. In this situation, a theory is constructed for the class of regressors that satisfy the so-called Grenander’s condition. However, while these regressors may be sufficiently general in empirical work or in the theoretical analysis of deterministically trending models, such as in Vogelsang (1998) and Perron and Yabu (2009), theoretical studies of asymptotic efficiency without this condition have rarely been seen.

Following the seminal work of G–R, there has been much attention to models with stochastic regressors in particular. For example, Krämer (1986) has proved the asymptotic equivalence of the OLS and the GLS estimator when the regressor is a univariate integrated process under stationary errors independent of the regressor. Subsequently, Phillips and Park (1988) extended these results when dealing with multiple regressions. Most recently, Krämer and Hassler (1998) studied the case where the regressors are fractionally integrated, while Shin and Oh (2002) generalized the class of integrated regressors to the class of unstable regressors containing seasonally integrated processes including high order integrations as special cases.

The present paper, however, reconsiders a classical G–R type regression, that is, a model with deterministic regressors. In this class of model, it is unclear whether the model is efficient unless Grenander’s condition is satisfied and the G–R result applied. This paper shows the existence of an asymptotically efficient model that does not satisfy the condition through proving the asymptotic efficiency of the model with a slowly varying (SV) regressor. That these kinds of regressors are asymptotically collinear with the constant term and the other SV regressors is owed to Phillips (2007). As a consequence, the asymptotic singularity of sample covariance matrix of regressors

arise, that is, the Grenander condition fails. Thus, the efficiency of the model with an SV regressor yields a simple but significant example of time series efficiency that does not satisfy the Grenander condition. Some results on SV regressors investigated by Phillips (2007) are fundamental to our theory, and hence are reviewed in the following section.

Before proceeding to the general result, we study the regression model with a certain error process $\{u_t\}$ as preparation.

$$y_t = \alpha + \beta \log t + u_t, \quad \text{or} \quad y = X\beta + u, \quad \text{say}$$

Note that the logarithmic function is a typical example of SV functions. If we let $D_n = \text{diag}[\sqrt{n}, \sqrt{n} \log n]$ and $F_n^{-1} = \text{diag}[\sqrt{n} \log^{-1} n, \sqrt{n}]$, then the sample moment matrices of the regressors $X_t = [1, \log t]$ and $X = [X'_1, \dots, X'_n]'$ asymptotically behave so that

$$D_n^{-1}(X'X)D_n^{-1} \rightarrow \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad F_n^{-1}(X'X)^{-1}F_n^{-1} \rightarrow \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}. \quad (1)$$

This is easily derived by a direct application of Lemma 1 described in the next section. Therefore, we may confirm that the classical approach in G–R is no longer valid. Our proof of the asymptotic equivalence of the OLS and GLS estimators is intuitively understood because we prove later that

$$X'\Gamma X \sim \sigma^2 X'X \quad \text{and} \quad X'\Gamma^{-1}X \sim \frac{1}{\sigma^2} X'X \quad (2)$$

holds, then we have the asymptotic equivalence of the variances,

$$(X'X)^{-1}X'\Gamma X(X'X)^{-1} \sim \sigma^2(X'X)^{-1} \quad \text{and} \quad (X'\Gamma^{-1}X)^{-1} \sim \sigma^2(X'X)^{-1},$$

where $\Gamma = \text{Var}(u)$ and σ^2 is the long-run variance of u if it is defined as a stationary process. A difficulty arises when we take the determinant for inverse, i.e.,

$$\det(X'X) = n \sum_{t=1}^n \log^2 t - \left(\sum_{t=1}^n \log t \right)^2.$$

The first and second asymptotically dominated terms of the difference offset each other and only the third dominated term survives. (See (5).) Therefore, a rather accurate approximation is required for (2).

In a regression model with an SV regressor, assumptions concerning the error terms are also important. Phillips (2007) and Mynbaev (2009) have defined $\{u_t\}$ as a linear process with some moment conditions to derive the asymptotic distribution of the OLS estimator. (See Remark 3.) Conversely, as we focus only on the second moments of the OLS and GLS estimators, a general stationary process with continuous and positive spectrum is required and is completely identical to the classical requirement in G–R. (See Section 2.1.)

The remainder of the paper is organized as follows: Section 2 includes some assumptions and provides some preliminary theory. We especially consider reviews of the G–R result and Phillips (2007) on SV regressors. Section 3 states the main theorem for asymptotic efficiency and Section 4 concludes. Appendixes A and B include the proofs for the main result.

2 PRELIMINARY RESULTS AND ASSUMPTIONS

2.1 Classics for Asymptotic Efficiency

Consider a time series regression model as follows:

$$y_t = \beta_0 x_{0,t} + \beta_1 x_{1,t} + u_t, \quad t = 1, \dots, n, \quad (3)$$

where $x_{i,t}$ for $i = 0, 1$ are nonstochastic regressors and u_t are stationary errors satisfying Assumption 2 below. We write $a_{ij,n}(h) = \sum_{t=1}^{n-h} x_{i,t+h} x_{j,t}$, a product sum of $x_{i,t}$'s for $t = 1, \dots, n$ and $h = 0, 1, \dots$. Further, we define the correlation matrix $R_n(h) = D_n^{-1} A_n(h) D_n^{-1}$ with $D_n = \text{diag}[\sqrt{a_{00,n}(0)}, \sqrt{a_{11,n}(0)}]$ and $A_n(h) = [a_{ij,n}(h)]$. In order to verify the asymptotic efficiency of their regression coefficients, these regressors $x_{i,t}$ should satisfy a set of reasonable conditions as follows.

Assumption 1 (Grenander) The regressors $x_{i,t}$ for $i, j = 0, 1$, $t = 1, \dots, n$ and $h = 0, 1, \dots$ satisfy the following four conditions:

(a) $\lim_{n \rightarrow \infty} a_{ii,n}(0) = \infty$,

- (b) $\lim_{n \rightarrow \infty} \frac{x_{i,n+1}^2}{d_{ii,n}(0)} = 0$,
- (c) For each element of $R_n(h)$, $\lim_{n \rightarrow \infty} r_{ij,n}(h)$ exists,
- (d) $\lim_{n \rightarrow \infty} R_n(0)$ is nonsingular.

If Assumption 1 is satisfied, there exists a Hermitian matrix $M(\lambda)$ with positive semidefinite increments such that

$$R(h) = \int_{-\pi}^{\pi} e^{ih\lambda} dM(\lambda),$$

where $i = \sqrt{-1}$. This Hermitian matrix $M(\lambda)$ is one of the keys to the result in G–R.

An assumption on the regression errors u_t in the model (3) is also required, but what we really need should be sufficiently general to include a very wide class of stationary processes.

Assumption 2 The error process $\{u_t\}$ for $t = 1, \dots, n$ is stationary with $Eu_t = 0$ and $Eu_t u_{t+h} = \gamma_h$ having a spectral density $f(\lambda)$ that is continuous and positive for all $\lambda \in [-\pi, \pi]$.

Remark 1 Throughout the paper, we let

$$\sigma_n^2 := n \text{Var}(\bar{u}) = \frac{1}{n} \sum_{s=1}^n \sum_{t=1}^n \gamma_{s-t} \quad \text{and} \quad \sigma^2 := 2\pi f(0),$$

where σ^2 is the long-run variance. If Assumption 2 holds, then we have $\sigma_n^2 \rightarrow \sigma^2$ as $n \rightarrow \infty$ (see Fuller (1996, p. 310)). The boundedness of the limit is ensured by the continuity of f . In addition, both the G–R result and our subsequent analysis require the positiveness of f because its reciprocal is needed for the expression of the GLS estimator.

Under such reasonable conditions, G–R derives the necessary and sufficient condition for the OLS estimator to be asymptotically efficient, or for the variances of the OLS and GLS estimators to be asymptotically equivalent.

Theorem 1 (Grenander–Rosenblatt) *Under Assumptions 1 and 2, the OLS estimator of the model (3) is asymptotically efficient relative to the GLS estimator if and only if $M(\lambda)$ increases at not more than two points of $\lambda \in [0, \pi]$ and the sum of the ranks of the increases in $M(\lambda)$ is two.*

For a proof and further discussion, see Anderson (1971, Sect. 10.2.3) for example. We should emphasize that this theorem only works for regressors that satisfy the Grenander condition (Assumption 1). Our aim is then to seek an asymptotically efficient model without satisfying this condition.

We should also note that the result holds for a general multiple regression model, but we concentrate on the model (3) because the corresponding result in Section 3 is given with a simple regression.

2.2 Slowly Varying Regressors

These kind of regressors are known to display asymptotic singularity in the sample moment matrix (see Phillips (2007)), such that the G–R result may not be applied to the asymptotic efficiency of the regression coefficients. Our main objective here is to prove the asymptotic equivalence of the variances of the OLS and GLS estimators in a model with a SV regressor and reveal the existence of an asymptotically efficient model that does not satisfy the Grenander condition. For this purpose, we start with a definition of SV functions. A positive valued function L on \mathbb{R}_+ is called slowly varying if it satisfies, for any $r > 0$, $L(rn)/L(n) \rightarrow 1$ as $n \rightarrow \infty$. In order to deal with such an SV function $L(t)$, the following Karamata’s representation theorem is well known and essential. That is, the function L is SV if and only if it may be written in the form

$$L(n) = c(n) \exp \left(\int_a^n \frac{\varepsilon(s)}{s} ds \right) \quad \text{for } n > a$$

for some $a > 0$, where $c(n) \rightarrow c \in (0, \infty)$ and $\varepsilon(n) \rightarrow 0$ as $n \rightarrow \infty$. Considering regression theory, however, we require a stronger assumption.

Assumption 3 L is an SV function satisfying the conditions below:

(a) $L(t)$ is a smoothly slowly varying (SSV) function with Karamata representation

$$L(n) = c \exp \left(\int_a^n \frac{\varepsilon(s)}{s} ds \right) \quad \text{for } n > a$$

for some $a > 0$, and where $c > 0$, $\varepsilon \in C^\infty$ and $\varepsilon(n) \rightarrow 0$ as $n \rightarrow \infty$.

(b) $|\varepsilon(n)|$ is SSV, and ε has Karamata representation

$$\varepsilon(n) = c_\varepsilon \exp \left(\int_a^n \frac{\eta(s)}{s} ds \right) \quad \text{for } n > a,$$

where $c_\varepsilon > 0$, $\eta \in C^\infty$, $|\eta(n)|$ is SSV and $\eta(n)^2 = o(\varepsilon(n)) \rightarrow 0$ as $n \rightarrow \infty$.

(c) L is SV with remainder ϕ , that is, as $n \rightarrow \infty$,

$$\frac{L(rn)}{L(n)} = 1 + O \left(\frac{1}{\phi(n)} \right),$$

where ϕ is such that, for some $\alpha > 0$ and $n \geq \alpha$,

$$\frac{1}{\alpha\phi(n)} \leq |\varepsilon(n)| \leq \frac{\alpha}{\phi(n)}.$$

(d) $L(n)$ is monotonically increasing.

Remark 2 Conditions (a) and (b) are more restrictive assumptions than in Karamata's representation. These conditions also appeared in Phillips (2007) and Mynbaev (2009). Mynbaev (2009) introduced condition (c) to ensure discussion of the asymptotic analysis of regressions was more rigorous. Many SV functions, including all the $L(n)$ tabulated in Table 1, possess the remainder $\phi = \varepsilon$. For further discussion of SV with a remainder, see Mynbaev (2009) and Bingham Goldie and Teugels (1987) Sects. 2.3 and 3.12. Condition (d) is for convenience.

Under Assumption 3, we have an important result useful for deriving asymptotic theory

$$\varepsilon(n) = \frac{nL'(n)}{L(n)} \rightarrow 0 \quad \text{and} \quad \eta(n) = \frac{n\varepsilon'(n)}{\varepsilon(n)} \rightarrow 0 \quad \text{as } n \rightarrow \infty. \quad (4)$$

This is easily obtained by the representation theorem. Consequently (4) produces some examples of $L(n)$ in Table 1. Conversely, typical SV functions L in Table 1 satisfy Assumption 3 (a)-(c). Another application of Assumption 3 leads to the following lemma. This is used for the asymptotic expansion of the sum of the SV functions in order to evaluate their limiting behavior.

Table 1: $\varepsilon(n)$ and $\eta(n)$ corresponding to $L(n)$. $\gamma > 0$.

$L(n)$	$\varepsilon(n)$	$\eta(n)$
$\log^\gamma n$	$\gamma/\log n$	$-1/\log n$
$1/\log^\gamma n$	$-\gamma/\log n$	$-1/\log n$
$\log \log n$	$1/(\log n \log \log n)$	$-1/(\log n \log \log n) - 1/\log n$
$1/\log \log n$	$-1/(\log n \log \log n)$	$-1/(\log n \log \log n) - 1/\log n$

Lemma 1 (Phillips) *If $L(t)$ satisfies Assumption 3, then we have:*

$$\begin{aligned} \frac{1}{n} \sum_{t=1}^n L(t)^k &= L(n)^k - kL(n)^k \varepsilon(n) + k^2 L(n)^k \varepsilon(n)^2 + kL(n)^k \varepsilon(n) \eta(n) \\ &\quad - k^3 [L(n)^k \varepsilon(n)^3 + 3L(n)^k \varepsilon(n)^2 \eta(n) + L(n)^k \varepsilon(n) \eta(n)^2] \\ &\quad + o(L(n)^k \varepsilon(n) [\varepsilon(n)^2 + \varepsilon(n) \eta(n) + \eta(n)^2]). \end{aligned}$$

As in Phillips (2007, p. 568), we particularly let,

$$\begin{aligned} L_{12}(n) &= L(n) - L(n)\varepsilon(n) + L(n)\varepsilon(n)^2 + L(n)\varepsilon(n)\eta(n) \\ &\quad + o(L(n)\varepsilon(n)[\varepsilon(n) + \eta(n)]), \\ L_{22}(n) &= L(n)^2 - 2L(n)^2\varepsilon(n) + 4L(n)^2\varepsilon(n)^2 + 2L(n)^2\varepsilon(n)\eta(n) \\ &\quad + o(L(n)^2\varepsilon(n)[\varepsilon(n) + \eta(n)]), \end{aligned} \tag{5}$$

which are the asymptotic expansions of $n^{-1} \sum_{t=1}^n L(t)$ and $n^{-1} \sum_{t=1}^n L(t)^2$ respectively owing to Lemma 1.

3 RESULT

3.1 Regression Model

Consider the following regression model, for $t = 1, \dots, n$,

$$y_t = \beta_0 + \beta_1 L(t) + u_t, \quad \text{or} \quad y = X\beta + u, \tag{6}$$

where $y = [y_1, \dots, y_n]'$, $\beta = [\beta_0, \beta_1]'$ and $X = [\iota, L]$ with $\iota = [1, \dots, 1]'$ and $L = [L(1), \dots, L(n)]'$. Moreover, we write $X_t = [1, L(t)]$. The error process $\{u_t\}$ is supposed to satisfy Assumption 2 and let $u = [u_1, \dots, u_n]'$ with notation

$$\text{Var}(u) = \begin{bmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{n-1} \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{n-1} & \gamma_{n-2} & \cdots & \gamma_0 \end{bmatrix} = [\Gamma_0, \Gamma_1, \dots, \Gamma_{n-1}] = \Gamma.$$

Using these notations, we may define the OLS and GLS estimators as $\hat{\beta}_{OLS} - \beta = (X'X)^{-1}X'u$ and $\hat{\beta}_{GLS} - \beta = (X'\Gamma^{-1}X)^{-1}X'\Gamma^{-1}u$, respectively.

Remark 3 $\{u_t\}$ in the model (6) with Assumption 2 is sufficiently general in that the G–R result requires the same conditions. Phillips (2007) and Mynbaev (2009) assume that $\{u_t\}$ is a linear process with more restrictive conditions as they consider it to derive the asymptotic distributions of the OLS estimator. The former requires one summability condition of the coefficients and the existence of a finite $2p$ th moment for $\{u_t\}$ for some $p > 2$, whereas the latter requires the absolute summability of the coefficients and uniform integrability on the innovations of $\{u_t\}$.

3.2 Asymptotic Variances

We now attempt to derive the asymptotic expression for the variance of the OLS and GLS estimators of (6), and show that the basic definition of asymptotic efficiency

$$F_n^{-1} \left[\text{Var}(\hat{\beta}_{OLS}) - \text{Var}(\hat{\beta}_{GLS}) \right] F_n^{-1} = o(1) \quad (7)$$

is true for some common standardizing matrix F_n^{-1} . The first result yields the asymptotic variance of the OLS estimator.

Lemma 2 *If $L(t)$ satisfies Assumption 3, then under Assumption 2, the OLS estimator of the model (6) has asymptotic variance*

$$\text{Var}(\hat{\beta}_{OLS}) = \sigma^2 \begin{bmatrix} \frac{1}{n\varepsilon(n)^2} & -\frac{1}{nL(n)\varepsilon(n)^2} \\ -\frac{1}{nL(n)\varepsilon(n)^2} & \frac{1}{nL(n)^2\varepsilon(n)^2} \end{bmatrix} (1 + o(1)).$$

This result is the same as the asymptotic variance derived as a by-product of the limiting normal distribution in Theorem 3.1 in Phillips (2007).

We next derive the asymptotic variance of the GLS estimator.

Lemma 3 *If $L(t)$ satisfies Assumption 3, then under Assumption 2, the GLS estimator of the model (6) has asymptotic variance*

$$\text{Var}(\hat{\beta}_{GLS}) = \sigma^2 \begin{bmatrix} \frac{1}{n\varepsilon(n)^2} & -\frac{1}{nL(n)\varepsilon(n)^2} \\ -\frac{1}{nL(n)\varepsilon(n)^2} & \frac{1}{nL(n)^2\varepsilon(n)^2} \end{bmatrix} (1 + o(1)).$$

To summarize, we state the theorem below.

Theorem 2 *Under Assumption 3, the OLS estimator of (6) is asymptotically efficient. In fact, (7) is true with the normalizing matrix*

$$F_n^{-1} = \begin{bmatrix} \sqrt{n}\varepsilon(n) & 0 \\ 0 & \sqrt{n}L(n)\varepsilon(n) \end{bmatrix}.$$

Remark 4 This result is given by the simple regression (6), although Phillips (2007) dealt with polynomial and multiple regressions as well as simple regression. With the same proof, we could extend our analysis to these general models. However, the computation required would be burdensome. Accordingly, we have only focused on the simple regression (6).

4 CONCLUSIONS

In the literature on stationary time series analysis, or the classical G–R result, regressors are required to satisfy the Grenander condition when we wish to observe the asymptotic efficiency of the regression coefficients of the model. In the present paper, we have revealed the existence of an asymptotically efficient model that does not satisfy the Grenander condition through proving the asymptotic equivalence of the OLS

and GLS estimators of the model with an SV regressor. The assumptions needed for the error term are so general and identical that the G–R result is required. Thus, we may consider the paper as a form of compensation for the G–R result.

References

- [1] Anderson, T. W. (1971), *The Statistical Analysis of Time Series*, Wiley.
- [2] Bingham, N. H., C. M. Goldie and J. L. Teugels (1987), *Regular Variation*, Cambridge University Press.
- [3] Brockwell, P. J. and R. A. Davis (1991), *Time Series: Theory and Methods*, 2nd ed. Springer.
- [4] Fuller, W. (1996), *Introduction to Statistical Time Series*, 2nd ed. Wiley.
- [5] Grenander, U. (1954), “On the estimation of regression coefficients in the case of an autocorrelated disturbance,” *The Annals of Mathematical Statistics*, **25**, 252–272.
- [6] Grenander, U. and M. Rosenblatt (1957), *Statistical Analysis of Stationary Time Series*, Wiley.
- [7] Krämer, W. (1986), “Least squares regression when the independent variable follows an ARIMA process,” *Journal of the American Statistical Association*, **81**, 150–154.
- [8] Krämer, W. and U. Hassler (1998), “Limiting efficiency OLS vs. GLS when the regressors are fractionally integrated,” *Economics Letters*, **60**, 285–290.
- [9] Mynbaev, K. T. (2009), “Central limit theorems for weighted sums of linear processes: L_p -approximability versus Brownian motion,” *Econometric Theory*, **25**, 748–763.
- [10] Perron, P. and Yabu (2009), “Estimating deterministic trends with an integrated or stationary noise component,” *Journal of Econometrics*, **151**, 56–69.

- [11] Phillips, P. C. B. (2007), “Regression with slowly varying regressors and nonlinear trends,” *Econometric Theory*, **23**, 557–614.
- [12] Phillips, P. C. B. and J. Y. Park (1988), “Asymptotic equivalence of ordinary least squares and generalized least squares in regressions with integrated regressors,” *Journal of the American Statistical Association*, **83**, 111–115.
- [13] Rosenblatt, M. (1956), “On the estimation of regression coefficients of a vector-valued time series with a stationary residual,” *The Annals of Mathematical Statistics*, **27**, 99–121.
- [14] Shaman, P. (1975), “An approximate inverse for the covariance matrix of moving average and autoregressive processes,” *The Annals of Statistics*, **3**, 532–538.
- [15] Shin, D. W. and M. S. Oh (2002), “Asymptotic efficiency of the ordinary least squares estimator for regressions with unstable regressors,” *Econometric Theory*, **18**, 1121–1138.
- [16] Vogelsang, T. (1998), “Trend function hypothesis testing in the presence of serial correlation,” *Econometrica*, **66**, 123–148.
- [17] Zygmund, A. (2003), *Trigonometric Series*, Cambridge.

APPENDIX A: A Lemma for the Main Proofs

Lemma 4 Let $\sigma_n^2 = n^{-1} \sum \sum_{u,s=1}^n \gamma_{s-u}$. For any $t = 1, \dots, n$ and $n \geq 1$, the following statements are true:

$$(a) \quad \sum_{u=1}^{t-1} \sum_{s=1}^n \gamma_{s-u} - (t-1)\sigma_n^2 < \infty,$$

$$(b) \quad \sum_{u=1}^{t-1} \sum_{s=1}^n L(s)\gamma_{s-u} - \left(\sum_{u=1}^{t-1} L(u) \right) \sigma_n^2 = o(tL(n)\varepsilon(n)).$$

Proof (a) Let $\tilde{g}_n(t-1)$ denote the left-hand side of (a). From (2), the autocovariance may be written as

$$\gamma_h = \int_{-\pi}^{\pi} e^{ih\lambda} f(\lambda) d\lambda, \quad (8)$$

where f is the continuous spectral density. Hence, we have

$$\begin{aligned} |\tilde{g}_n(t-1)| &= \left| \sum_{u=1}^{t-1} \sum_{s=1}^n \gamma_{s-u} - \frac{t-1}{n} \sum_{u=1}^n \sum_{s=1}^n \gamma_{s-u} \right| \\ &= \left| \int_{-\pi}^{\pi} \left(\sum_{u=1}^{t-1} e^{-iu\lambda} \sum_{s=1}^n e^{is\lambda} - \frac{t-1}{n} \sum_{u=1}^n e^{-iu\lambda} \sum_{s=1}^n e^{is\lambda} \right) f(\lambda) d\lambda \right| \\ &= \left| \int_{-\pi}^{\pi} \left(\sin \frac{(t-1)\lambda}{2} \sin \frac{n\lambda}{2} \cos \frac{(n-t+1)\lambda}{2} - \frac{t-1}{n} \sin^2 \frac{n\lambda}{2} \right) \frac{f(\lambda)}{\sin^2 \frac{\lambda}{2}} d\lambda \right| \\ &\leq \max_{-\pi \leq \lambda \leq \pi} |f(\lambda)| \int_{-\pi}^{\pi} \left| \sin \frac{(t-1)\lambda}{2} \sin \frac{n\lambda}{2} \cos \frac{(n-t+1)\lambda}{2} - \frac{t-1}{n} \sin^2 \frac{n\lambda}{2} \right| \frac{d\lambda}{\sin^2 \frac{\lambda}{2}}. \end{aligned} \quad (9)$$

This integrand is clearly found to be bounded on $[-\pi, \pi]$ for all $t = 1, \dots, n$ and $n \geq 1$ except at $\lambda = 0$. Only at this point, the ratio is of indeterminate form. Even if the point $\lambda = 0$ is included, however, boundedness can be proved uniformly on $[-\pi, \pi]$ as follows:

Transforming the product of the trigonometric functions to their summations, and applying l'Hospital's rule twice, we obtain

$$\begin{aligned}
& \lim_{\lambda \rightarrow 0} \left(\sin \frac{(t-1)\lambda}{2} \sin \frac{n\lambda}{2} \cos \frac{(n-t+1)\lambda}{2} - \frac{t-1}{n} \sin^2 \frac{n\lambda}{2} \right) \frac{1}{\sin^2 \frac{\lambda}{2}} \\
&= \lim_{\lambda \rightarrow 0} \left(-\frac{\cos n\lambda}{4} + \frac{\cos(t-1-n)\lambda}{4} \right. \\
&\quad \left. + \frac{1 - \cos(t-1)\lambda}{4} - \frac{2(t-1)(1 - \cos n\lambda)}{4n} \right) \frac{1}{(1 - \cos \lambda)/2} \\
&= \lim_{\lambda \rightarrow 0} \left(\frac{n \sin n\lambda}{4} - \frac{(t-1-n) \sin(t-1-n)\lambda}{4} \right. \\
&\quad \left. + \frac{(t-1) \sin(t-1)\lambda}{4} - \frac{2n(t-1) \sin n\lambda}{4n} \right) \frac{1}{(\sin \lambda)/2} \quad (10) \\
&= \lim_{\lambda \rightarrow 0} \left(\frac{n^2 \cos n\lambda}{4} - \frac{(t-1-n)^2 \cos(t-1-n)\lambda}{4} \right. \\
&\quad \left. + \frac{(t-1)^2 \cos(t-1)\lambda}{4} - \frac{2n(t-1) \cos n\lambda}{4} \right) \frac{1}{(\cos \lambda)/2} \\
&= \left(\frac{n^2}{4} - \frac{(t-1-n)^2}{4} + \frac{(t-1)^2}{4} - \frac{2n(t-1)}{4} \right) \frac{1}{1/2} \\
&= 0
\end{aligned}$$

identically for all $t = 1, \dots, n$ and $n \geq 1$. Thus, the integral in (9) is finite for all $t = 1, \dots, n$ and $n \geq 1$, and this gives the proof of (a).

(b) Let $\tilde{h}_n(t-1)$ denote the left-hand side of (b). Applying summation by parts and collecting terms gives

$$\begin{aligned}
\tilde{h}_n(t-1) &= \sum_{v=1}^n L(v) \sum_{u=1}^{t-1} \gamma_{v-u} - \frac{\sum_{u=1}^{t-1} L(u)}{n} \sum_{u=1}^n \sum_{v=1}^n \gamma_{v-u} \\
&= L(n) \sum_{v=1}^n \sum_{u=1}^{t-1} \gamma_{v-u} - \sum_{s=2}^n \sum_{v=1}^{s-1} \sum_{u=1}^{t-1} \gamma_{v-u} \Delta L(s) - \frac{\sum_{u=1}^{t-1} L(u)}{n} \sum_{u=1}^n \sum_{v=1}^n \gamma_{v-u} \quad (11) \\
&= \sum_{s=2}^n \left(\sum_{v=1}^n \sum_{u=1}^{t-1} \gamma_{v-u} - \sum_{v=1}^{s-1} \sum_{u=1}^{t-1} \gamma_{v-u} - \frac{\sum_{u=1}^{t-1} L(u)}{nL(n)} \sum_{u=1}^n \sum_{v=1}^n \gamma_{v-u} \right) \Delta L(s).
\end{aligned}$$

If we use (8) like (a), the parentheses in (11) may be expressed as an integration of some trigonometric functions

$$\int_{-\pi}^{\pi} \left(\sin \frac{(t-1)\lambda}{2} \sin \frac{n\lambda}{2} \cos \frac{(n-t+1)\lambda}{2} - \sin \frac{(t-1)\lambda}{2} \sin \frac{(s-1)\lambda}{2} \cos \frac{(s-t)\lambda}{2} - \frac{\sum_{u=1}^{t-1} L(u)}{nL(n)} \sin^2 \frac{n\lambda}{2} \right) \frac{f(\lambda)}{\sin^2 \frac{n\lambda}{2}} d\lambda. \quad (12)$$

If we write the integrand of (12) as $T_{n,s,t}(\lambda)f(\lambda)$, (i.e., $T_{n,s,t}(\lambda)$ is a polynomial of the ratio of the trigonometric functions), then the absolute value of (12) is

$$\left| \int_{-\pi}^{\pi} T_{n,s,t}(\lambda)f(\lambda)d\lambda \right| \leq \max_{-\pi \leq \lambda \leq \pi} |f(\lambda)| \int_{-\pi}^{\pi} |T_{n,s,t}(\lambda)| d\lambda. \quad (13)$$

As in (a), the integrand $|T_{n,s,t}(\lambda)|$ is bounded in $[-\pi, \pi]$ except at the indeterminate point $\lambda = 0$. In this case, however, the behavior of the integrand $|T_{n,s,t}(\lambda)|$ around $\lambda = 0$ is different from that of (a) in that it depends on s, t and n , and we now observe this. For deriving the form of the first and second terms of $T_{n,s,t}(\lambda)$ around $\lambda = 0$, the computational method is the same as (10), although it is also more tedious. For the third term, we note that

$$\begin{aligned} \frac{\sum_{u=1}^{t-1} L(u)}{nL(n)} &= \frac{(t-1)L(t-1)(1 + O(\varepsilon(t)))}{nL(n)} \\ &= \frac{t-1}{n}(1 + O(\varepsilon(n)))(1 + O(\varepsilon(t))) \end{aligned}$$

for large n because of Assumption 3 (c) and Lemma 1. Thus, after collecting terms, we can observe for small $|\lambda|$ and large n that

$$T_{n,s,t}(\lambda) \sim t - st + s - 1 + tnO(\varepsilon(t)).$$

Returning to the integral (13), we split the area of the integral $(-\delta/n, \delta/n)$ and $[-\pi, \pi] \setminus (-\delta/n, \delta/n)$ for any fixed $\delta > 0$. Then the last integral in (13) becomes

$$\begin{aligned} \int_{[-\pi, \pi]} |T_{n,s,t}(\lambda)| d\lambda &= \left(\int_{(-\delta/n, \delta/n)} + \int_{[-\pi, \pi] \setminus (-\delta/n, \delta/n)} \right) |T_{n,s,t}(\lambda)| d\lambda \\ &= \frac{\delta}{n} O(t - st + s - 1 + tn\varepsilon(t)) + O(1) \\ &= \delta O(t\varepsilon(t)) = o(t\varepsilon(t)) \end{aligned} \quad (14)$$

since $\delta > 0$ is arbitrary. Finally, from (11) and (14), we obtain

$$\begin{aligned} \left| \tilde{h}_n(t-1) \right| &\leq \sum_{s=2}^n \left(\int_{[-\pi, \pi]} |T_{n,s,t}(\lambda)| d\lambda \right) \Delta L(s) = o(t\varepsilon(t)) \sum_{s=2}^n \Delta L(s) \\ &= o(tL(n)\varepsilon(t)). \end{aligned}$$

This is the desired result. \blacksquare

APPENDIX B: Proofs for Section 3

Proof of Lemma 2 We can write

$$X'\Gamma X = \begin{bmatrix} \iota'\Gamma\iota & \iota'\Gamma L \\ \iota'\Gamma L & L'\Gamma L \end{bmatrix} \quad \text{and} \quad \sigma_n^2 = \frac{1}{n} \sum_{s=1}^n \sum_{t=1}^n \gamma_{s-t},$$

and first show that

- (i) $\iota'\Gamma\iota = \sigma_n^2 n$,
- (ii) $\iota'\Gamma L = \sigma_n^2 \iota' L + o(nL(n)\varepsilon(n)^2)$, and
- (iii) $L'\Gamma L = \sigma_n^2 L' L + o(nL(n)^2\varepsilon(n)^2)$.

If these equations are true, the variance leads to

$$\begin{aligned}
\text{Var}(\hat{\beta}_{OLS}) &= (X'X)^{-1}X'\Gamma X(X'X)^{-1} \\
&= \begin{bmatrix} n & \iota'L \\ \iota'L & L'L \end{bmatrix}^{-1} \begin{bmatrix} \iota'\Gamma\iota & \iota'\Gamma L \\ L'\Gamma\iota & L'\Gamma L \end{bmatrix} \begin{bmatrix} n & \iota'L \\ L'\iota & L'L \end{bmatrix}^{-1} \\
&= \frac{1}{(nL'L - (\iota'L)^2)^2} \begin{bmatrix} L'L & -\iota'L \\ -L'\iota & n \end{bmatrix} \\
&\quad \times \sigma_n^2 \begin{bmatrix} n & \iota'L + o(nL(n)\varepsilon(n)^2) \\ L'\iota + o(nL(n)\varepsilon(n)^2) & L'L + o(nL(n)^2\varepsilon(n)^2) \end{bmatrix} \begin{bmatrix} L'L & -\iota'L \\ -L'\iota & n \end{bmatrix} \quad (15) \\
&= \frac{\sigma_n^2}{n^4L(n)^4\varepsilon(n)^4(1+o(1))} \\
&\quad \times \begin{bmatrix} n^3L(n)^4\varepsilon(n)^2(1+o(1)) & -n^3L(n)^3\varepsilon(n)^2(1+o(1)) \\ -n^3L(n)^3\varepsilon(n)^2(1+o(1)) & n^3L(n)^2\varepsilon(n)^2(1+o(1)) \end{bmatrix} \\
&= \frac{\sigma_n^2}{1+o(1)} \begin{bmatrix} \frac{1}{n\varepsilon(n)^2} & -\frac{1}{nL(n)\varepsilon(n)^2} \\ -\frac{1}{nL(n)\varepsilon(n)^2} & \frac{1}{nL(n)^2\varepsilon(n)^2} \end{bmatrix} (1+o(1)),
\end{aligned}$$

where we have used the fact that

$$nL'L - (\iota'L)^2 = n^2L(n)^2\varepsilon(n)^2(1+o(1))$$

due to Lemma 1. Thus, it suffices to prove (i), (ii) and (iii).

(i) is trivial by definition of σ_n^2 . Next, for (ii), we can write

$$|\iota'\Gamma L - \sigma_n^2\iota'L| = \left| \sum_{t=1}^n L(t) \left(\sum_{s=1}^n \gamma_{s-t} - \sigma_n^2 \right) \right| = \left| \sum_{t=1}^n L(t)g_n(t) \right|, \quad (16)$$

where $g_n(t) = \sum_{s=1}^n \gamma_{s-t} - \sigma_n^2$ and let $\tilde{g}_n(t) = \sum_{u=1}^t g_n(u)$. Applying summation by parts to (16), we have, together with the monotonicity of L ,

$$\left| \tilde{g}_n(n)L(n) - \sum_{t=2}^n \tilde{g}_n(t-1)\Delta L(t) \right| \leq \sum_{t=2}^n |\tilde{g}_n(t-1)| \Delta L(t), \quad (17)$$

where $\Delta L(t) = L(t) - L(t-1)$. This notation may be used repeatedly. Here we remark that $\max_{1 \leq t \leq n} |\tilde{g}_n(t)|$ is bounded because of Lemma 4 (a). Therefore, for a sufficiently large n , (17) is less than or equal to:

$$\max_{1 \leq t \leq n} |\tilde{g}_n(t)| \sum_{t=2}^n \Delta L(t) = \max_{1 \leq t \leq n} |\tilde{g}_n(t)| (L(n) - L(1)) = O(L(n)), \quad (18)$$

which is certainly $o(nL(n)\varepsilon(n)^2)$.

We prove (iii) in a similar way. We write

$$|L'\Gamma L - \sigma_n^2 L'L| = \left| \sum_{t=1}^n L(t) \left(\sum_{s=1}^n L(s) \gamma_{s-t} - \sigma_n^2 L(t) \right) \right| = \left| \sum_{t=1}^n L(t) h_n(t) \right|, \quad (19)$$

where $h_n(t)$ signifies $\sum_{s=1}^n L(s) \gamma_{s-t} - \sigma_n^2 L(t)$ and let $\tilde{h}_n(t) = \sum_{u=1}^t h_n(u)$. Applying summation by parts to (19) yields

$$\begin{aligned} & \left| L(n) \tilde{h}_n(n) - \sum_{t=2}^n \tilde{h}_n(t-1) \Delta L(t) \right| \\ & \leq \left| L(n) \tilde{h}_n(n) \right| + \left| \sum_{t=2}^n \tilde{h}_n(t-1) \Delta L(t) \right| = A(n) + B(n), \quad \text{say.} \end{aligned}$$

Because of the symmetry of γ_{s-t} and the result of (ii), the first term $A(n)$ is calculated so that

$$\begin{aligned} A(n) &= L(n) \left| \sum_{t=1}^n \left(\sum_{s=1}^n L(s) \gamma_{s-t} - \sigma_n^2 L(t) \right) \right| \\ &= L(n) \left| \sum_{t=1}^n L(t) \left(\sum_{s=1}^n \gamma_{s-t} - \sigma_n^2 \right) \right| = O(L(n)^2). \end{aligned} \quad (20)$$

Meanwhile, owing to Lemma 4 (b) directly, the second term $B(n)$ is written as

$$\begin{aligned} B(n) &= \left| \sum_{t=2}^n \tilde{h}_n(t-1) \Delta L(t) \right| \leq \sum_{t=2}^n |o(tL(n)\varepsilon(t))| \Delta L(t) \\ &= o \left(L(n) \sum_{t=2}^n n \frac{t}{n} \varepsilon \left(n \frac{t}{n} \right) \Delta L \left(n \frac{t}{n} \right) \right) \\ &= o \left(nL(n) \int_0^1 r \varepsilon(nr) dL(nr) \right) \end{aligned} \quad (21)$$

under Assumption 3. Changing variables and the formula (4) imply that (21) leads to

$$\begin{aligned}
o\left(L(n) \int_0^n t\varepsilon(t) dL(t)\right) &= o\left(L(n) \int_0^n t\varepsilon(t) L'(t) dt\right) \\
&= o\left(L(n) \int_0^n t\varepsilon(t) \frac{L(t)\varepsilon(t)}{t} dt\right) \\
&= o(nL(n)^2\varepsilon(n)^2).
\end{aligned} \tag{22}$$

From (20) and (22), the desired result of (iii) is given by

$$|L'\Gamma L - \sigma_n^2 L'L| \leq A(n) + B(n) = o(nL(n)^2\varepsilon(n)^2). \tag{23}$$

In consequence, we obtain the asymptotic form of $\text{Var}(\hat{\beta}_{OLS})$ as in (15) from (i), (ii) and (iii). ■

Proof of Lemma 3 First, it should be proved for the element γ^{s-t} of Γ^{-1} that

$$\begin{aligned}
(i) \quad & \iota'\Gamma^{-1}\iota = \omega_n^2 n, \\
(ii) \quad & \iota'\Gamma^{-1}L = \omega_n^2 \iota'L + o(nL(n)\varepsilon(n)^2), \quad \text{and} \\
(iii) \quad & L'\Gamma^{-1}L = \omega_n^2 L'L + o(nL(n)^2\varepsilon(n)^2),
\end{aligned}$$

where

$$\omega_n^2 := \frac{1}{n} \sum_{s=1}^n \sum_{t=1}^n \gamma^{s-t}.$$

They can be shown from the proof of Lemma 2 as long as the spectral density of γ^{s-t} is continuous on $[-\pi, \pi]$. But this is true for a sufficiently large n because of the fact given in Shaman (1975) that Γ^{-1} can be asymptotically replaced by the matrix Γi , whose (s, t) element is defined by

$$\gamma^{i_{s-t}} := \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} e^{i(s-t)\lambda} \frac{1}{f(\lambda)} d\lambda.$$

Here we note that $f > 0$ on $[-\pi, \pi]$ by Assumption 2. That is, the spectral density of γ^{s-t} is given by $(2\pi)^{-2}f(\lambda)^{-1}$ in the asymptotic sense, and hence is continuous under Assumption 2.

Thus, we obtain the expression from (i), (ii) and (iii) that

$$\begin{aligned}
\text{Var}(\hat{\beta}_{GLS}) &= (X'\Gamma^{-1}X)^{-1} = \begin{bmatrix} \iota'\Gamma^{-1}\iota & \iota'\Gamma^{-1}L \\ L'\Gamma^{-1}\iota & L'\Gamma^{-1}L \end{bmatrix}^{-1} \\
&= \frac{\omega_n^{-2}}{nL'L - (\iota'L)^2 + o(n^2L(n)^2\varepsilon(n)^2)} \\
&\quad \times \begin{bmatrix} L'L + o(nL(n)^2\varepsilon(n)^2) & -\iota'L + o(nL(n)\varepsilon(n)^2) \\ -L'\iota + o(nL(n)\varepsilon(n)^2) & n \end{bmatrix} \\
&= \frac{\omega_n^{-2}}{n^2L(n)^2\varepsilon(n)^2(1+o(1))} \begin{bmatrix} nL(n)^2(1+o(1)) & -nL(n)(1+o(1)) \\ -nL(n)(1+o(1)) & n \end{bmatrix} \\
&= \frac{\omega_n^{-2}}{1+o(1)} \begin{bmatrix} \frac{1}{n\varepsilon(n)^2} & -\frac{1}{nL(n)\varepsilon(n)^2} \\ -\frac{1}{nL(n)\varepsilon(n)^2} & \frac{1}{nL(n)^2\varepsilon(n)^2} \end{bmatrix} (1+o(1)),
\end{aligned}$$

Finally, we specify the asymptotic form of ω_n^2 , but we know that the value converges to the long-run variance, or 2π times the spectrum evaluated at the origin. Therefore, it follows that

$$\omega_n^2 \rightarrow 2\pi \frac{1}{4\pi^2 f(0)} = \frac{1}{2\pi f(0)} = \frac{1}{\sigma^2},$$

which yields the desired result. \blacksquare