Hitotsubashi University

Global COE Hi-Stat Discussion Paper Series 209

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

Regression with a Slowly Varying Regressor in the Presence of a Unit Root

Yoshimasa Uematsu

October 2011

Hi-Stat Discussion Paper

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

Regression with a Slowly Varying Regressor in the Presence of a Unit Root

Yoshimasa Uematsu*

Graduate School of Economics, Hitotsubashi University

Abstract

This paper considers the regression model with a slowly varying (SV) regressor in the presence of a unit root in serially correlated disturbances. This regressor is known to be asymptotically collinear with the constant term; see Phillips (2007). Under nonstationarity, we find that the estimated coefficients of the constant term and the SV regressor are asymptotically normal, but neither is consistent. Further, we derive the limiting distribution of the unit root test statistic. We may here observe that the finite sample approximation to the limiting one is not monotone and it is poor due to the influence of the collinear regressor. In order to construct a well-behaved test statistic, we recommend dropping the constant term intentionally from the regression and computing the statistics, which are still consistent under the true model having the constant term. The powers and sizes of these statistics are found to be well-behaved

*The author thanks Eiji Kurozumi and all the participants of the annual meeting of the Kansai Econometric Society at Osaka University and the workshop at Hitotsubashi University for helpful suggestions. The author also thanks Research Fellowships of the Japan Society for the Promotion of Science for Young Scientists for financial support.

through simulation studies. Finally, these results are extended to general Phillips and Perron-type statistics.

1 Introduction

Considerable research papers on deterministic time-trending models have been produced. Even if we restrict our attention to the works that focus on problems related to a (near) unit root, there are still many papers involved. Considering formulations of trends in these papers, we observe that most of these studies employed a general vector of time-trending regressors like polynomials, but ruled out trends that are asymptotically collinear with a constant term or the other trends; see, for example, Vogelsang (1998) and Bunzel and Vogelsang (2005).

In this background, Phillips (2007) established the theory on stationary models possessing slowly varying (SV) regressors, which are asymptotically collinear. For example, the typical model is the following logarithmic trend model

$$y_t = \alpha + \beta \log t + u_t \quad \text{for} \quad t = 1, \ldots, n, \tag{1}$$

where $\{u_t\}$ is stationary and satisfies some conditions. If we let $D_n = \mathrm{diag}[\sqrt{n}, \sqrt{n}\log n]$ and $F_n^{-1} = \mathrm{diag}[\sqrt{n}\log^{-1} n, \sqrt{n}]$, then the sample moment matrices of the regressors $X_t = [1, \log t]$ and $X = [X_1', \ldots, X_n']'$ behave asymptotically so that

$$D_n^{-1}(X'X)D_n^{-1} \to \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad F_n^{-1}(X'X)^{-1}F_n^{-1} \to \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \tag{2}$$

from Phillips (2007) or Lemma 2.1 below. (2) implies the singularity of the asymptotic covariance matrix of the OLS estimator $(\hat{\alpha}, \hat{\beta})$ in the regression (1). Phillips (2007) also revealed that the OLS estimator $(\hat{\alpha}, \hat{\beta})$ is consistent and asymptotically normally distributed, but the convergence rates of $\hat{\alpha}$ and $\hat{\beta}$ are affected by the presence of the logarithmic trend. In view of asymptotic arguments in stationary cases, Phillips (2007) relied on uniform strong approximation of partial sums by Brownian motions, but the condition was rather restrictive. Mynbaev (2009), on the other hand, introduced the "L_p-approximation" technique in this connection. The sophisticated idea is to approximate sequences of deterministic weights with functions of a continuous argument and apply the Central Limit Theorem (CLT). This enables us to establish the results of Phillips (2007) under less stringent conditions.

The present paper extends the asymptotic theory for Phillips' model to the integrated case. In addition, unit root tests in the presence of such an SV regressor are also investigated. Then, a few problems arise in the proceeding study. First, the OLS estimator of the trend coefficient $\hat{\beta}$ is inconsistent in addition to that of the constant term $\hat{\alpha}$, so that the regression is almost meaningless if the errors are integrated; see Section 3.1. This phenomenon might be understood that an SV regressor is classified into a constant in the asymptotic sense and this feature is amplified under integrated errors. Of course, an analysis like that of Canjels and Watson (1997) no longer makes sense. Thus, we should emphasize the necessity of unit root tests when it comes to employing an SV regressor. Second, another problem emerges when we derive asymptotic test statistics. If we construct Phillips and Perron (PP)-type test statistics, their finite sample distributions hardly approach the limiting ones. This makes it difficult to test a unit root that is based on the limiting critical values; see Section 4.1. Despite the adversity, we present a solution to cope with this problem by using the misspecified regression model in which we drop a constant term on purpose. Such an intentionally misspecified procedure is asymptotically justified under the true model with a constant term and brings about a significant improvement in terms of the size and power in finite sample situations; see Section 4.2.

The plan of the paper is as follows. Section 2 includes some assumptions and preliminary theories for SV functions based on the results obtained by Phillips (2007). Section 3 states the main results, namely the weak convergence of the regression coefficients derived under integrated errors. We also calculate the limiting distributions of the unit root coefficient and the corresponding t-statistic. Section 4 studies the properties of the unit root test statistics derived in Section 3 using simulations involving a procedure to improve the performance of the tests. Section 5 concludes. Appendix A gives a lemma used in the proofs. The proofs of the analytical results derived in Sections 3 and 4 are presented in Appendix B.

2 Assumptions and Preliminary Results

Our main objective is to analyze the regression model with an SV regressor under a unit root assumption. For this purpose, we start with the definition of SV functions. A positively valued function on $\mathbb{R}_+$ is called slowly varying if it satisfies, for any $r > 0$, $L(rn)/L(n) \to 1$ as $n \to \infty$. To deal with such an SV function $L(n)$, the so-called Karamata's representation theorem is well-known and essential. That is, the function $L(n)$ varies slowly if and only if it is written in the form

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

for some $a > 0$, where $c(t) \to c \in (0, \infty)$ and $\varepsilon(n) \to 0$ as $t \to \infty$. Considering regression theory, however, we require a stronger assumption on L in terms of its smoothness.

Assumption 2.1 (SSV) L is an SV function satisfying the conditions below:

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

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

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

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

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

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

(c) $L(n)$ is SV with remainder $\phi(n)$, that is, as $n \to \infty$,

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

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

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

Remark 2.1 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 that the asymptotic analysis of the 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) in Sections 2.3 and 3.12. convenience.

Under Assumption 2.1, we have an important result that is useful for deriving asymptotic results as follows:

$$\varepsilon(n) = \frac{nL'(n)}{L(n)} \to 0 \quad \text{and} \quad \eta(n) = \frac{n\varepsilon'(n)}{\varepsilon(n)} \to 0 \quad \text{as} \quad n \to \infty. \tag{3}$$

This is easily obtained by the representation theorem. Consequently, (3) produces some examples of $L(t)$ in Table 1. Conversely, typical SV functions L in Table 1 satisfy Assumption 2.1. Another application of Assumption 2.1 leads to the following lemma.

Table 1: $\varepsilon(t)$ and $\eta(t)$ associated with some $L(t)$. $\gamma > 0$.

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

Lemma 2.1 (Phillips) *If $L(t)$ satisfies Assumption 2.1, then we have*

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

This is used for asymptotic expansion of the sum of SV functions to evaluate the limiting behavior of estimators and is frequently used in the proofs of our results.

Another tool for developing the asymptotic theory is obtained in de Haan and Resnick (1998), and introduced in Phillips (2007) and Mynbaev (2009). This enables us to replace a general SV function with a logarithmic function asymptotically, as in Lemma A in Appendix A.

Lemma 2.2 (de Haan and Resnick) *If $L(t)$ satisfies Assumption 2.1, then, for all $r > 0$,*

$$L(rn) = L(n) + L(n)\varepsilon(n)\log r[1 + o(1)]$$

uniformly in $r \in [a, b]$ for any $0 < a < b < \infty$.

In the remainder of this section, we review the weak convergence results under stationary disturbances given by Phillips (2007). First, the following linear process assumption for error terms is sufficient for the asymptotic theory.

Assumption 2.2 (LP) For all $t > 0$, u_t has Wold representation

$$u_t = C(L)e_t = \sum_{j=0}^{\infty} c_j e_{t-j}, \quad \sum_{j=0}^{\infty} j|c_j| < \infty, \quad C(1) \neq 0,$$

with $e_t \sim i.i.d.(0, \sigma_e^2)$ and $E|u_t|^{2p} < \infty$ for some $p > 2$.

We let $\bar{L} = n^{-1} \sum_{t=1}^{n} L(t)$, $S_t = \sum_{s=1}^{t} u_s$ and $\bar{u} = n^{-1} S_n$. Then, under Assumption 2.2, it is sufficient that the functional CLT for S_t works well so that $n^{-1/2} S_{[n\cdot]} \to_d \sigma w(\cdot)$, where w is a standard Wiener measure and the variance is $\sigma^2 = \sigma_e^2 C(1)^2$. However, Phillips (2007) relied on the uniformly strong approximation by $w(\cdot)$ for his proofs. In other words,

$$\sup_{1 \leq t \leq n} \left| \frac{S_{t-1}}{\sigma\sqrt{n}} - w\left(\frac{t-1}{n}\right) \right| = o_{a.s.}\left(\frac{1}{n^{1/2-1/p}}\right) \tag{4}$$

is applied for some integer $p > 2$.

Remark 2.2 Mynbaev (2009) pointed out that the Phillips assumption on disturbances is relatively restrictive, and relaxed it to the absolutely summable condition with uniformly integrable innovations $\{e_t\}$. Under the condition, the results are derived by the CLT based on "L_p-approximability" of SV regressors. We rely on the Phillips way to provide the intuition for observing the influence of an SV regressor under a nonstationary assumption.

Lemma 2.3 (Phillips) *If $L(t)$ satisfies Assumption 2.1 and u_t satisfies Assumption 2.2, then, as $n \to \infty$,*

$$(i) \quad \frac{1}{\sqrt{n}L(n)} \sum_{t=1}^{n} L(t) u_t \xrightarrow{d} \sigma w(1) \overset{d}{=} N(0, \sigma^2),$$

$$(ii) \quad \frac{1}{\sqrt{n}L(n)\varepsilon(n)} \sum_{t=1}^{n} (L(t) - \bar{L}) u_t \xrightarrow{d} \sigma \int_0^1 (1 + \log r) dw(r) \overset{d}{=} N(0, \sigma^2).$$

Lemma 2.3 leads to the asymptotic results for a regression. If we consider the following simple regression model

$$y_t = \alpha + \beta L(t) + u_t \quad \text{for} \quad t = 1, \dots, n,$$

where u_t is assumed to satisfy Assumption 2.2, then the asymptotic distribution of the OLS estimator is obtained by Lemma 2.1 and 2.3.

Theorem 2.1 (Phillips) *If $L(t)$ satisfies Assumption 2.1 and u_t satisfies Assumption 2.2, then we have*

$$\begin{bmatrix} \sqrt{n}\varepsilon(n)(\hat{\alpha} - \alpha) \\ \sqrt{n}L(n)\varepsilon(n)(\hat{\beta} - \beta) \end{bmatrix} \xrightarrow{d} N\left(0,\, \sigma^2 \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\right).$$

We observe the singularity of the asymptotic covariance matrix. This is caused by the asymptotic collinearity of the regressors such as (2).

Instead of Assumption 2.2, we adopt the following assumption in the rest of this paper to consider a unit root case.

Assumption 2.3 (UR) The process $\{u_t\}$ possesses a unit root under the null hypothesis $\rho = 1$ in $u_t = \rho u_{t-1} + v_t$, where $\{v_t\}$ is a linear process having the same Wold representation given in Assumption 2.2 with the innovations $e_t \sim i.i.d.(0, \sigma_e^2)$ and $E|v_t|^{2p} < \infty$ for some $p > 2$.

Under Assumption 2.3, we let $\sigma_L^2 = n\text{Var}(\bar{v}) = \sigma_e^2 C(1)^2$ and $\sigma_S^2 = \text{Var}(v_t) = \sigma_e^2 \sum_{j=0}^{\infty} c_j^2$, called the long-run and short-run variance, respectively. If we suppose v_t is i.i.d. random variables, on the other hand, we know that $n\text{Var}(\bar{e}) = \text{Var}(e_t) = \sigma_e^2$. This simplification is used in simulation studies to exclude the effect of the long-run variance estimation error and focus on the influence of an SV regressor.

3 Analytical Results

The main purpose of this section is to reveal the asymptotic behavior of the regression with an SV regressor in the presence of a unit root. In particular, the following simple regression model is considered:

$$y_t = \alpha + \beta L(t) + u_t, \qquad t = 1, \dots, n, \tag{5}$$

where the SV regressor $L(t)$ and disturbances $\{u_t\}$ are supposed to satisfy Assumptions 2.1 and 2.3, respectively.

3.1 Weak convergence of the regression coefficients

We also denote $F_1(r)$, the one-folded integrated Brownian motion defined by $\int_0^r w(s)ds$. The derivation of the asymptotic distribution of the OLS estimator $(\hat{\alpha}, \hat{\beta})$ in (5) requires the following lemma:

Lemma 3.1 *Under Assumption 2.1 and Assumption 2.3, we have, as $n \to \infty$,*

$$(i) \quad \frac{1}{n\sqrt{n}L(n)} \sum_{t=1}^{n} L(t)u_t \xrightarrow{d} \sigma_L \int_0^1 w(r)dr,$$

$$(ii) \quad \frac{1}{n\sqrt{n}L(n)\varepsilon(n)} \sum_{t=1}^{n} (L(t) - \bar{L})u_t \xrightarrow{d} \sigma_L \int_0^1 \left(w(r) - \frac{F_1(r)}{r} \right) dr.$$

We should note that the limiting distribution in Lemma 3.1 (i) is found to be identical to $N(0, \sigma_L^2/3)$. Similarly, the limiting law in (ii) is $N(0, 2\sigma_L^2/27)$. According to Phillips (2007) or Lemma 2.1, note, moreover, that $\sum_{t=1}^{n}(L(t) - \bar{L})^2 = nL(n)^2\varepsilon(n)^2[1 + O(\varepsilon(n))]$ holds for large n. Thus, a direct application of Lemma 3.1 gives the next theorem.

Theorem 3.1 *Under Assumption 2.1 and Assumption 2.3, the OLS estimator $(\hat{\alpha}, \hat{\beta})$*

is asymptotically distributed as follows:

$$\begin{bmatrix} \dfrac{\varepsilon(n)}{\sqrt{n}}\left(\hat{\alpha}-\alpha\right) \\ \\ \dfrac{L(n)\varepsilon(n)}{\sqrt{n}}\left(\hat{\beta}-\beta\right) \end{bmatrix} \xrightarrow{d} N\left(0,\ \frac{2\sigma_L^2}{27}\begin{bmatrix} 1 & -1 \\ \\ -1 & 1 \end{bmatrix}\right).$$

It should be emphasized that, because any SV function $L(n)$ possesses asymptotic order $o(\sqrt{n})$, the OLS estimators, $\hat{\alpha}$ and $\hat{\beta}$, cannot be consistently estimated in the model (5). This result contrasts with the case where the simple trend t is employed. Considering models with an SV regressor, we therefore remark that the existence of a unit root leads to a meaningless regression and that testing for a unit root is indispensable.

3.2 Weak convergence of the unit root coefficient

Using the results in Theorem 3.1, we derive the limiting distribution of the OLS estimator $\hat{\rho}$ in the residual-based regression $\hat{u}_t = \rho\hat{u}_t + v_t$ with $\rho = 1$ under Assumption 2.3. These residuals $\hat{u}_t$ are obtained from the regression model (5), so that we have:

$$\hat{u}_t = u_t - \bar{u} - (L(t) - \bar{L})\left(\hat{\beta}-\beta\right) \quad \text{for} \ \ t = 1, \ldots, n. \tag{6}$$

From (6), the scaled OLS estimator $\hat{\rho}$ is obtained by

$$\begin{aligned} n(\hat{\rho}-1) &= \frac{1}{n}\sum_{t=2}^{n} \hat{u}_{t-1}\left(\hat{u}_t - \hat{u}_{t-1}\right) \Bigg/ \left[\frac{1}{n^2}\sum_{t=2}^{n}\hat{u}_{t-1}^2\right] \\ &= \left[\frac{1}{2n}(\hat{u}_n^2 - \hat{u}_1^2) - \frac{1}{2n}\sum_{t=2}^{n}(\hat{u}_t - \hat{u}_{t-1})^2\right] \Bigg/ \left[\frac{1}{n^2}\sum_{t=1}^{n}\hat{u}_t^2 - \frac{1}{n^2}\hat{u}_n^2\right], \end{aligned}$$

where the scale coefficient n is tentative. This asymptotic behavior depends on the four terms $\hat{u}_n^2$, $\hat{u}_1^2$, $\sum_{t=2}^{n}(\hat{u}_t - \hat{u}_{t-1})^2$ and $\sum_{t=1}^{n}\hat{u}_t^2$. Their weak convergences are presented in the following lemma.

Lemma 3.2 *Under Assumption 2.1 and Assumption 2.3, we have, as* $n \to \infty$*,*

$$
\begin{aligned}
&(i) \quad \frac{1}{n}\hat{u}_n^2 = O_p(1),\\
&(ii-a) \quad \frac{L(n)^2\varepsilon(n)^2}{n}\hat{u}_1^2 \xrightarrow{d} L(1)^2\sigma_L^2 U_1 \quad if \quad L(n) \to 0,\\
&(ii-b) \quad \frac{\varepsilon(n)^2}{n}\hat{u}_1^2 \xrightarrow{d} \sigma_L^2 U_1 \quad if \quad L(n) \to \infty,\\
&(iii) \quad \frac{1}{n}\sum_{t=2}^{n}(\hat{u}_t - \hat{u}_{t-1})^2 = O_p(1),\\
&(iv) \quad \frac{1}{n^2}\sum_{t=1}^{n}\hat{u}_t^2 \xrightarrow{d} \sigma_L^2 V_1,
\end{aligned}
$$

where

$$
\begin{aligned}
U_1 &= \left\{\int_0^1 \left(w(r) - \frac{F_1(r)}{r}\right) dr\right\}^2 \quad and\\
V_1 &= \int_0^1 w(r)^2 dr - \left(\int_0^1 w(r)dr\right)^2 - \left\{\int_0^1 \left(w(r) - \frac{F_1(r)}{r}\right) dr\right\}^2.
\end{aligned}
$$

Since the dominated term is found to be $\hat{u}_1^2$, by Lemma 3.2, we conclude that the limiting distribution of the unit root coefficient test statistic, $n\varepsilon(n)^2(\hat{\rho} - 1)$, is given by the next theorem.

Theorem 3.2 *Suppose that Assumption 2.1 and Assumption 2.3 are satisfied. If* $L(n) \to \infty$ *as* $n \to \infty$*, then it follows that*

$$
n\varepsilon(n)^2\left(\hat{\rho} - 1\right) \xrightarrow{d} -\frac{U_1}{2V_1}.
$$

This limiting distribution is free from the nuisance parameters, σ_S^2 and σ_L^2, but it turns out to be useless in that the finite sample approximation is poor; see Section 4.1. The corresponding scaled t-statistic $\varepsilon(n)^2 t_{\hat{\rho}} = \varepsilon(n)^2(\hat{\rho} - 1)/s.e.(\hat{\rho})$ for testing the null hypothesis, $H_0 : \rho = 1$, is also obtained.

Corollary 3.1 *Suppose that Assumption 2.1 and Assumption 2.3 are satisfied. If* $L(n) \to \infty$ *as* $n \to \infty$*, then it follows that:*

$$
\varepsilon(n)^2 t_{\hat{\rho}} \xrightarrow{d} -\frac{\sigma_L}{\sigma_S}\frac{U_1}{2\sqrt{V_1}}.
$$

Remark 3.1 The test statistic, $t_{\hat{\rho}}$, in Corollary 3.1 requires the computation of $s.e.(\hat{\rho})$ or $\hat{\sigma}_S^2$, which is a consistent estimator of $\sigma_S^2 = \mathrm{Var}(v_t)$. However, the natural candidate estimator, $n^{-1}\sum_{t=2}^{n}(\hat{u}_t - \hat{u}_{t-1})^2$, is not consistent for σ_S^2 but has a nondegenerate distribution, in the limit, as in Lemma 3.2 (iii). We suggest a method to construct a consistent estimator, $\hat{\sigma}_S^2$, in Remark 4.1.

4 Properties of the Unit Root Test Statistics

Theorem 3.2 and Corollary 3.1 have yielded the limiting distributions of the unit root regression coefficient and corresponding t-statistic. In this section, we examine the effect of the SV regressor on these statistics by simulations. To observe this, we employ Assumption 2.3, but restrict $\{v_t\}$ being i.i.d. random variables with mean zero and variance unity to exclude the influence caused by the long-run variance estimation.

If we need to embody the form of the regressor $L(t)$ for simulation studies, we always use $\log t$. The number of replications is $10,000$ unless otherwise noted.

4.1 Finite sample behaviors

In this subsection, we first observe the finite sample behaviors of the simulated cumulative distribution functions (CDFs) in Theorem 3.2 and Corollary 3.1. Figures 1 and 2 show the finite sample and limiting CDFs of these statistics. For each figure, the finite sample CDFs are expressed in dotted, dashed and solid lines for each sample size 100, 500 and $50,000$, and the limiting CDF is expressed in bold lines.

These graphs indicate that the approaching manner of the finite sample CDFs to the limiting ones is not monotonic. That is, finite sample CDFs approach the limiting ones from the left to the right in their upper tails first. Then, in their lower

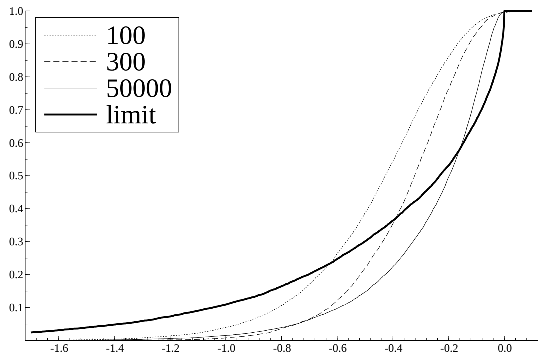

Figure 1: CDFs of $n\varepsilon(n)^2(\hat{\rho}-1)$.

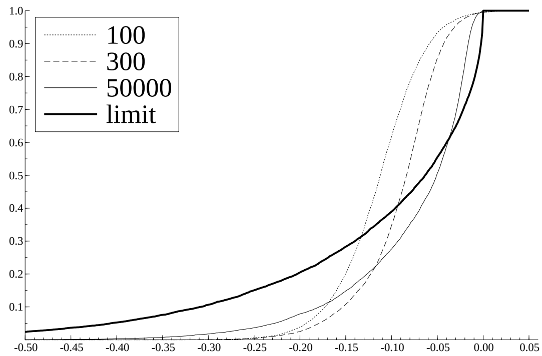

Figure 2: CDFs of $\varepsilon(n)^2 t_{\hat{\rho}}$.

tails, the CDFs are attracted from below to the above limits more slowly. As a consequence, fatal size distortions are provoked in a unit root testing based on the limiting critical values in their lower tails. Table 2 shows the percentage points of the limiting distributions. From these graphs and the limiting percentage points, we are convinced that 5% empirical sizes are very close to zero and the tests based on them are almost impossible.

Table 2: Percentage points of the limiting distributions.

	Probability (%)							
Statistic	1.0	2.5	5.0	10.0	90.0	95.0	97.5	99.0
$n\varepsilon(n)^2(\hat{\rho}-1)$	-2.13	-1.69	-1.38	-1.05	-0.01	-0.00	-0.00	-0.00
$\varepsilon(n)^2 t_{\hat{\rho}}$	-0.63	-0.50	-0.41	-0.31	-0.00	-0.00	-0.00	-0.00

So far we have certainly assumed that the regression model contains a constant term. In this case, of course, the OLS estimator requires that the sum of the squared deviations from the sample mean is such that $n^{-1}\sum_{t=1}^{n}\left(L(t)-\bar{L}\right)^2 = n^{-1}\sum_{t=1}^{n} L(t)^2 - (\bar{L})^2$. Using Lemma 2.1, we observe that the first and second asymptotically leading terms of the difference offset each other and only the third term survives. Therefore, the asymptotic order of this is $O(L(n)^2\varepsilon(n)^2)$, which fluctuates greatly. One solution may be obtained by avoiding such computation.

4.2 Improvement of the finite sample performance

In this subsection, we consider the regression without the constant term. The model having no constant term is defined as:

$$y_t = \beta L(t) + u_t \quad \text{for} \quad t = 1, \ldots, n, \tag{7}$$

with the errors satisfying Assumption 2.3. The following results are parallel to those in the preceding section.

Theorem 4.1 *Under Assumption 2.1 and 2.3, the OLS estimator of (7) is asymptotically distributed as follows:*

$$\frac{L(n)}{\sqrt{n}}\left(\hat{\beta} - \beta\right) \xrightarrow{d} N\left(0, \frac{\sigma_L^2}{3}\right).$$

Lemma 4.1 *Under Assumption 2.1, 2.3 and the model (7), we have, as $n \to \infty$,*

$$\begin{aligned}
(i) \quad & \frac{1}{n}\hat{u}_n^2 \xrightarrow{d} \sigma_L^2 U_2, \\
(ii-a) \quad & \frac{L(n)^2}{n}\hat{u}_1^2 = O_p(1) \quad if \quad L(n) \to 0, \\
(ii-b) \quad & \frac{1}{n}\hat{u}_1^2 = o_p(1) \quad if \quad L(n) \to \infty, \\
(iii) \quad & \frac{1}{n}\sum_{t=2}^{n}(\hat{u}_t - \hat{u}_{t-1})^2 \xrightarrow{p} \sigma_S^2, \\
(iv) \quad & \frac{1}{n^2}\sum_{t=1}^{n}\hat{u}_t^2 \xrightarrow{d} \sigma_L^2 V_2,
\end{aligned}$$

where

$$U_2 = \left\{w(1) - \int_0^1 w(r)dr\right\}^2 \quad and \quad V_2 = \int_0^1 w(r)^2 dr - \left(\int_0^1 w(r)dr\right)^2.$$

Theorem 4.2 *Under Assumption 2.1, 2.3 and the model (7), if we further suppose* $L(t) \to \infty$ *as* $n \to \infty$*, then it follows that, as* $n \to \infty$*,*

$$n\,(\hat{\rho} - 1) \xrightarrow{d} \frac{U_2 - \sigma_S^2/\sigma_L^2}{2V_2}.$$

Corollary 4.1 *Under Assumption 2.1, 2.3 and the model (7), if we further suppose* $L(t) \to \infty$ *as* $n \to \infty$*, then it follows that, as* $n \to \infty$*,*

$$t_{\hat{\rho}} \xrightarrow{d} \frac{\sigma_L}{\sigma_S} \frac{U_2 - \sigma_S^2/\sigma_L^2}{2\sqrt{V_2}}.$$

Unfortunately, it may seldom be appropriate to suppose that the true process has no constant term. However, it is worth analyzing the situation where the true process is given by (5), which possesses a constant term, but the no-constant model (7) is employed for regression. Then, we find that it works well asymptotically from the following fact.

Theorem 4.3 *Under Assumption 2.1 and Assumption 2.3, we suppose that the true data-generating process (DGP) is given by (5), but (7) is employed for regression. then we still have the same asymptotic result given in Theorem 4.1 with the effect of a constant term declining at the rate* $O(n^{-1/2})$.

We note that under the assumptions in Theorem 4.3, the test statistics constructed by the same way also have the same asymptotic behavior as in Theorem 4.2 and Corollary 4.1. This suggests that, even if the true process includes a nonzero constant term, regression without it may still be beneficial under integrated errors, and may provide a good test statistic. We study the properties of such a test statistic in the rest of this subsection.

Remark 4.1 If we let $\hat{\sigma}_S^2 = n^{-1} \sum_{t=2}^{n} (\hat{u}_t - \hat{u}_{t-1})^2$ in Lemma 4.1, which is constructed based on the philosophy of Theorem 4.3, $\hat{\sigma}_S^2$ is used for a consistent estimator of σ_S^2.

Remark 4.2 If we use the regression model $y_t = \alpha + u_t$ against the true process (5), a similar result is obtained. In this case, however, the declining rate of the irrelevant constant term becomes $O_p(L(n)n^{-1/2})$, which is slightly slower than that obtained in Theorem 3.5, $O_p(n^{-1/2})$.

Using the no-constant model (7) for regression with (5) being the DGP (α is set to unity for the simulation), we have Figures 3 and 4 and Table 3 under Assumption 2.3 with v_t i.i.d., just as in the preceding subsection. In this case, we can see that their finite sample approximation seems to be significantly improved.

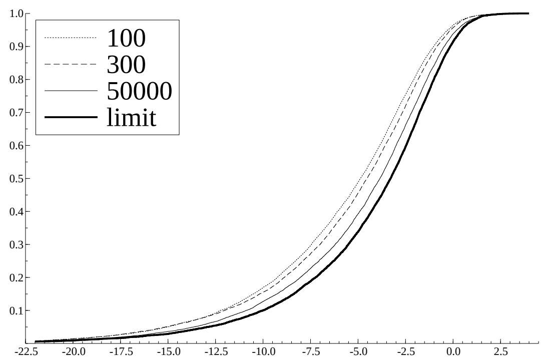

Figure 3: CDF of $n(\hat{\rho} - 1)$ based on (7).

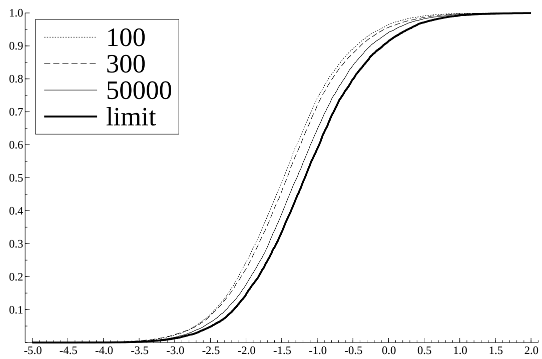

Figure 4: CDF of $t_{\hat{\rho}}$ based on (7).

Table 3: Percentage points of the limiting distributions.

	Probability (%)							
Statistic	1.0	2.5	5.0	10.0	90.0	95.0	97.5	99.0
$n(\hat{\rho} - 1)$	-19.69	-15.88	-12.91	-10.02	-0.15	0.45	0.95	1.45
$t_{\hat{\rho}}$	-3.07	-2.75	-2.48	-2.17	-0.08	0.27	0.55	0.91

We see that we obtain good performances for the powers and sizes of these test statistics. Table 4 indicates the empirical sizes and Figures 5 and 6 show the size-adjusted powers. They illustrate that these tests are suitable for practical use, despite a little size distortion remaining. Hence, we conclude that the influence of an SV regressor on the test statistics can be removed and such statistics are recommended.

Table 4: Empirical size of tests (%).

	# observations			
Statistic	100	200	300	500
$n(\hat{\rho}-1)$	8.8	8.6	8.1	7.6
$t_{\hat{\rho}}$	11.3	9.6	8.9	8.2

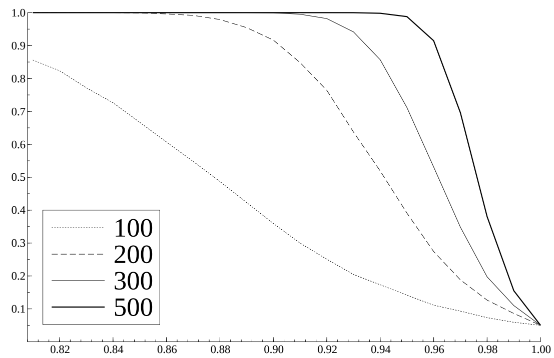

Figure 5: Power of DF-$\hat{\rho}$ with (7).

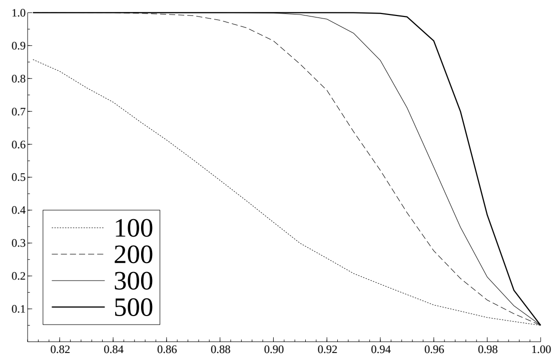

Figure 6: Power of DF-t with (7).

4.3 Generalization of the unit root test statistics

The weak convergence results in the preceding section have nuisance parameters in their limit. Finally, we introduce Phillips and Perron (1988) (PP)-type statistics under assumption 2.3 based on the discussion in the preceding subsection.

Theorem 4.4 *Construct the following two statistics based on the regression model (7) against the natural DGP (5):*

$$Z_{\hat{\rho}} = n(\hat{\rho}-1) - \frac{n^2(\hat{\sigma}_L^2 - \hat{\sigma}_S^2)}{2\sum_{t=1}^n \hat{u}_{t-1}^2} \quad \textit{and} \quad Z_t = \frac{\hat{\sigma}_S}{\hat{\sigma}_L} t_{\hat{\rho}} - \frac{n(\hat{\sigma}_L^2 - \hat{\sigma}_S^2)}{2\hat{\sigma}_L \sqrt{\sum_{t=1}^n \hat{u}_{t-1}^2}},$$

where

$$\hat{\sigma}_S^2 = \frac{1}{n}\sum_{t=2}^{n}(\Delta\hat{u}_t)^2 \quad and \quad \hat{\sigma}_L^2 = \hat{\sigma}_S^2 + \frac{2}{n}\sum_{j=1}^{k}\left(1-\frac{j}{k+1}\right)\sum_{t=j+1}^{n}\Delta\hat{u}_t\Delta\hat{u}_{t-j}$$

with $k = o(n^{1/4})$. *Then, under Assumption 2.1 and 2.3, it follows that, as* $n \to \infty$,

$$Z_{\hat{\rho}} \xrightarrow{d} \frac{U_2 - 1}{2V_2} \quad and \quad Z_t \xrightarrow{d} \frac{U_2 - 1}{2\sqrt{V_2}}.$$

These limiting distributions have the same percentage points of $n(\hat{\rho} - 1)$ *and* $t_{\hat{\rho}}$, *respectively, in Table 3.*

Remark 4.3 It may be possible to reduce the size distortion caused by the estimation of the long-run variance by using a method such as that of Perron and Ng (1996). However, this is beyond the scope of this paper, and is left to future studies.

5 Conclusions

We have studied the model with an SV regressor in the presence of integrated errors and found three main results. First, regression coefficients are asymptotically normally distributed, but they are not consistent. We thus observe that there is a contrast between a simple trend and an SV trend. Second, ordinal unit root test statistics based on AR(1) errors behave badly due to the existence of an SV regressor, and it is not recommended that tests of hypotheses be based on them. This may be caused by asymptotic collinearity of the regressors. Third, in spite of this difficulty, correction of sizes is possible by eliminating the constant term from the regression model. It is shown that the statistics constructed by this approach are still consistent even when the DGP has a constant term. Applying this result, we finally give PP-type test statistics.

References

[1] Abramowitz, M. and I. A. Stegun (1972), *Handbook of Mathematical Functions*, Dover.

[2] Bingham, N. H., C. M. Goldie and J. L. Teugels (1987), *Regular Variation*, Cambridge University Press.

[3] Buzel, H. and T. J. Vogelsang (2005), "Powerful trend function tests that are robust to strong serial correlation, with an application to the Prebisch–Singer hypothesis," *Journal of the American Statistical Association*, **23**, 381–394.

[4] Canjels, E. and M. W. Watson (1997), "Estimating deterministic trends in the presence of serially correlated errors," *The Review of Economics and Statistics*, **79**, 184–200.

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

[6] Doornik, J. A. (2007), *Object-Oriented Matrix Programming Using Ox*, 3rd ed. London: Timberlake Consultants Press and Oxford: www.doornik.com/

[7] Durlauf, S. N. and P. C. B. Phillips (1988), "Trends versus random walks in time series analysis," *Econometrica*, **56**, 1333–1354.

[8] Grenander, U. and M. Rosenblatt (1957), *Statistical Analysis of Stationary Time Series,* Wiley.

[9] Hamilton, J. D. (1994), *Time Series Analysis*, Princeton.

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

[11] Newey, W. and K. West (1987), "A simple, positive semidefinite, heteroskedasticity and autocorrelation consistent covariance matrix," *Econometrica*, **55**, 703–705.

[12] Perron, P. and S. Ng (1996), "Useful modifications to some unit root tests with dependent errors and their local asymptotic properties," *Review of Economic studies*, **63**, 435–63.

[13] Perron, P., and T. Yabu (2009), "Estimating deterministic trends with an integrated or stationary noise component," *Journal of Econometrics*, **151**, 56–69.

[14] Phillips, P. C. B. (1987), "Time series regression with a unit root," *Econometrica*, **55**, 277–301.

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

[16] Phillips, P. C. B. and P. Perron (1988), "Testing for a unit root in time series regression," *Biometrika*, **75**, 335–346.

[17] Phillips, P. C. B. and Y. Sun (2003), "Regression with an evaporating logarithmic trend," *Econometric Theory*, **19**, 692–701.

[18] Tanaka, K. (1996), *Time Series Analysis: Nonstationary and Noninvertible Distribution Theory*, Wiley.

[19] Vogelsang, T. (1998), "Trend function hypothesis testing in the presence of serial correlation," *Econometrica*, **66**, 123–148.

APPENDIX A: A Lemma for Proofs

Lemma A *Assume that L(t) satisfies assumption SSV. For sufficiently large n, it follows that*

$$\sum_{t=2}^{n}(L(t)-L(t-1))^2 < L(n)^2\varepsilon(n)^2[1+o(1)].$$

Proof of Lemma A We start with the fact called second-order SV in Phillips (2007, eq. (10)) and, more precisely, in de Haan and Resnick (1996). This implies that for sufficiently large n,

$$L(nr) = L(n) + L(n)\varepsilon(n)\log r[1+o(1)] \tag{8}$$

for arbitrary $r > 0$. If we set $r = (t-1)/n$ and $r = t/n$ in (8), $(L(t) - L(t-1))^2$ can be evaluated as

$$\begin{aligned}(L(t)-L(t-1))^2 &= L(n)^2\varepsilon(n)^2[1+o(1)]^2\left(\log\left(\frac{t}{n}\right)-\log\left(\frac{t-1}{n}\right)\right)^2\\ &= L(n)^2\varepsilon(n)^2[1+o(1)]\log^2\left(1-\frac{1}{t}\right)\end{aligned}$$

The last logarithmic function has the representation so that

$$\begin{aligned}\log^2\left(1-\frac{1}{t}\right) &= \left(-\frac{1}{t}-\frac{1}{2t^2}-\frac{1}{3t^3}-\dots\right)^2\\ = \frac{1}{t^2}+\frac{1}{t^3}+\frac{11}{12t^4}+\frac{5}{6t^5}+\dots &< \sum_{k=2}^{\infty}\frac{1}{t^k} = \frac{1}{t(t-1)}.\end{aligned}$$

Hence, we have:

$$(L(t)-L(t-1))^2 < L(n)^2\varepsilon(n)^2[1+o(1)]\frac{1}{t(t-1)}$$

for sufficiently large n, so that

$$\begin{aligned}\sum_{t=2}^{n}(L(t)-L(t-1))^2 &< L(n)^2\varepsilon(n)^2[1+o(1)]\sum_{t=2}^{n}\frac{1}{t(t-1)}\\ &= L(n)^2\varepsilon(n)^2[1+o(1)]\left(1-\frac{1}{n}\right)\\ &= L(n)^2\varepsilon(n)^2[1+o(1)]. \quad \blacksquare\end{aligned}$$

APPENDIX B: Proofs for Sections 3 and 4

Proof of Lemma 3.1 For (i), we make a decomposition as:

$$\frac{1}{n\sqrt{n}L(n)}\sum_{t=1}^{n}L(t)u_t = \frac{1}{n}\sum_{t=1}^{n}\frac{u_t}{\sqrt{n}} - \frac{1}{nL(n)}\sum_{t=2}^{n}\sum_{s=1}^{t}\frac{u_s}{\sqrt{n}}(L(t)-L(t-1)). \tag{9}$$

Assumption 2.3 enables us to use the strong approximation of $u_t/\sqrt{n}$ for Brownian motion as in (4). Hence, the first term of (9) is

$$\frac{1}{n}\sum_{t=1}^{n}\frac{u_t}{\sqrt{n}} = \sigma_L\int_0^1 w(s)ds + o_p(1). \tag{10}$$

The second term of (9) is

$$\begin{aligned}
&\frac{1}{nL(n)}\sum_{t=2}^{n}\sum_{s=1}^{t}\frac{u_s}{\sqrt{n}}\left(L\left(n\frac{t}{n}\right)-L\left(n\frac{t-1}{n}\right)\right)\\
&=\frac{\sigma_L}{L(n)}\int_0^1\int_0^r w(s)ds\,dL(nr)+o_p(1)\\
&=\frac{\sigma_L}{L(n)}\int_0^1\frac{F_1(r)}{r}\frac{nrL'(nr)}{L(nr)}L(nr)dr+o_p(1)\\
&=\frac{\sigma_L}{L(n)}\int_0^1\frac{F_1(r)}{r}L(nr)\varepsilon(nr)dr+o_p(1)\\
&=\varepsilon(n)\sigma\int_0^1\frac{F_1(r)}{r}dr+o_p(1)=O_p(\varepsilon(n))+o_p(1)=o_p(1),
\end{aligned} \tag{11}$$

where $F_1(r)$ is a one-fold integrated Brownian motion defined by $\int_0^r w(s)ds$. Note that the fourth equality in (11) holds from the fact in Phillips (2007, eq. (74)). In consequence, (10) and (11) give the result.

For (ii), we basically take the same manipulation as (i). Decomposing the product

$L(t)u_t$ as in (9) and remembering the asymptotic expansion of $\bar{L}$, we have:

$$\begin{aligned}
&\frac{1}{n\sqrt{n}}\sum_{t=1}^{n}(L(t)-\bar{L})u_t \\
&= \frac{L(n)}{n}\sum_{t=1}^{n}\frac{u_t}{\sqrt{n}} - \frac{1}{n}\sum_{t=2}^{n}\sum_{s=1}^{t}\frac{u_s}{\sqrt{n}}\left(L(t)-L(t-1)\right) \\
&\qquad - [L(n) - L(n)\varepsilon(n) + O(L(n)\varepsilon(n)^2)]\frac{1}{n}\sum_{t=1}^{n}\frac{u_t}{\sqrt{n}} \qquad (12)\\
&= -\frac{1}{n}\sum_{t=2}^{n}\sum_{s=1}^{t}\frac{u_s}{\sqrt{n}}\left(L(t)-L(t-1)\right) + L(n)\varepsilon(n)[1+O(\varepsilon(n))]\frac{1}{n}\sum_{t=1}^{n}\frac{u_t}{\sqrt{n}}.
\end{aligned}$$

From (i), (12) leads asymptotically to

$$\begin{aligned}
&-\sigma_L\int_0^1 F_1(r)dL(nr) + L(n)\varepsilon(n)\left(\sigma_L\int_0^1 w(r)dr + O_p(\varepsilon(n))\right) + o_p(1) \\
&= L(n)\varepsilon(n)\sigma_L\left(-\int_0^1\frac{F_1(r)}{r}dr + \int_0^1 w(r)dr + O_p(\varepsilon(n))\right) + o_p(1) \\
&= L(n)\varepsilon(n)\sigma_L\int_0^1\left(w(r)-\frac{F_1(r)}{r}\right)dr[1+O_p(\varepsilon(n))] + o_p(1).
\end{aligned}$$

Note that F_1 is differentiable in mean square, so that we have:

$$\sigma_L\int_0^1\frac{F_1(r)}{r}dr = -\sigma_L\int_0^1\log r\, w(r)dr.$$

Thus, we conclude that (12) divided by $L(n)\varepsilon(n)$ reduces asymptotically to:

$$\begin{aligned}
&\sigma_L\int_0^1\left(w(r)-\frac{F_1(r)}{r}\right)dr = \sigma_L\int_0^1(\log r+1)\,w(r)dr \\
&= -\sigma_L\int_0^1 r\log r dw(r) \stackrel{d}{=} N\left(0,\sigma_L^2\int_0^1 r^2\log^2 r dr\right) = N\left(0,\frac{2\sigma_L^2}{27}\right). \qquad \blacksquare
\end{aligned}$$

Proof of Lemma 3.1 The marginal limiting distribution of $\hat{\beta}$ is clear from Lemma 3.1. The remaining claim is the distribution of $\hat{\alpha}$ and their joint behavior. It is easy to see that:

$$\begin{aligned}
\frac{\varepsilon(n)}{\sqrt{n}}(\hat{\alpha}-\alpha) &= \frac{\varepsilon(n)}{n\sqrt{n}}\sum_{t=1}^{n}u_t - \frac{\varepsilon(n)}{\sqrt{n}}\frac{\bar{L}\sum_{t=1}^{n}L(t)u_t - \bar{L}^2\sum_{t=1}^{n}u_t}{\sum_{t=1}^{n}(L(t)-\bar{L})^2} \\
&= O_p(\varepsilon(n)) - \frac{\varepsilon(n)\bar{L}}{\sqrt{n}}(\hat{\beta}-\beta) \\
&= o_p(1) - \frac{L(n)\varepsilon(n)}{\sqrt{n}}(\hat{\beta}-\beta)[1+O(\varepsilon(n))] \\
&= -\frac{L(n)\varepsilon(n)}{\sqrt{n}}(\hat{\beta}-\beta) + o_p(1).
\end{aligned}$$

Thus, symmetry on the origin of the limiting normal distribution of $\hat{\beta} - \beta$ implies the desired result. ■

Proof of Lemma 3.2 (i) From (6) and Theorem 3.1, applying the functional Central Limit Theorem, together with the continuous mapping theorem, yields:

$$\begin{aligned}
\frac{1}{n}\hat{u}_n^2 &= \left\{ \frac{u_n}{\sqrt{n}} - \left(\frac{1}{n} \sum_{t=1}^{n} \frac{u_t}{\sqrt{n}} \right) - (L(n) - \bar{L})(\hat{\beta} - \beta) \right\}^2 \\
&= \left\{ \frac{u_n}{\sqrt{n}} - \left(\frac{1}{n} \sum_{t=1}^{n} \frac{u_t}{\sqrt{n}} \right) - \frac{L(n)\varepsilon(n)[1 + O(\varepsilon(n))]}{\sqrt{n}}(\hat{\beta} - \beta) \right\}^2 \\
&\xrightarrow{d} \sigma_L^2 \left\{ w(1) - \int_0^1 w(r)dr - \int_0^1 \left(w(r) - \frac{F_1(r)}{r} \right) dr \right\}^2 .
\end{aligned}$$

as stated. ■

(ii) By a similar manner to the proof of Lemma 3.2, we obtain:

$$\begin{aligned}
\frac{1}{n}\hat{u}_1^2 &= \left\{ \frac{u_1}{\sqrt{n}} - \left(\frac{1}{n} \sum_{t=1}^{n} \frac{u_t}{\sqrt{n}} \right) - \frac{L(1) - \bar{L}}{\sqrt{n}}(\hat{\beta} - \beta) \right\}^2 \\
&= \left\{ O_p\left(\frac{1}{\sqrt{n}} \right) - O_p(1) - \frac{L(1) - L(n) + L(n)\varepsilon(n)[1 + O(\varepsilon(n))]}{\sqrt{n}}(\hat{\beta} - \beta) \right\}^2 .
\end{aligned} \tag{13}$$

This rate of convergence depends on the asymptotic behavior of $L(n)$. That is, if $L(n) \to 0$, (13) times $L(n)^2\varepsilon(n)^2$ is $O_p(1)$ so that

$$\begin{aligned}
\frac{L(n)^2\varepsilon(n)^2}{n}\hat{u}_1^2 &= \left\{ o_p(1) - L(1)\frac{L(n)\varepsilon(n)}{\sqrt{n}}(\hat{\beta} - \beta)(1 + o_p(1)) \right\}^2 \\
&\xrightarrow{d} L(1)^2\sigma_L^2 \left\{ \int_0^1 \left(w(r) - \frac{F_1(r)}{r} \right) dr \right\}^2 .
\end{aligned} \tag{14}$$

If $L(n) \to \infty$ to the contrary, (13) times $\varepsilon(n)^2$ is $O_p(1)$. In fact, we obtain:

$$\begin{aligned}
\frac{\varepsilon(n)^2}{n}\hat{u}_1^2 &= \left\{ o_p(1) + \frac{L(n)\varepsilon(n)}{\sqrt{n}}(\hat{\beta} - \beta)(1 + o_p(1)) \right\}^2 \\
&\xrightarrow{d} \sigma_L^2 \left\{ \int_0^1 w(r)dr - \int_0^1 \frac{F_1(r)}{r}dr \right\}^2 .
\end{aligned} \tag{15}$$

Hence, (14) and (15) give the conclusion. ■

(iii) We have:

$$\begin{aligned}
&\frac{1}{n}\sum_{t=2}^{n}(\hat{u}_t-\hat{u}_{t-1})^2 \\
&=\frac{1}{n}\sum_{t=2}^{n}v_t^2-\frac{2(\hat{\beta}-\beta)}{\sqrt{n}}\sum_{t=2}^{n}\frac{v_t}{\sqrt{n}}\Delta L(t)+\left(\frac{\hat{\beta}-\beta}{\sqrt{n}}\right)^2\sum_{t=2}^{n}(\Delta L(t))^2.
\end{aligned} \tag{16}$$

The first term converges to σ_S^2 in probability. For the second term, its absolute value is less than or equal to:

$$2\left|\frac{\hat{\beta}-\beta}{\sqrt{n}}\right|\left(\frac{1}{n}\sum_{t=2}^{n}v_t^2\right)^{1/2}\left(\sum_{t=2}^{n}(\Delta L(t))^2\right)^{1/2}=O_p(1) \tag{17}$$

by Schwarz inequality and Lemma A. The last term in (16) is similarly found to be $O_p(1)$. We therefore conclude that (16) is $O_p(1)$. ■

(iv) If we use $\sum_{t=1}^{n}L(t)=nL(n)[1+O(\varepsilon(n))]$, $\sum_{t=1}^{n}L(t)^2=nL(n)^2[1+O(\varepsilon(n))]$ and Lemma 3.1 (i), it follows that:

$$\begin{aligned}
\frac{1}{n^2}\sum_{t=1}^{n}\hat{u}_t^2&=\frac{1}{n}\sum_{t=1}^{n}\left(\frac{u_t}{\sqrt{n}}-\left(\frac{1}{n}\sum_{s=1}^{n}\frac{u_s}{\sqrt{n}}\right)-\frac{\hat{\beta}-\beta}{\sqrt{n}}(L(t)-\bar{L})\right)^2\\
&=\frac{1}{n}\sum_{t=1}^{n}\left(\frac{u_t}{\sqrt{n}}\right)^2+\left(\frac{1}{n}\sum_{s=1}^{n}\frac{u_s}{\sqrt{n}}\right)^2+\left(\frac{\hat{\beta}-\beta}{\sqrt{n}}\right)^2\frac{1}{n}\sum_{t=1}^{n}(L(t)-\bar{L})^2\\
&\quad-2\left(\frac{1}{n}\sum_{t=1}^{n}\frac{u_t}{\sqrt{n}}\right)^2-\frac{L(n)\varepsilon(n)(\hat{\beta}-\beta)}{\sqrt{n}}\frac{2}{n\sqrt{n}L(n)\varepsilon(n)}\sum_{t=1}^{n}(L(t)-\bar{L})u_t\\
&=\frac{1}{n}\sum_{t=1}^{n}\left(\frac{u_t}{\sqrt{n}}\right)^2-\left(\frac{1}{n}\sum_{s=1}^{n}\frac{u_s}{\sqrt{n}}\right)^2-\left(\frac{L(n)\varepsilon(n)(\hat{\beta}-\beta)}{\sqrt{n}}\right)^2[1+O(\varepsilon(n))]\\
&\xrightarrow{d}\sigma_L^2\int_0^1 w(r)^2dr-\sigma_L^2\left(\int_0^1 w(r)dr\right)^2-\sigma_L^2\left(\int_0^1\left(w(r)-\frac{F_1(r)}{r}\right)dr\right)^2,
\end{aligned}$$

which completes the proof. ■

Proof of Theorem 3.2 and Corollary 3.1 It is clear from Lemma 3.2. ■

Proof of Theorem 4.1 The result follows from Lemma 3.1. ■

Proof of Lemma 4.1 (i) Note that

$$\hat{u}_t = u_t - (\hat{\beta} - \beta)L(t)$$

under the model (7). Following from Theorem 3.3, the result can be derived:

$$\frac{1}{n}\hat{u}_n^2 = \left(\frac{1}{\sqrt{n}}u_n - \frac{L(n)}{\sqrt{n}}\left(\hat{\beta} - \beta\right)\right)^2 \xrightarrow{d} \sigma_L^2\left(w(1) - \int_0^1 w(r)dr\right)^2. \quad \blacksquare$$

(ii) It follows from $u_0 = 0$ and Theorem 3.3 that:

$$\frac{1}{n}\hat{u}_1^2 = \left(\frac{1}{\sqrt{n}}u_1 - \frac{L(1)}{\sqrt{n}}\left(\hat{\beta} - \beta\right)\right)^2 = \left(O_p\left(n^{-1/2}\right) - O_p\left(L(n)^{-1}\right)\right)^2.$$

Hence, this converges to zero in probability if $L(n)$ tends to infinity, whereas if $L(n) \to 0$, we have:

$$\frac{L(n)^2}{n}\hat{u}_1^2 = \left(\frac{L(n)}{\sqrt{n}}u_1 - \frac{L(1)L(n)}{\sqrt{n}}\left(\hat{\beta} - \beta\right)\right)^2 \xrightarrow{d} \sigma_L^2\left(-L(1)\int_0^1 w(r)dr\right)^2. \quad \blacksquare$$

(iii) Since $\hat{\beta}$ is not $O_p(\sqrt{n}/(L(n)\varepsilon(n)))$ but $O_p(\sqrt{n}/L(n))$ under the model (7), the second term of (16) becomes $O_p(\varepsilon(n))$ by Schwarz inequality, as in (17). The third term of (16) is similarly found to be $O_p(\varepsilon(n)^2)$. Therefore, we have:

$$\frac{1}{n}\sum_{t=2}^{n}(\hat{u}_t - \hat{u}_{t-1})^2 \xrightarrow{p} \sigma_S^2. \quad \blacksquare$$

(iv) It follows from Lemma 3.1 (i) and Theorem 3.1 that:

$$\begin{aligned}\frac{1}{n^2}\sum_{t=1}^{n}\hat{u}_t^2 = \frac{1}{n}\sum_{t=1}^{n}\left(\frac{u_t}{\sqrt{n}}\right)^2 - 2\frac{L(n)}{\sqrt{n}}\left(\hat{\beta} - \beta\right)\frac{1}{n\sqrt{n}L(n)}\sum_{t=1}^{n}L(t)u_t \\ + \frac{L(n)^2}{n}\left(\hat{\beta} - \beta\right)^2[1 + O(\varepsilon(n))] \\ \xrightarrow{d} \sigma_L^2\int_0^1 w(r)^2dr - 2\sigma_L^2\left(\int_0^1 w(r)dr\right)^2 + \sigma_L^2\left(\int_0^1 w(r)dr\right)^2.\end{aligned}$$

Collecting terms gives the result. ■

Proof of Theorem 4.2 and Corollary 4.1 It is clear from Lemma 4.1. ■

Proof of Theorem 4.3 We consider the following situation:

$$\text{True DGP :} \qquad y_t = \alpha + \beta L(t) + u_t,$$

$$\text{Regression :} \qquad y_t = \beta L(t) + u_t.$$

Then, the OLS estimator is:

$$\begin{aligned}
\hat{\beta} = \frac{\sum_{t=1}^n L(t) y_t}{\sum_{t=1}^n L(t)^2} &= \alpha \frac{\sum_{t=1}^n L(t)}{\sum_{t=1}^n L(t)^2} + \beta + \frac{\sum_{t=1}^n L(t) u_t}{\sum_{t=1}^n L(t)^2} \\
&= \alpha \frac{nL(n)[1 + O(\varepsilon(n))]}{nL(n)^2[1 + O(\varepsilon(n))]} + \beta + O_p\left(\frac{\sqrt{n}}{L(n)}\right) \\
&= \alpha \frac{1 + o(1)}{L(n)[1 + o(1)]} + \beta + O_p\left(\frac{\sqrt{n}}{L(n)}\right),
\end{aligned}$$

by Lemma 3.1 and Theorem 3.1. Collecting terms and scaling by $L(n)/\sqrt{n}$, we have $n^{-1/2}L(n)(\hat{\beta} - \beta) = O(n^{-1/2}) + O_p(1)$. Thus, it leads to the same limiting result as in Theorem 3.1 (i). The conclusions for the other statistics in this situation are also derived in the same way via Lemma 4.1. ■

Proof of Theorem 4.4 It suffices to prove the consistency of the estimated long-run variance $\hat{\sigma}_L^2$. Because the last three terms of

$$\begin{aligned}
\frac{1}{n} \sum_{t=j+1}^n \Delta \hat{u}_t \Delta \hat{u}_{t-j} = \frac{1}{n} \sum_{t=j+1}^n \{ & v_t v_{t-j} - (\hat{\beta} - \beta) v_t \Delta L(t-j) \\
& - (\hat{\beta} - \beta) v_{t-j} \Delta L(t) + (\hat{\beta} - \beta)^2 \Delta L(t) \Delta L(t-j) \}
\end{aligned}$$

are $o_p(1)$ from Schwarz inequality as in Lemma 3.2 (iii). Therefore, we have

$$\frac{1}{n} \sum_{t=j+1}^n \{ \Delta \hat{u}_t \Delta \hat{u}_{t-j} - \Delta u_t \Delta u_{t-j} \} = o_p(1).$$

Combining the fact that $\hat{\sigma}_S^2 - \sigma_S^2 = o_p(1)$, we conclude that $\hat{\sigma}_L^2$ is consistent for σ_L^2 from Theorem 2 in Newey and West (1987). $\blacksquare$