Spurious Regression and Residual-Based Tests for Cointegration in Panel Data When the Cross-Section and Time-Series Dimensions are Comparable

Chihwa Kao*
Center for Policy Research
and Department of Economics
Syracuse University
Syracuse, NY 13244
(315) 443-3233
Fax: (315) 443-1081
Email: cdkao@maxwell.syr.edu

First Draft: January 10, 1995
This Draft: September 2, 1996

Abstract

In the first half of the paper we study spurious regressions in panel data when the cross-section and time-series dimensions are comparable. Asymptotic properties of the least-squares dummy variable (LSDV) estimator and other conventional statistics are examined. We show that the LSDV estimator, $\widehat{\beta}$, is consistent for its true value, but the t-statistic, t_β, diverges so that inferences about the regression coefficient, β, are wrong with the probability that goes to one as $N \to \infty$ and $T \to \infty$. The asymptotics of $\widehat{\beta}$ are also different from those of the spurious regression in the pure time series. This has an important consequence for residual-based cointegration tests in panel data, because the null distribution of residual-based cointegration tests depends on the asymptotics of $\widehat{\beta}$.

In the second half of the paper we study residual-based tests for cointegration regression in panel data. We study Dickey-Fuller (DF) tests and an augmented Dickey-Fuller (ADF) test to test the null of no cointegration. Asymptotic distributions of the tests are derived and Monte

---

[*] I thank two referees, an associate editor and an editor, Richard Blundell, for helpful comments. I also thank participants of workshop seminars at Hong Kong Univeristy of Science and Technology, Academia Sinica, and SUNY-Albany for helpful comments and suggestions. I thank Bangtian Chen for his research assistantce on an earlier draft of this paper. Thanks also go to Jodi Woodson and Martha Bonney for correcting my English and carefully checking the manuscript to enhance its readability. An electronic version of the paper in postscript format can be retrieved from http://web.syr.edu/~cdkao.

Carlo experiments are conducted to evaluate finite sample properties of the proposed tests. The simulation results suggest that the DF_ρ^* and DF_t^* tests have better size and power properties than the DF_ρ, DF_t, and ADF tests when σ is small. However, when σ is large, ADF clearly dominates the rest.

Key Words: *Panel Data; Spurious Regression; LSDV; Residual-Based Tests.*

1 Introduction

There is immense interest in testing for unit roots and cointegration in time-series data but not much attention has been paid to testing the unit roots and cointegration of panel data at either empirical or theoretical levels. Apparently, the only theoretical studies in this area are Breitung and Meyer (1994), Quah (1994), Levin and Lin (1992, 1993), Im, Pesaran and Shin (1995), Kao and Chen (1995a, b) and Pedroni (1995). Breitung and Meyer (1994) derived the asymptotic normality of the Dickey-Fuller test statistic for panel data with a large cross-section dimension and a small time-series dimension. Quah (1994) studied the unit roots tests for panel data that have simultaneous extensive cross-section and time-series variation. Levin and Lin (1992, 1993) recently derived the limiting distributions for unit roots on panel data and showed that the power of these tests increases dramatically as the cross-section dimension increases. Im et al. (1995) critiqued the Levin and Lin Panel unit root statistics and proposed alternatives. Kao and Chen (1995a) studied the asymptotic results for a least-squares dummy variable (LSDV) estimator in a cointegrated regression in panel data. They showed that the LSDV estimator and t-ratio are asymptotically normally distributed. They also confirmed the existence of second-order bias (e.g., due to endogeneity or errors-in-variables). Kao and Chen (1995b) proposed residual based tests for cointegration under a set of restricted assumptions.. On the other hand, Park and Ogaki (1991) derived asymptotic distributions for cointegration coefficient estimators and related t-statistics for panel data using canonical cointegrating regression (CCR) transformations. Although they take a seemingly unrelated regression (SUR) approach rather using N dimensional asymptotics, many of the issues are similar. Peasran and Smith (1995) are not directly concerned with cointegration but do touch on a number of related issues, including the potential problems of homogeneity misspecification for cointegrated panels.

The first half of this paper examines a spurious regression in panel data when the cross-section and time-series dimensions are comparable. Asymptotic properties of the LSDV estimator and other conventional statistics are examined. I show that the LSDV estimator, $\widehat{\beta}$, is consistent for its true value, but the t-statistic, t_β, diverges so that inferences about the regression coefficient, β, are wrong with the probability that goes to one as $N \to \infty$ and $T \to \infty$. The asymptotics of $\widehat{\beta}$ are also different from those of the spurious regression in the pure time series. This has an important consequence for residual-based cointegration tests in panel data, because the null distribution of residual-based cointegration tests depends on the asymptotics of $\widehat{\beta}$. The second half of the paper examines the asymptotic null distribution of residual-based cointegration tests in panel data.

This paper is related to recent work by Pedroni (1995) that came to my attention when this work was completed. Pedroni (1995) derived asymptotic distributions for residual based tests of cointegration for both homogenous and heterogenous panels. This paper discusses the fixed effect model, which Pedroni (1995) does not discuss.

Section 2 introduces the model and derives the asymptotic distributions of the LSDV estimator and various conventional statistics from the spurious regression in panel data. Section 3 derives the asymptotic distributions for the residual-based tests for the cointegration. Section 4 describes the estimation of long-run variances. Section 5 reports the simulation results. Concluding remarks are given in Section 6. The proofs in the text are collected in the Appendix.

A word on notation. We define $\Omega^{1/2}$ be any matrix such that $\Omega = \left(\Omega^{1/2}\right)\left(\Omega^{1/2}\right)'$. We use " $\Rightarrow$ " to denote weak convergence, $\xrightarrow{p}$ to denote convergence in probability, $[x]$ to denote the largest integer $\leq x$, and $I(1)$ to signify a time series that is integrated of order one.

2 Large Sample Asymptotics

In developing our theory, we first assume the cross-section dimension is a monotonic function of the time-series dimension so that the law of large numbers (e.g., Theorem 6.2, Billingsley, 1986) and the central limit theorem (e.g., Theorem 27.2, Billingsley, 1986) for triangular arrays can be applied.

2.1 Independent Random Walk

Let y_{it} and x_{it} be independent random walks for all i,

$$y_{it} = y_{it-1} + u_{it}$$

and

$$x_{it} = x_{it-1} + \varepsilon_{it} \tag{1}$$

in which $y_{i0} = 0, x_{i0} = 0$.

Suppose these two unrelated $I(1)$ series, y_{it} and x_{it}, are incorrectly estimated by least squares for all i using panel data; the spurious LSDV regression model is

$$y_{it} \;=\; \alpha_i \;+\; \beta x_{it} \;+\; e_{it},\ i=1,...,N,\ t=1,...,T. \tag{2}$$

The LSDV estimators of α_i and β are

$$\widehat{\beta}=\frac{\sum_{i=1}^{N}\sum_{t=1}^{T}y_{it}(x_{it}-\overline{x}_i)}{\sum_{i=1}^{N}\sum_{t=1}^{T}(x_{it}-\overline{x}_i)^2}$$

and

$$\widehat{\alpha}_i=\overline{y}_i-\widehat{\beta}\overline{x}_i. \tag{3}$$

Thus,

$$\begin{aligned}\widehat{\beta} &= \frac{\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T^2}\sum_{t=1}^{T}(x_{it}-\overline{x}_i)y_{it}}{\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T^2}\sum_{t=1}^{T}(x_{it}-\overline{x}_i)^2}\\ &= \frac{\frac{1}{N}\sum_{i=1}^{N}\zeta_{1iT}}{\frac{1}{N}\sum_{i=1}^{N}\zeta_{2iT}}\\ &= \frac{\xi_{1NT}}{\xi_{2NT}},\end{aligned} \tag{4}$$

where $\overline{x}_i=\frac{1}{T}\sum_{t=1}^{T}x_{it}$, $\overline{y}_i=\frac{1}{T}\sum_{t=1}^{T}y_{it}$, $\zeta_{1iT}=\frac{1}{T^2}\sum_{t=1}^{T}(x_{it}-\overline{x}_i)\,y_{it}$, $\zeta_{2iT}=\frac{1}{T^2}\sum_{t=1}^{T}(x_{it}-\overline{x}_i)^2$, $\xi_{1NT}=\frac{1}{N}\sum_{i=1}^{N}\zeta_{1iT}$, and $\xi_{2NT}=\frac{1}{N}\sum_{i=1}^{N}\zeta_{2iT}$. For each $0\leq r\leq 1$, we define $[rT]$ as the integer part of rT. $W_i(r)$ and $V_i(r)$ are independent standard Wiener processes obtained from ε_{it} and u_{it}, respectively, i.e.,

$$\frac{1}{\sigma_\varepsilon\sqrt{T}}\sum_{t=1}^{[rT]}\varepsilon_{it}\Rightarrow W_i(r) \tag{5}$$

and

$$\frac{1}{\sigma_u\sqrt{T}}\sum_{t=1}^{[rT]}u_{it}\Rightarrow V_i(r) \tag{6}$$

for $0\leq r\leq 1$.

Assumption 1 *The* $(\varepsilon_{it},u_{it})^{'}$ *are assumed to be independent cross* i*.*

Remark 1 *It is true that the assumption of independence across i is pretty strong. However, the assumption of independence across i is needed in order to satisfy the requirement of the central limit theorem for triangle arrays. Moreover, as pointed out by Quah (1994), modeling cross-sectional dependence is involved because individual observations in cross-section have no natural ordering. It is possible to extend the current model to allow a degree of dependency across i using the approaches adapted by Im et al. (1995) and Pedroni (1995). However, it goes beyond the scope of this paper.*

Assumption 2 *$u_{it} \sim iid(0, \sigma_u^2)$, $\varepsilon_{it} \sim iid(0, \sigma_\varepsilon^2)$, $E\left(u_{it}^4\right) < \infty$, and $E\left(\varepsilon_{it}^4\right) < \infty$.*

Before going into first theorem of the paper, we need to consider some preliminary results. All limits in Lemmas 1 and 2 are taken as $T \to \infty$, except where otherwise noted.

Lemma 1 *If Assumptions 1 and 2 hold, then*

(a) $\zeta_{1iT} \Rightarrow \sigma_\varepsilon \sigma_u \int_0^1 W_i(r) V_i(r) dr - \left[\sigma_\varepsilon \int_0^1 W_i(r) dr\right] \left[\sigma_u \int_0^1 V_i(r) dr\right] \equiv \zeta_{1i}$,

(b) $\zeta_{2iT} \Rightarrow \sigma_\varepsilon^2 \left[\int_0^1 W_i^2(r) dr - \left\{\int_0^1 W_i(r) dr\right\}^2\right] \equiv \zeta_{2i}$,

(c) $E\left[\zeta_{1iT}\right] \equiv \mu_{1T} = E\left[\zeta_{1i}\right] \equiv \mu_1 = 0$,

(d) $E\left[\zeta_{2i}\right] \equiv \mu_2 = \frac{\sigma_\varepsilon^2}{6}$,

(e) $E\left[\zeta_{2iT}\right] \equiv \mu_{2T} = \mu_2 + O(T^{-1})$,

(f) $Var\left[\zeta_{1iT}\right] \equiv \sigma_{1T}^2 = \frac{\sigma_\varepsilon^2 \sigma_u^2}{90} + O(T^{-1})$,

(g) $Var\left[\zeta_{2iT}\right] \equiv \sigma_{2T}^2 = \frac{\sigma_\varepsilon^4}{45} + O(T^{-1})$,

(h) $Var\left[\zeta_{1i}\right] \equiv \sigma_1^2 = \frac{\sigma_\varepsilon^2 \sigma_u^2}{90}$,

(i) $Var\left[\zeta_{2i}\right] \equiv \sigma_2^2 = \frac{\sigma_\varepsilon^4}{45}$,

(j) $\xi_{1NT} \xrightarrow{p} \mu_1 = 0$,

(k) $\xi_{2NT} \xrightarrow{p} \mu_2 = \frac{\sigma_\varepsilon^2}{6}$.

All limits in Theorems 1 and 2 are taken as $N \to \infty$ and $T \to \infty$, except where otherwise noted. We are now in a position to state the following theorem:

Theorem 1 *Suppose the conditions of Lemma 1 are satisfied; then*

(a) $\widehat{\beta} \xrightarrow{p} 0$,

(b) $\sqrt{N}\widehat{\beta} \Rightarrow N\left(0, \frac{2}{5}\frac{\sigma_u^2}{\sigma_\varepsilon^2}\right)$ *or* $\widehat{\beta} = O_p(N^{-1/2})$,

(c) $T^{-1/2} t_\beta \Rightarrow N\left(0, \frac{2}{5}\right)$ *or* $t_\beta = O_p(T^{1/2})$,

(d) $R^2 \xrightarrow{p} 0$,

(e) $DW = O_p(T^{-1})$,

(f) $TDW \xrightarrow{p} 6$.

Theorem 1 shows that the LSDV estimator, $\widehat{\beta}$, is consistent for its true value, 0, but the t-statistic, t_β, diverges so that inferences about β are wrong with the probability that goes to one as $N \to \infty$ and $T \to \infty$. Moreover, R^2 and DW converge to zero in probability.

Theorem 1 will be extended to allow for much weaker assumptions, e.g., serial correlation, weak exogeneity, though independence still holds across all $i = 1, ..., N$ in the next section.

2.2 Relaxing Some Assumptions

The error process assumptions made in the Section 2.1 are quite restrictive. Indeed, assuming constant variances across i is also restrictive. The real job is how to verify the Lindeberg condition. The Lemma 4.2 (a version of central limit theorem (CLT) for triangular arrays) presented by Levin and Lin (1993) is not ready to derive the asymptotics when the variances are not constant across i. The difficulty is that uniform integrability is not guaranteed from the moment conditions in their lemma. Moreover, their conditions in the lemma do not yield the Lindberg condition. Hence, before technical issues can be resolved, I think the assumption of constant variances across i is an acceptable compromise for the purpose of this paper.

On other hand imposing strong exogeneity on the explanatory variable and serial independence on the error process is also quite unrealistic in practice. In this section, we relax these assumptions and derive the asymptotics. I will show in this section that the extension consists of correcting for both serial correlation and weak exogeneity, which can be done by replacing σ_v^2 with the long-run conditional variance σ_{0u}^2. Define $w_t = (u_{it}, \varepsilon_{it})'$. The long-run covariance matrix of w_{it} is given by

$$\Omega = \lim_{T\to\infty} \frac{1}{T} E\left(\sum_{t=1}^{T} w_{it}\right)\left(\sum_{t=1}^{T} w_{it}\right)' = \Sigma + \Gamma + \Gamma' \equiv \begin{bmatrix} \sigma_{0u}^2 & \sigma_{0u\varepsilon} \\ \sigma_{0u\varepsilon} & \sigma_{0\varepsilon}^2 \end{bmatrix}, \tag{7}$$

where

$$\Gamma = \lim_{T\to\infty} \frac{1}{T} \sum_{k=1}^{T-1} \sum_{t=k+1}^{T} E\left(w_{it} w_{it-k}'\right) \equiv \begin{bmatrix} \Gamma_u & \Gamma_{u\varepsilon} \\ \Gamma_{\varepsilon u} & \Gamma_{\varepsilon} \end{bmatrix} \tag{8}$$

and

$$\Sigma = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} E\left(w_{it} w_{it}'\right) \equiv \begin{bmatrix} \sigma_u^2 & \sigma_{u\varepsilon} \\ \sigma_{u\varepsilon} & \sigma_{\varepsilon}^2 \end{bmatrix}. \tag{9}$$

We assume that w_{it} satisfies the invariance principle of Theorem 2.1 of Phillips and Durflauf (1986), so that

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} w_{it} \Rightarrow \Omega^{1/2} \mathbf{W}_i(r) \text{ as } T \to \infty \tag{10}$$

where $\mathbf{W}(r)$ is a 2×1 dimensional standard Wiener process with

$$\mathbf{W}_i(r) = \begin{bmatrix} V_i(r) \\ W_i(r) \end{bmatrix}.$$

Assumption 3 *(e.g., Phillips,1986)*

1. $E\left[w_{it}\right] = 0$ *for all* t,
2. $\sup\limits_t E\left|w_{it}\right|^p$ *for some* $2 \leq p < \infty$,
3. $\{w_{it}\}_1^\infty$ *is strong mixing with mixing number* η_m *satisfying* $\sum_{m=1}^\infty \eta_m^{1-2/p} < \infty$.

Let

$$\begin{aligned} y_{it}^* &= y_{it} - \sigma_{0u\varepsilon}\sigma_{0\varepsilon}^{-2} x_{it}, \\ x_{it}^* &= \sigma_{0\varepsilon}^{-1} x_{it}, \end{aligned} \tag{11}$$

with $\widehat{e}_{it}^* = y_{it}^* - \widehat{\alpha}_i^* - \widehat{\beta}^* x_{it}^*$. The residuals from LSDV estimation of (2) are identical to those from LSDV of (11) in that

$$\begin{aligned} y_{it} - \widehat{\alpha}_i - \widehat{\beta} x_{it} &= y_{it}^* - \widehat{\alpha}_i^* - \widehat{\beta}^* x_{it}^* \\ &= y_{it} - \sigma_{0u\varepsilon}\sigma_{0\varepsilon}^{-2} x_{it} - \widehat{\alpha}_i^* - \widehat{\beta}^* \left(\sigma_{0\varepsilon}^{-1/2} x_{it}\right) \\ &= y_{it} - \widehat{\alpha}_i^* - \left(\widehat{\beta}^* \sigma_{0\varepsilon}^{-1} + \sigma_{0u\varepsilon}\sigma_{0\varepsilon}^{-2}\right) x_{it} \end{aligned} \tag{12}$$

with

$$\begin{aligned} \widehat{\alpha}_i &= \widehat{\alpha}_i^* \\ \widehat{\beta} &= \widehat{\beta}^* \sigma_{0\varepsilon}^{-1} + \left(\sigma_{0\varepsilon}^{-2}\right)\sigma_{0u\varepsilon}. \end{aligned}$$

This implies that

$$\begin{aligned} \widehat{\beta}^* &= \sigma_{0\varepsilon}\widehat{\beta} - \sigma_{0\varepsilon}^{-1}\sigma_{0u\varepsilon} \\ &= \sigma_{0\varepsilon}\widehat{\beta} - \sigma_{0\varepsilon}^{-1}\sigma_{0u\varepsilon}. \end{aligned}$$

Define

$$\sigma_{0v}^2 = \sigma_{0u}^2 - \sigma_{0u\varepsilon}^2 \sigma_{0\varepsilon}^{-2} \tag{13}$$

and

$$\sigma_v^2 = \sigma_u^2 - \sigma_{u\varepsilon}^2 \sigma_\varepsilon^{-2}.$$

Notice that

$$\begin{bmatrix} \frac{y_{it}^*}{\sigma_{0v}} \\ x_{it}^* \end{bmatrix} = L' \begin{bmatrix} y_{it} \\ x_{it} \end{bmatrix},$$

where

$$L^{'} = \begin{bmatrix} \frac{1}{\sigma_{0v}} & -\frac{1}{\sigma_{0v}}\sigma_{0u\varepsilon}\sigma_{0\varepsilon}^{-2} \\ 0 & \sigma_{0\varepsilon}^{-1} \end{bmatrix}.$$

It follows that (e.g., Hamilton, 1994)

$$\mathbf{W}_i^*(r) = L^{'}\Omega^{1/2}\mathbf{W}_i(r)$$

is a Wiener process with an identity covariance, $\mathbf{I}$, with

$$\mathbf{W}_i^*(r) = \begin{bmatrix} V_i^*(r) \\ W_i^*(r) \end{bmatrix}.$$

It follows that

$$\begin{aligned}
&\tfrac{1}{\sqrt{T}}\textstyle\sum_{t=1}^{[Tr]} L^{'} w_{it} \\
&= \tfrac{1}{\sqrt{T}}\textstyle\sum_{t=1}^{[Tr]} \begin{bmatrix} \frac{1}{\sigma_{0v}} & -\frac{1}{\sigma_{0v}}\sigma_{0u\varepsilon}\sigma_{0\varepsilon}^{-2} \\ 0 & \sigma_{0\varepsilon}^{-1} \end{bmatrix} \begin{bmatrix} u_{it} \\ \varepsilon_{it} \end{bmatrix} \\
&= \tfrac{1}{\sqrt{T}}\textstyle\sum_{t=1}^{[Tr]} \begin{bmatrix} \frac{u_{it}}{\sigma_{0v}} & -\frac{\varepsilon_{it}}{\sigma_{0v}}\sigma_{0u\varepsilon}\sigma_{0\varepsilon}^{-2} \\ \frac{\varepsilon_{it}}{\sigma_{0\varepsilon}} & \end{bmatrix} \\
&= \tfrac{1}{\sqrt{T}}\textstyle\sum_{t=1}^{[Tr]} \begin{bmatrix} \frac{u_{it}^*}{\sigma_{0v}} \\ \varepsilon_{it}^* \end{bmatrix} \\
&\Rightarrow \mathbf{W}_i^*(r) \text{ as } T \to \infty,
\end{aligned}$$

where

$$u_{it}^* = \quad u_{it} \quad -\varepsilon_{it}\sigma_{0u\varepsilon}\sigma_{0\varepsilon}^{-2}$$

and

$$\varepsilon_{it}^* = \frac{\varepsilon_{it}}{\sigma_{0\varepsilon}}.$$

Define

$$\sqrt{N}\frac{\widehat{\beta}^*}{\sigma_{0v}} = \frac{\sqrt{N}\xi_{3NT}}{\xi_{4NT}},$$

where $\zeta_{3iT} = \frac{1}{T^2}\sum_{t=1}^{T}\left(x_{it}^* - \overline{x}_i^*\right)\frac{y_{it}^*}{\sigma_{0v}}$, $\zeta_{4iT} = \frac{1}{T^2}\sum_{t=1}^{T}\left(x_{it}^* - \overline{x}_i^*\right)^2$, $\xi_{3NT} = \frac{1}{N}\sum_{i=1}^{N}\zeta_{3iT}$, and $\xi_{4NT} = \frac{1}{N}\sum_{i=1}^{N}\zeta_{4iT}$. Before we derive the asymptotics we summarize some limit theory for the LSDV estimator.

Lemma 2 *If Assumptions 1 and 3 hold, then*

(a) $\zeta_{3iT} \Rightarrow \int_0^1 W_i^*(r)V_i^*(r)dr - \int_0^1 W_i^*(r)dr \int_0^1 V_i^*(r)dr \equiv \zeta_{3i}$,

(b) $\zeta_{4iT} \Rightarrow \int_0^1 W_i^*(r)^2 dr - \left\{\int_0^1 W_i^*(r)dr\right\}^2 \equiv \zeta_{4i}$,

(c) $E\left[\zeta_{3iT}\right] \equiv \mu_{3T} = E\left[\zeta_{3i}\right] \equiv \mu_3 = 0$,

(d) $E\left[\zeta_{4i}\right] \equiv \mu_4 = \frac{1}{6}$,

(e) $E\left[\zeta_{4iT}\right] \equiv \mu_{4T} = \mu_4 + O(T^{-1})$,

(f) $Var\left[\zeta_{3iT}\right] \equiv \sigma_{3T}^2 = \frac{1}{90} + O(T^{-1})$,

(g) $Var\left[\zeta_{4iT}\right] \equiv \sigma_{4T}^2 = \frac{1}{45} + O(T^{-1})$,

(h) $Var\left[\zeta_{3i}\right] \equiv \sigma_3^2 = \frac{1}{90}$,

(i) $Var\left[\zeta_{4i}\right] \equiv \sigma_4^2 = \frac{1}{45}$,

(j) $\xi_{3NT} \xrightarrow{p} \mu_3 = 0$,

(k) $\xi_{4NT} \xrightarrow{p} \mu_4 = \frac{1}{6}$.

Using similar techniques in Theorem 1, we can show the results similar to the ones in Theorem 2.

Theorem 2 *Suppose the conditions of Lemma 2 are satisfied; then*

(a) $\widehat{\beta} \xrightarrow{p} \frac{\sigma_{0u\varepsilon}}{\sigma_{0\varepsilon}^2}$,

(b) $\sqrt{N}\left(\widehat{\beta} - \frac{\sigma_{0u\varepsilon}}{\sigma_{0\varepsilon}^2}\right) \Rightarrow N\left(0, \frac{2\sigma_{0v}^2}{5\sigma_{0\varepsilon}^2}\right)$,

(c) $T^{-1/2}t_{\beta} - \frac{T^{-1/2}\sigma_{0\varepsilon}^{-2}\sigma_{0u\varepsilon}\sqrt{\sum_{i=1}^{N}\sum_{t=1}^{T}(x_{it}-\overline{x}_i)^2}}{s} \Rightarrow N\left(0, \frac{2}{5}\right)$,

(d) $R^2 \xrightarrow{p} \frac{\sigma_{0u\varepsilon}^2}{\sigma_{0v}^2\sigma_{0\varepsilon}^2 + \sigma_{0u\varepsilon}^2}$,

(e) $DW = O_p(T^{-1})$,

(f) $TDW \xrightarrow{p} 6$.

The results of Theorem 1 and 2 can also be extended to the multiple regression, provided that $\{x_{it}\}$ are not cointegrated. The asymptotics of $\widehat{\beta}$ are different from those of a spurious regression in the pure time series, and this difference has an important consequence for residual-based cointegration tests using panel data, because the null distribution of residual-based cointegration tests depends on the asymptotics of $\widehat{\beta}$. This point is reported in the next section.

3 Residual-Based Tests for Cointegration

In this section we derive the null distributions of residual-based cointegration tests using Dickey-Fuller (DF) tests and augmented Dickey-Fuller (ADF) when applied to the model in Section 2.2. The model in Section 2.1 is a special case of the general class in Section 2.2.

3.1 Dickey-Fuller Test

The DF test can be applied to the residuals using

$$\widehat{e}_{it} = \rho \widehat{e}_{it-1} + v_{it}, \tag{14}$$

where $\widehat{e}_{it}$ is the estimate of e_{it} from (2). The OLS of ρ, $\widehat{\rho}$, is

$$\widehat{\rho} = \frac{\sum_{i=1}^{N} \sum_{t=2}^{T} \widehat{e}_{it} \widehat{e}_{it-1}}{\sum_{i=1}^{N} \sum_{t=2}^{T} \widehat{e}_{it-1}^{2}}. \tag{15}$$

The null hypothesis that $\rho = 1$ in regression (14) is tested by:

$$\begin{aligned} & \sqrt{N}T(\widehat{\rho} - 1) \\ &= \frac{\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \frac{1}{T} \sum_{t=2}^{T} \widehat{e}_{it-1} \triangle \widehat{e}_{it}}{\frac{1}{N} \sum_{i=1}^{N} \frac{1}{T^2} \sum_{t=2}^{T} \widehat{e}_{it-1}^{2}} \\ &= \frac{\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \frac{1}{T} \sum_{t=2}^{T} \widehat{e}_{it-1}^{*} \triangle \widehat{e}_{it}^{*}}{\frac{1}{N} \sum_{i=1}^{N} \frac{1}{T^2} \sum_{t=2}^{T} \left(\widehat{e}_{it-1}^{*}\right)^2} \\ &= \frac{\sqrt{N} \frac{1}{N} \sum_{i=1}^{N} \zeta_{5iT}}{\frac{1}{N} \sum_{i=1}^{N} \zeta_{6iT}} \\ &= \frac{\sqrt{N} \xi_{5NT}}{\xi_{6NT}}, \end{aligned}$$

where $\zeta_{5iT} = \frac{1}{T} \sum_{t=2}^{T} \widehat{e}_{it-1}^{*} \triangle \widehat{e}_{it}^{*}$, $\zeta_{6iT} = \frac{1}{T^2} \sum_{t=2}^{T} \widehat{e}_{it-1}^{*2}$, $\xi_{5NT} = \frac{1}{N} \sum_{i=1}^{N} \zeta_{5iT}$, and $\xi_{6NT} = \frac{1}{N} \sum_{i=1}^{N} \zeta_{6iT}$.

Theorem 3 *Suppose Assumptions 1 and 3 hold. Then,*

$$\sqrt{N}T(\widehat{\rho} - 1) - \frac{\sqrt{N}\mu_{5T}}{\mu_{6T}} \Rightarrow N\left(0, 3 + \frac{7.2\sigma_v^4}{\sigma_{0v}^4}\right)$$

and

$$t_\rho - \frac{\sqrt{N}\mu_{5T}}{s\sqrt{\mu_{6T}}} \Rightarrow N\left(0, \frac{\sigma_{0v}^2}{2\sigma_v^2} + \frac{3\sigma_v^2}{10\sigma_{0v}^2}\right),$$

where $\mu_{5T} = E[\zeta_{5iT}]$ *and* $\mu_{6T} = E[\zeta_{6iT}]$.

Corollary 3 *Assume Assumptions 1 and 3 are satisfied and* $\frac{\sqrt{N}}{T} \to 0$ *as* $N \to \infty$ *and* $T \to \infty$. *Then, we have*

$$\sqrt{N}T(\widehat{\rho} - 1) + \frac{3\sqrt{N}\sigma_v^2}{\sigma_{0v}^2} \Rightarrow N\left(0, 3 + \frac{7.2\sigma_v^4}{\sigma_{0v}^4}\right)$$

and

$$t_\rho + \frac{\sqrt{6N}\sigma_v}{2\sigma_{0v}} \Rightarrow N\left(0, \frac{\sigma_{0v}^2}{2\sigma_v^2} + \frac{3\sigma_v^2}{10\sigma_{0v}^2}\right).$$

Note that with the strong exogeneity and in the absence of serial correlation we have $\sigma_u^2 = \sigma_{0u}^2 = \sigma_v^2 = \sigma_{0v}^2$. This case was previously studied by Kao and Chen (1995b). We then have the following corollary:

Corollary 4 *Suppose Assumptions 1 and 3 hold. Then,*

$$\sqrt{N}T\left(\widehat{\rho}-1\right)-\frac{\sqrt{N}\mu_{5T}}{\mu_{6T}}\Rightarrow N\left(0,10.2\right)$$

and

$$\sqrt{1.25}\left\{t_{\rho}-\frac{\sqrt{N}\mu_{5T}}{\sqrt{\mu_{6T}}}\right\}\Rightarrow N\left(0,1\right).$$

3.2 Augmented Dickey-Fuller (ADF) Test

The Dickey-Fuller test in Section 3.1 was based on a simple OLS regression of $\widehat{e}_{it}$ on its own lagged value. Correction for serial correlation was made to the OLS and t-statistic. Alternatively, the lagged changes in the residuals can be added to the regression of (14):

$$\widehat{e}_{it} \quad = \quad \rho\widehat{e}_{it-1}+\textstyle\sum_{j=1}^{p}\varphi_j \bigtriangleup \widehat{e}_{it-j}+v_{itp}, \tag{16}$$

where p is chosen so that the residuals v_{itp} are serially uncorrelated. The ADF test statistic discussed in this section is the usual t-statistic of $\rho=1$ the regression (16). Let X_{ip} and X_{ip}^{*} be the matrix of observation on the p regressors $(\triangle\widehat{e}_{it-1},\triangle\widehat{e}_{it-2},\cdots,\triangle\widehat{e}_{it-p})$ and $\left(\triangle\widehat{e}_{it-1}^{*},\triangle\widehat{e}_{it-2}^{*},\cdots,\triangle\widehat{e}_{it-p}^{*}\right)$ respectively, e_i and e_i^{*} equal the vector of observations of $\widehat{e}_{it-1}$ and $\widehat{e}_{it-1}^{*}$ respectively, v_i equal the vector of observations of $v_{itp}, Q_i = I-X_{ip}(X_{ip}'X_{ip})^{-1}X_{ip}^{'}$, $Q_i^{*}=I-X_{ip}^{*}(X_{ip}^{*\prime}X_{ip}^{*})^{-1}X_{ip}^{*'}$, and $s_v^2=\frac{1}{NT}\sum\limits_{i=1}^{N}\sum\limits_{t=1}^{T}\widehat{v}_{itp}^{2}$, where $\widehat{v}_{itp}$ is an estimate of v_{itp}. Under the null hypothesis of no cointegration, the ADF test takes the form

$$t_{ADF}=\frac{\left(\widehat{\rho}-1\right)\left[\sum\limits_{i=1}^{N}\left(e_i^{'}Q_ie_i\right)\right]^{\frac{1}{2}}}{s_v},$$

where $\widehat{\rho}$ is the OLS estimate of ρ. Note that

$$\left(\widehat{\rho}-1\right)=\left[\sum_{i=1}^{N}\left(e_i'Q_ie_i\right)\right]^{-1}\left[\sum_{i=1}^{N}\left(e_i'Q_iv_i\right)\right].$$

Now,

$$\begin{aligned} & t_{ADF} \\ &= \frac{\sum_{i=1}^{N}\left(e_i' Q_i v_i\right)}{s_v\sqrt{\sum_{i=1}^{N}\left(e_i' Q_i e_i\right)}} \\ &= \frac{\sqrt{N}\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T}\left(e_i' Q_i v_i\right)}{s_v\sqrt{\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T^2}\left(e_i' Q_i e_i\right)}} \\ &= \frac{\sqrt{N}\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T}\left(e_i^{*\prime} Q_i^* v_i\right)}{s_v\sqrt{\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T^2}\left(e_i^{*\prime} Q_i^* e_i^*\right)}} \\ &= \frac{\sqrt{N}\xi_{7NT}}{s_v\sqrt{\xi_{8NT}}}, \end{aligned}$$

where $\xi_{7NT} = \frac{1}{N}\sum_{i=1}^{N}\zeta_{7iT}$, $\zeta_{7iT} = \frac{1}{T}\left(e_i^{*\prime} Q_i^* v_i\right)$, $\xi_{8NT} = \frac{1}{N}\sum_{i=1}^{N}\zeta_{8iT}$, and $\zeta_{8iT} = \frac{1}{T^2}\left(e_i^{*\prime} Q_i^* e_i^*\right)$.

Theorem 4 *Suppose that the assumptions of Theorem 3 are satisfied and* $p \to \infty$. *Then, we have*

$$t_{ADF} - \frac{\sqrt{N}\mu_{7T}}{s_v\sqrt{\mu_{8T}}} \Rightarrow N\left(0, \frac{\sigma_{0v}^2}{2\sigma_v^2} + \frac{3\sigma_v^2}{10\sigma_{0v}^2}\right)$$

as $N \to \infty$ *and* $T \to \infty$, *where* $\mu_{7T} = E\left[\zeta_{7iT}\right]$ *and* $\mu_{8T} = E\left[\zeta_{8iT}\right]$.

4 Estimation of Long-Run Parameters

The asymptotic distributions given in Theorem 2 depend on unknown parameters (σ_u^2, σ_ε^2, $\sigma_{u\varepsilon}$, $\sigma_{0\varepsilon}^2$, σ_{0u}^2 and $\sigma_{0u\varepsilon}$). We now define

$$w_{it} = \left(u_{it}, \varepsilon_{it}\right)'.$$

The long-run covariance matrix of w_{it} is given by

$$\Omega = \lim_{T\to\infty}\frac{1}{T}E\left(\sum_{t=1}^{T} w_{it}\right)\left(\sum_{t=1}^{T} w_{it}\right)' = \Sigma + \Gamma + \Gamma' \equiv \begin{bmatrix} \sigma_{0u}^2 & \sigma_{0u\varepsilon} \\ \sigma_{0u\varepsilon} & \sigma_{0\varepsilon}^2 \end{bmatrix},$$

where $\Gamma = \lim_{T\to\infty}\frac{1}{T}\sum_{k=1}^{T-1}\sum_{t=k+1}^{T} E\left(w_{it}w_{it-k}'\right) \equiv \begin{bmatrix} \Gamma_u & \Gamma_{u\varepsilon} \\ \Gamma_{\varepsilon u} & \Gamma_\varepsilon \end{bmatrix}$ and $\Sigma = \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T} E\left(w_{it}w_{it}'\right) \equiv$

$\begin{bmatrix} \sigma_u^2 & \sigma_{u\varepsilon} \\ \sigma_{u\varepsilon} & \sigma_\varepsilon^2 \end{bmatrix}$, the usual covariance matrix. Once the estimates of w_{it}, $\widehat{w}_{it}$, are obtained, we can use

$$\widehat{\Sigma} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\widehat{w}_{it}\widehat{w}_{it}^{'} \tag{17}$$

to estimate Σ. Ω can be estimated by

$$\widehat{\Omega} = \frac{1}{N}\sum_{i=1}^{N}\left\{\frac{1}{T}\sum_{t=1}^{T}\widehat{w}_{it}\widehat{w}_{it}^{'} + \frac{1}{T}\sum_{\tau=1}^{l}\varpi_{\tau l}\sum_{t=\tau+1}^{T}\left(\widehat{w}_{it}\widehat{w}_{it-\tau}^{'} + \widehat{w}_{it-\tau}\widehat{w}_{it}^{'}\right)\right\}, \tag{18}$$

where $\varpi_{\tau l}$ is a weight function or a kernel. Usually kernels are truncated by the bandwidth parameter l so that $\varpi_{\tau l} = 0$ for $\tau > l$. Using Phillips and Durlauf (1986) and the law of large numbers for triangular arrays, $\widehat{\Sigma}$ and $\widehat{\Omega}$ can be shown to be consistent for Σ and Ω.

5 Some Monte Carlo Results

5.1 Spurious Regression

In this section, we report some results from Monte Carlo experiments that examined the finite-sample properties of the spurious LSDV estimator, $\widehat{\beta}$, t_β, R^2, and DW. The simulations were performed by a Sun SparcServer 1000 using the software GAUSS 3.27. The results we report are based on $10,000$ replications. The data were generated by creating $T + 1,000$ observations and discarding the first $1,000$ observations to remove the effect of the initial conditions.

The data generating process (DGP) was

$$y_{it} = y_{it-1} + u_{it}$$

and

$$x_{it} = x_{it-1} + \varepsilon_{it}$$

for $i = 1, ..., N$ and $t = 1, ..., T$, where the innovation sequences $(u_{it}, \varepsilon_{it})$ were generated from a bivariate normal with independence across both individuals and time periods, i.e.,

$$\begin{bmatrix} u_{it} \\ \varepsilon_{it} \end{bmatrix} \overset{iid}{\sim} N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}\right).$$

Random numbers for $(u_{it}, \varepsilon_{it})$ were generated by the GAUSS procedure RNDNS. For each experiment, we calculated the mean values of $\widehat{\beta}$, t_β, R^2, DW, the mean estimated standard error (ESE), $P(|t| > 1.96)$ (the rejection frequency of the t-test of $H_0 : \beta = 0$), and the sample standard deviation (SSD) of $\widehat{\beta}$ where

$$SSD(\widehat{\beta}) = \left\{ \frac{1}{M-1} \sum_{i=1}^{M} [\widehat{\beta}_i - \overline{\beta}]^2 \right\}^{1/2} ,$$

$\overline{\beta} = \frac{1}{M} \sum_{i=1}^{M} \widehat{\beta}_i$, and M was the number of replications. For sample size, we considered different settings for T and N.

The Monte Carlo results are reported in Tables 1, 2, and 3. In the case where $N = T = 30$, the probability of rejecting H_0, $P(|t_\beta| > 1.96)$, is approximately .5723, which is far from the "conventional" value of .05. Similar discrepancies exist for all our other values of N and T. This means that, even when the null hypothesis is true, we will wrongly reject it most of the time.

We know from Table 1 that the mean values of $\widehat{\beta}$ almost always stay constant around zero, but $\overline{\beta} \pm 2 \times SSD$ converges to the mean value as panel size increases. Unlike in the pure time-series case, $\widehat{\beta}$ converges to zero in probability instead of being a random variable. Table 1 also reveals some discrepancy between the SSD and the ESE. However, the discrepancy between these two measures is not as severe as the discrepancy in the pure time series because both the SSD and the ESE diminish as T increases in panel data.

The mean value of the t-statistic changes little as $T(N)$ grows from 10 to 70, but the SSD increases rapidly. The likelihood of $|t_\beta| > 1.96$ becomes higher and higher so that the critical value at the 5% significance level increases consequently as T increases. Therefore, as in a pure time series, the spurious regression problem becomes worse as the panel size grows. This is confirmed by Table 1, which records the rejection frequency of the t-test for the sample sizes from 10 to 80. The rejection rate of the null hypothesis of no relationship between x_{it} and y_{it} increases steadily with T. At $N = T = 50$; the rejection rate is as high as 66%.

We also examined how the finite sample properties of the spurious LSDV regression change if N and T are different. We first fixed N at 30 and changed T. Then, we let T be fixed at 30 and changed N. Table 2 summarizes the results for different time-series dimensions with $N = 30$. Table 3 summarizes the results for different cross-section dimensions with $T = 30$. We found that an increase in cross-section decreases the SSD of $\widehat{\beta}$; varying cross-section size has little effect on the t-statistic; and the t-statistic is closely related to the time series dimension. The reason why the t statistic does not converge to the standard normal distribution is that the t statistic does not have a zero mean even when T is fixed (at 30). Hence, increasing N will not help the t statistic converge to a meaningful distribution (here a standard normal distribution) without some sort of normalization. When T grows, the SSD of the t-statistic and the rejection rate of the null hypothesis based on the conventional critical value increase significantly.

The results obtained from the pure time series tell us that R^2 may be quite high in the spurious regression. However, this is not the case using panel data. Tables 1 and 2 reveal that R^2 remains very low and decreasing, which is consistent with the result in Theorem 1. The results on the Durbin-Watson statistic are also what we expected. Clearly, as shown in Table 1, DW converges to zero in probability. Furthermore, TDW stays constant at about 6 on average.

5.2 Residual-Based Tests

To examine the finite sample properties of the proposed tests, we conducted Monte Carlo experiments which are similar to the design of Engel and Granger (1987) and are used in Gonzalo (1994) and Haug (1996). The data generating process (DGP) was:

$$y_{it} - \alpha_i - \beta x_{it} = z_{it}$$

$$a_1 y_{it} - a_2 x_{it} = w_{it}$$

$$z_{it} = \rho z_{it-1} + e_{zit}$$

$$w_{it} = w_{it-1} + e_{wit}$$

$$e_{wit} = \phi_{it} + \pi \phi_{it-1}$$

and

$$\begin{bmatrix} e_{zit} \\ \phi_{it} \end{bmatrix} \overset{iid}{\sim} N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \theta\sigma \\ \theta\sigma & \sigma^2 \end{bmatrix} \right).$$

The RNDN procedure in GAUSS was used to generate the random numbers. The data were generated by creating $T + 1,000$ observations and discarding the first $1,000$ observations to remove the effect of the initial conditions. The results we report are based on $10,000$ replications. We generate α_i from a uniform distribution, $U[0, 10]$, and set $\beta = 2$.

For the purposes of the this paper, we consider the following values for the parameters in the DGP: $a_1 = 1$, $a_2 = -1$, $\rho = (.95, 1)$, $\sigma = (.25, 1, 4)$, $\pi = (-0.8, 0, 0.8)$, and $\theta = (-0.5, 0, 0.5)$. When $\rho = 1$, y_{it} and x_{it} are not cointegrated, and when $|\rho| < 1$, they are cointegrated.

The estimate of the long-run covariance matrix

$$\widehat{\Omega} = \begin{pmatrix} \widehat{\sigma}^2_{0u} & \widehat{\sigma}_{0u\varepsilon} \\ \widehat{\sigma}_{0u\varepsilon} & \widehat{\sigma}^2_{0\varepsilon} \end{pmatrix}$$

and the long run conditional variance

$$\widehat{\sigma}^2_{0v} = \widehat{\sigma}^2_{0u} - \widehat{\sigma}^2_{0u\varepsilon} \widehat{\sigma}^{-2}_{0\varepsilon}.$$

were obtained by using the procedure KERNEL in COINT 2.0 with a Bartlett window of lag length 5.

Define

$$DF_{\rho} = \frac{\sqrt{N}T\left(\widehat{\rho}-1\right)+3\sqrt{N}}{\sqrt{10.2}},$$

$$DF_{t} = \sqrt{1.25}t_{\rho}+\sqrt{1.875N},$$

$$DF_{\rho}^{*} = \frac{\sqrt{N}T\left(\widehat{\rho}-1\right)+\frac{3\sqrt{N}\widehat{\sigma}_{v}^{2}}{\widehat{\sigma}_{0v}^{2}}}{\sqrt{3+\frac{7.2\widehat{\sigma}_{v}^{4}}{\widehat{\sigma}_{0v}^{4}}}},$$

$$DF_{t}^{*} = \frac{t_{\rho}+\frac{\sqrt{6N}\widehat{\sigma}_{v}}{2\widehat{\sigma}_{0v}}}{\sqrt{\frac{\widehat{\sigma}_{0v}^{2}}{2\widehat{\sigma}_{v}^{2}}+\frac{3\widehat{\sigma}_{v}^{2}}{10\widehat{\sigma}_{0v}^{2}}}},$$

and

$$ADF = \frac{t_{ADF}+\frac{\sqrt{6N}\widehat{\sigma}_{v}}{2\widehat{\sigma}_{0u}}}{\sqrt{\frac{\widehat{\sigma}_{0v}^{2}}{2\widehat{\sigma}_{v}^{2}}+\frac{3\widehat{\sigma}_{v}^{2}}{10\widehat{\sigma}_{0v}^{2}}}},$$

where

$$\widehat{\sigma}_{v}^{2} = \widehat{\sigma}_{u}^{2}-\widehat{\sigma}_{u\varepsilon}^{2}\widehat{\sigma}_{\varepsilon}^{-2}.$$

We expect that $DF_{\rho}^{*}, DF_{t}^{*}$, and ADF converge to $N(0,1)$ asymptotically. We also put DF_{ρ} and DF_{t} here for comparison and the two lags are selected for ADF.

Tables 4 contains estimates of the size of tests at the a 5% level when the asymptotic critical value, 1.645, from the $N(0,1)$ distribution was used. Five different tests are considered - $DF_{\rho}, DF_{t}, DF_{\rho}^{*}, DF_{t}^{*}$, and ADF - for different sample sizes with $\theta = 0$, no moving average component ($\pi = 0.0$), $\sigma = 1$, and endogenous x_{it} ($a_1 = 1.0$). We note that all tests have a large size distortion when T is small (e.g., $T = 10$) even at $N = 300$, but the size distortion begins to disappear quickly when the T is increased to 25 for all N. The empirical sizes of DF_{ρ}^{*} and DF_{t}^{*} are close to 0.05 when T and N are both large. But the empirical size of the ADF test is greater than 0.09 at each sample sizes for T and N. Overall, it is found that DF_{ρ}^{*} and DF_{t}^{*} outperform the rest in terms of the size distortion.

Tables 5 reports the unadjusted powers of the five tests when the asymptotic critical value is used and $\rho = .90$ for different sample sizes with 0 contemporaneous correlation ($\theta = 0$), no moving average component ($\pi = 0.0$), $\sigma = 1$ and endogenous x_{it} ($a_1 = 1.0$). All tests have little power when $T = 10$ and when N is small, as one would expect. However, DF_{ρ}^{*} displayed little power even when N is large. When T is increased to 25 and above, DF_{ρ}^{*} dominates DF_{t}^{*}, and the ADF tests for all N. Interestingly, DF_{ρ} and DF_{t}, though they are misspecified, perform pretty well in terms of size distortion and power.

To provide further insight into other specifications, we consider the size and power of these five tests for three different values of θ, π, and σ in Tables 6 - 8. With negative π, all five tests are severely distorted, as one would expect. In fact, with $\pi = -0.8$, all tests except ADF reject no cointegration 100% of the time for a nominal 5% level for sample sizes $T = 50$ and $N = 50$. The empirical sizes of the DF_ρ and DF_t tests are smaller than the nominal sizes of the tests (at 5%), when $\pi = 0.8$ at $N = T = 50$. For $\pi = 0.8$, the size distortion of all five tests begins to improve as σ is increased from 0.25 to 1 and 4. For $\pi = 0.0$, all five tests show little variation in size across the various values of θ and σ.

For $\sigma \geq 1$, we note that the unadjusted power of tests is extremely high for all values of θ, and π. For $\sigma = 0.25$ and $\pi = 0$, the power of all tests is very low. When σ is increased to 1 and 4, the power increases dramatically.

Table 8 reports the size-adjusted powers with $\rho = 0.85$ such that each test has the same rejection frequency of 5% when the null hypothesis is true. For $\sigma = 0.25$ and $\pi = -0.8$, the power of all tests is low and the null hypothesis of no cointegration is often not rejected. With $\pi = 0$, ADF has the lowest power for all θ and, on the other hand, DF_ρ^* has the highest power. With $\pi = 0$, the power of all tests except ADF is high. When σ is increases to 1, the power of all tests increases to viturally 1 except when $\theta = -0.8$. In fact, the power of all tests remains very low, zeros for most cases when $\theta = -0.8$. Interestingly and to my surprise, ADF became very powerful when $\sigma = 4$ and $\theta \geq 0$.

The results in Tables 4-8 suggest that the DF_ρ^* and DF_t^* tests have better size and power properties than the DF_ρ, DF_t, and ADF tests when σ is small. However, when σ is large, ADF clearly dominates the rest.

6 Conclusion

The first half of this paper develops a framework for understanding the behavior of spurious panel regression. I have provided an asymptotic theory for the behavior of the LSDV estimator in a model which attempts to estimate the panel regression when the dependent variable and independent variable are actually independent $I(1)$ processes. In particular, the t-statistic diverges in spite of the fact that the LSDV estimator converges to zero in probability. Moreover, DW converges to zero in probability. The asymptotic results and simulation results in this paper are useful beyond explaining the spurious effects on panel regression. Asymptotic distributions of residual-based tests depend on the LSDV estimator from the spurious regression because residual-based tests for cointegration in panel data take the spurious regression as the null hypothesis.

In the second half of this paper, we proposed tests for the null hypothesis of no cointegration in panel data and derived asymptotic distribution for each test. The simulations show that the distributions of DF_ρ^*, DF_t^*, and ADF can be far different from $N(0,1)$ suggested by the theory when the underlying process contains an MA component. From the Monte Carlo results, it is apparent that the DF_ρ and DF_t tests are substantially robust despite the model misspecification. The results in Tables 4-8 suggest that the DF_ρ^* and DF_t^* tests have better size and power properties than the DF_ρ, DF_t, and ADF tests.

Appendix

A Proof of Lemma 1

Proof. (a) and (b) are immediate from Phillips (1986) and Banerjee et al. (1993). For example, using Lemma 1 in Phillips (1986) and (3.31) in Banerjee et al. (1993), we have for a fixed i

$$\begin{aligned} \zeta_{2iT} &= \tfrac{1}{T^2} \textstyle\sum_{t=1}^{T} (x_{it} - \overline{x}_i)^2 \\ &= \tfrac{1}{T^2} \textstyle\sum_{t=1}^{T} x_{it}^2 - \tfrac{1}{T}\overline{x}_i^2 \\ &\Rightarrow \sigma_\varepsilon^2 \textstyle\int_0^1 W_i^2(r)dr - \sigma_\varepsilon^2 \left\{ \int_0^1 W_i(r)dr \right\}^2 \\ &= \sigma_\varepsilon^2 \left[\textstyle\int_0^1 W_i^2(r)dr - \left\{ \int_0^1 W_i(r)dr \right\}^2 \right] \\ &\equiv \zeta_{2i}. \end{aligned} \tag{19}$$

Now

$$\zeta_{1iT} = \frac{1}{T^2} \sum_{t=1}^{T} (x_{it} - \overline{x}_i) y_{it} = \frac{1}{T^2} \sum_{t=1}^{T} x_{it} y_{it} - \left(T^{-1/2} \overline{x}_i \right) \left(T^{-1/2} \overline{y}_i \right). \tag{20}$$

Then, we have

$$T^{-1/2} \overline{x}_i \Rightarrow \sigma_\varepsilon \int_0^1 W_i(r)dr, \tag{21}$$

$$T^{-1/2} \overline{y}_i \Rightarrow \sigma_u \int_0^1 V_i(r)dr, \tag{22}$$

and

$$\frac{1}{T^2} \sum_{t=1}^{T} x_{it} y_{it} \Rightarrow \sigma_\varepsilon \sigma_u \int_0^1 W_i(r) V_i(r) dr. \tag{23}$$

The limiting distribution of ζ_{1iT} can be deduced by using (21) - (23)

$$\zeta_{1iT} \Rightarrow \zeta_{1i},$$

where

$$\zeta_{1i} \equiv \sigma_\varepsilon \sigma_u \int_0^1 W_i(r) V_i(r) dr - \left[\sigma_\varepsilon \int_0^1 W_i(r) dr \right] \left[\sigma_u \int_0^1 V_i(r) dr \right].$$

To prove (c), we use the fact that

$$E\left[\int_0^1 W_i(r)V_i(r)dr\right] = 0$$

and

$$E\left[\int_0^1 W_i(r)dr\right] = E\left[\int_0^1 V_i(r)dr\right] = 0,$$

and, therefore,

$$E\left[\zeta_{1i}\right] \equiv \mu_1 = 0. \tag{24}$$

Also,

$$E\left[\zeta_{1iT}\right] = \mu_{1T} = 0 \tag{25}$$

because of the independence of x_{it} and y_{it}.

Parts (d) and (g) can be shown easily as follows:

$$E\left[\zeta_{2i}\right] = \frac{\sigma_\varepsilon^2}{6} \equiv \mu_2 \tag{26}$$

and

$$Var\left[\zeta_{2i}\right] = \frac{\sigma_\varepsilon^4}{45} \equiv \sigma_2^2. \tag{27}$$

To prove (e), we note that

$$\begin{aligned} \mu_{2T} &= E\left[\zeta_{2iT}\right] \\ &= \tfrac{1}{T^2}\left\{\sum_{t=1}^T E\left[x_{it}\right]^2 - \tfrac{1}{T}E\left(\sum_{t=1}^T x_{it}\right)^2\right\} \\ &= \tfrac{\sigma_\varepsilon^2}{T^2}\left\{\tfrac{T(T+1)}{2} - \tfrac{T(T+1)(2T+1)}{6T}\right\} \\ &= \tfrac{\sigma_\varepsilon^2}{6} + O(T^{-1}) = \mu_2 + O(T^{-1}). \end{aligned}$$

To prove (f), we first rewrite

$$\begin{aligned} Var\left(\zeta_{1iT}\right) &= Var\left(\tfrac{1}{T^2}\sum_{t=1}^T \left(x_{it} - \overline{x}_i\right) y_{it}\right) \\ &= \tfrac{1}{T^4}E\left[\left(\sum_{t=1}^T x_{it}y_{it}\right)^2\right] - \tfrac{2}{T^5}E\left[\left(\sum_{t=1}^T x_{it}y_{it}\right)\left(\sum_{t=1}^T x_{it}\right)\left(\sum_{t=1}^T y_{it}\right)\right] \\ &\quad + \tfrac{1}{T^6}E\left[\left(\sum_{t=1}^T x_{it}\right)^2\left(\sum_{t=1}^T y_{it}\right)^2\right]. \end{aligned}$$

After some tedious algebra, we can show that

$$\frac{1}{T^4}E\left[\left(\sum_{t=1}^T x_{it}y_{it}\right)^2\right] = \frac{\sigma_\varepsilon^2\sigma_u^2}{6} + O\left(T^{-1}\right),$$

$$\frac{2}{T^5}E\left[\left(\sum_{t=1}^{T}x_{it}y_{it}\right)\left(\sum_{t=1}^{T}x_{it}\right)\left(\sum_{t=1}^{T}y_{it}\right)\right]=\frac{4\sigma_\varepsilon^2\sigma_u^2}{15}+O\left(T^{-1}\right),$$

and

$$\frac{1}{T^6}E\left[\left(\sum_{t=1}^{T}x_{it}\right)^2\left(\sum_{t=1}^{T}y_{it}\right)^2\right]=\frac{\sigma_\varepsilon^2\sigma_u^2}{9}+O\left(T^{-1}\right).$$

It follows that

$$Var\left(\zeta_{1iT}\right)=\left(\frac{1}{6}-\frac{4}{15}+\frac{1}{9}\right)\sigma_\varepsilon^2\sigma_u^2+O\left(T^{-1}\right)=\frac{\sigma_\varepsilon^2\sigma_u^2}{90}+O\left(T^{-1}\right),$$

proving (f). Next, we prove (h). First,

$$\begin{aligned}Var\left(\zeta_{1i}\right) &= \sigma_\varepsilon^2\sigma_u^2E\left[\left(\int_0^1 W_i(r)V_i(r)dr-\int_0^1 W_i(r)dr\int_0^1 V_i(r)dr\right)^2\right]\\ &= \sigma_\varepsilon^2\sigma_u^2\{E\left[\left(\int_0^1 W_i(r)V_i(r)dr\right)^2\right]+E\left[\left(\int_0^1 W_i(r)dr\int_0^1 V_i(r)dr\right)^2\right]\\ &\quad -2E\left[\left(\int_0^1 W_i(r)V_i(r)dr\right)\left(\int_0^1 W_i(r)dr\int_0^1 V_i(r)dr\right)\right]\}.\end{aligned}$$

Again, we can show after tedious algebra,

$$E\left[\left(\int_0^1 W_i(r)V_i(r)dr\right)^2\right]=\frac{1}{6},$$

$$E\left[\left(\int_0^1 W_i(r)dr\int_0^1 V_i(r)dr\right)^2\right]=\frac{1}{9},$$

and

$$E\left[\left(\int_0^1 W_i(r)V_i(r)dr\right)\left(\int_0^1 W_i(r)dr\int_0^1 V_i(r)dr\right)\right]=\frac{2}{15}.$$

Hence,

$$Var\left(\zeta_{1i}\right)=\frac{\sigma_\varepsilon^2\sigma_u^2}{90},$$

establishing (h).

Part (i) is easy. Finally,

$$\xi_{1NT}\equiv\frac{1}{N}\sum_{i=1}^{N}\zeta_{1iT}\xrightarrow{p}\mu_1=0$$

and

$$\xi_{2NT}\equiv\frac{1}{N}\sum_{i=1}^{N}\zeta_{2iT}\xrightarrow{p}\mu_2=\frac{\sigma_\varepsilon^2}{6}$$

as $N\to\infty$ and $T\to\infty$, as required for (j) and (k) by the law of large numbers for triangular arrays. ■

B Proof of Theorem 1

Proof. By Lemma 3 we know that

$$\widehat{\beta} = \frac{\xi_{1NT}}{\xi_{2NT}} \xrightarrow{p} \frac{\mu_1}{\mu_2} = 0 \tag{28}$$

proving (a). We note that, unlike the pure time-series case, $\widehat{\beta}$ does converge in probability to zero here.

Note that the sequence

$$\sqrt{N}\xi_{1NT} \quad = \quad \tfrac{1}{\sqrt{N}} \textstyle\sum_{i=1}^{N} \zeta_{1iT}$$

is a triangular array sequence, and therefore a central limit theorem of a triangular array is needed. The central limit theorem of a triangular array (e.g., see Theorem 27.2, Billingsley, 1986, p. 369) can be applied if the Lindeberg Condition is satisfied; that is, for any $\epsilon > 0$,

$$\lim_{T\to\infty} \frac{1}{s_T^2} \sum_{i=i}^{N(T)} E\left[I\left(|\zeta_{1iT}| > \epsilon s_T\right)\left(\zeta_{1iT}\right)^2\right] = 0 \tag{29}$$

where $s_T^2 = \sum_{i=1}^{N(T)} var(\zeta_{1iT})$, and $I(.)$ represents an indicator function taking the value one if the condition in the brackets is satisfied and zero otherwise. It is easy to see that (29) is satisfied since $s_T^2 = \sum_{i=1}^{N(T)} var(\zeta_{1iT}) = N(T)\sigma_{1T}^2 \to \infty$ and $\sigma_{1T}^2 = \sigma_1^2 + O(T^{-1}) < \infty$. Since the Lindeberg condition holds, we can readily see that the $\sqrt{N}\xi_{1NT}$ will converge to a normal variable with an appropriate normalization and that ξ_{2NT} will converge to μ_2 in probability by the law of large numbers for triangular arrays. Using the Slutsky theorem and Lemma 3, we obtain

$$\sqrt{N}\widehat{\beta} = \frac{\sqrt{N}\xi_{1NT}}{\xi_{2NT}} \Rightarrow N\left(0, \frac{\sigma_1^2}{\mu_2^2}\right) = N\left(0, \frac{2}{5}\frac{\sigma_u^2}{\sigma_\varepsilon^2}\right),$$

as required for (b). Thus, after rescaling by $\sqrt{N}$, $\widehat{\beta}$ converges in distribution to a normal random variable.

We observe that

$$T^{-1/2}\widehat{\alpha}_i = T^{-1/2}\overline{y}_i - \widehat{\beta}T^{-1/2}\overline{x}_i.$$

From (28),

$$T^{-1/2}\widehat{\alpha}_i \Rightarrow \sigma_u \int_0^1 V_i(r)dr - 0 \cdot \sigma_\varepsilon \int_0^1 W_i(r)dr = \sigma_u \int_0^1 V_i(r)dr \sim N\left(0, \frac{\sigma_u^2}{3}\right) \tag{30}$$

for all i as $T \to \infty$. We define

$$s^2 = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left\{\left(y_{it} - \overline{y}_i\right) - \widehat{\beta}\left(x_{it} - \overline{x}_i\right)\right\}^2 \tag{31}$$

and then (e.g., Phillips, 1986, p. 331)

$$\begin{aligned} T^{-1}s^2 &= \tfrac{1}{N}\textstyle\sum_{i=1}^{N}\left[\tfrac{1}{T^2}\sum_{t=1}^{T}\left\{(y_{it}-\overline{y}_i)-\widehat{\beta}\,(x_{it}-\overline{x}_i)\right\}^2\right] \\ &= \tfrac{1}{N}\textstyle\sum_{i=1}^{N}\left[\tfrac{1}{T^2}\sum_{t=1}^{T}(y_{it}-\overline{y}_i)^2-\widehat{\beta}^2\tfrac{1}{T^2}\sum_{t=1}^{T}(x_{it}-\overline{x}_i)^2\right] \\ &= \tfrac{1}{N}\textstyle\sum_{i=1}^{N}\left\{\tfrac{1}{T^2}\sum_{t=1}^{T}(y_{it}-\overline{y}_i)^2\right\}-\widehat{\beta}^2\tfrac{1}{N}\sum_{i=1}^{N}\left\{\tfrac{1}{T^2}\sum_{t=1}^{T}(x_{it}-\overline{x}_i)^2\right\}, \end{aligned} \tag{32}$$

where

$$\tfrac{1}{T^2}\textstyle\sum_{t=1}^{T}(y_{it}-\overline{y}_i)^2 \quad\Rightarrow\quad \sigma_u^2\left[\int_0^1 V_i(r)^2dr-\left(\int_0^1 V_i(r)dr\right)^2\right]\equiv\eta_i$$

as $T\to\infty$. It follows that (e.g., Levin and Lin, 1992, Appendix 1)

$$E[\eta_i]=\frac{\sigma_u^2}{6} \tag{33}$$

and

$$var[\eta_i]=\frac{\sigma_u^4}{45}. \tag{34}$$

Therefore,

$$\frac{1}{N}\sum_{i=1}^{N}\left\{\frac{1}{T^2}\sum_{t=1}^{T}(y_{it}-\overline{y}_i)^2\right\}\xrightarrow{p}\frac{\sigma_u^2}{6}.$$

Similarly,

$$\frac{1}{N}\sum_{i=1}^{N}\left\{\frac{1}{T^2}(x_{it}-\overline{x}_i)^2\right\}\xrightarrow{p}\frac{\sigma_\varepsilon^2}{6}.$$

Obviously,

$$\widehat{\beta}^2\frac{1}{N}\sum_{i=1}^{N}\left\{\frac{1}{T^2}(x_{it}-\overline{x}_i)^2\right\}\xrightarrow{p}0.$$

Hence,

$$T^{-1}s^2\xrightarrow{p}\frac{\sigma_u^2}{6}. \tag{35}$$

Thus, s^2 diverges instead of converging. Now,

$$
\begin{aligned}
T^{-1/2}t_\beta &= \frac{T^{-1/2}\widehat{\beta}\sqrt{\sum_{i=1}^{N}\sum_{t=1}^{T}(x_{it}-\overline{x}_i)^2}}{s} \\
&= \frac{\sqrt{N}\widehat{\beta}\sqrt{\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T^2}\sum_{t=1}^{T}(x_{it}-\overline{x}_i)^2}}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left[\frac{1}{T^2}\sum_{t=1}^{T}(y_{it}-\overline{y}_i)^2-\widehat{\beta}^2\frac{1}{T^2}(x_{it}-\overline{x}_i)^2\right]}} \\
&= \frac{\sqrt{N}\widehat{\beta}\sqrt{\frac{1}{N}\sum_{i=1}^{N}\zeta_{2iT}}}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}\eta_{iT}-\widehat{\beta}^2\frac{1}{N}\sum_{i=1}^{N}\zeta_{2iT}}} \\
&\Rightarrow \frac{N\left(0,\frac{\sigma_1^2}{\mu_2^2}\right)\sqrt{\frac{\sigma_\varepsilon^2}{6}}}{\sqrt{\frac{\sigma_u^2}{6}}} = N\left(0,\frac{\sigma_1^2}{\mu_2^2}\right)\frac{\sigma_\varepsilon}{\sigma_u} \\
&= N\left(0,\frac{\sigma_1^2}{\mu_2^2}\frac{\sigma_\varepsilon^2}{\sigma_u^2}\right) \\
&= N\left(0,\tfrac{2}{5}\right)
\end{aligned}
$$

as required for part (c). That is, $T^{-1/2}t_\beta$ has a non-degenerate distribution, $t_\beta = O_p(T^{1/2})$, which is the same as in a pure time-series case. This implies that the t-statistic for $\widehat{\beta}$ has a divergent distribution.

The coefficient of determination is

$$
\begin{aligned}
R^2 &= \frac{\sum_{i=1}^{N}\sum_{t=1}^{T}(\widehat{y}_{it}-\overline{y}_i)^2}{\sum_{i=1}^{N}\sum_{t=1}^{T}(y_{it}-\overline{y}_i)^2} \\
&= \frac{\widehat{\beta}^2\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T^2}\sum_{t=1}^{T}(x_{it}-\overline{x}_i)^2}{\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T^2}\sum_{t=1}^{T}(y_{it}-\overline{y}_i)^2} \\
&= \frac{\widehat{\beta}^2\frac{1}{N}\sum_{i=1}^{N}\zeta_{2iT}}{\frac{1}{N}\sum_{i=1}^{N}\eta_{iT}} \\
&\xrightarrow{p} 0\cdot\frac{\sigma_\varepsilon^2}{\sigma_u^2} = 0,
\end{aligned}
$$

where

$$
\eta_{iT} \Rightarrow \sigma_u^2\left[\int_0^1 V_i(r)^2dr - \left(\int_0^1 V_i(r)dr\right)^2\right].
$$

This proves part (d).

Next, we consider the behavior of the Durbin-Watson (DW) statistic when $H_0 : \beta = 0$ is true. The DW statistic is

$$
\begin{aligned}
DW &= \frac{\sum_{i=1}^{N}\sum_{t=2}^{T}\left(\widehat{e}_{it}-\widehat{e}_{it-1}\right)^2}{\sum_{i=1}^{N}\sum_{t=1}^{T}\widehat{e}_{it}^2} \\
&= \frac{1}{T}\frac{\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T}\sum_{t=2}^{T}\left(u_{it}-\widehat{\beta}\varepsilon_{it}\right)^2}{\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T^2}\sum_{t=1}^{T}\left(y_{it}-\overline{y}_i-\widehat{\beta}(x_{it}-\overline{x}_i)\right)^2}.
\end{aligned} \tag{36}
$$

Now,

$$\begin{aligned}\frac{1}{N}\sum_{i=1}^{N}\left\{\frac{1}{T}\sum_{t=2}^{T}\left(u_{it}-\widehat{\beta}\varepsilon_{it}\right)^{2}\right\} &= \frac{1}{N}\sum_{i=1}^{N}\left\{\frac{1}{T}\sum_{t=2}^{T}u_{it}^{2}-\widehat{\beta}\frac{1}{T}\sum_{t=2}^{T}u_{it}\varepsilon_{it}+\frac{1}{T}\widehat{\beta}^{2}\sum_{t=2}^{T}\varepsilon_{it}^{2}\right\} \\ &= \frac{1}{N}\sum_{i=1}^{N}\left\{\frac{1}{T}\sum_{t=2}^{T}u_{it}^{2}\right\}-\widehat{\beta}\frac{1}{N}\sum_{i=1}^{N}\left\{\frac{1}{T}\sum_{t=2}^{T}u_{it}\varepsilon_{it}\right\}+ \\ &\quad \widehat{\beta}^{2}\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T}\sum_{t=2}^{T}\varepsilon_{it}^{2} \\ &\xrightarrow{p} \sigma_{u}^{2}.\end{aligned} \tag{37}$$

because

$$\widehat{\beta}=o_p(1),$$

$$\frac{1}{N}\sum_{i=1}^{N}\left\{\frac{1}{T}\sum_{t=2}^{T}u_{it}\varepsilon_{it}\right\}=o_p(1),$$

and

$$\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T}\sum_{t=2}^{T}\varepsilon_{it}^{2}=O_p(1).$$

From (35),

$$\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T^{2}}\sum_{t=1}^{T}\left(y_{it}-\overline{y}_{i}-\widehat{\beta}\left(x_{it}-\overline{x}_{i}\right)\right)^{2}\xrightarrow{p}\frac{\sigma_{u}^{2}}{6}. \tag{38}$$

Thus,

$$DW=O_p(T^{-1})$$

or

$$DW\xrightarrow{p}0. \tag{39}$$

It also follows that

$$TDW\xrightarrow{p}6. \tag{40}$$

Therefore, we establish parts (e) and (f). ■

C Proof of Lemma 2

Proof. First we note that

$$\begin{aligned}\begin{bmatrix} T^{-3/2}\sum_{t=1}^{T}\frac{y_{it}^{*}}{\sigma_{0v}} \\ T^{-3/2}\sum_{t=1}^{T}x_{it}^{*}\end{bmatrix} &= T^{-3/2}\sum_{t=1}^{T}L^{'}\begin{bmatrix} y_{it} \\ x_{it}\end{bmatrix} \\ &\Rightarrow \int_{0}^{1}\mathbf{W}_{i}^{*}(r)dr,\end{aligned} \tag{41}$$

and

$$\begin{array}{l} \left[\begin{array}{ll} \frac{1}{T^2}\sum_{t=1}^{T}\frac{(y_{it}^*)^2}{\sigma_{0v}^2} & \frac{1}{T^2}\sum_{t=1}^{T}\frac{y_{it}^* x_{it}^*}{\sigma_{0v}} \\ \frac{1}{T^2}\sum_{t=1}^{T}\frac{x_{it}^* y_{it}^*}{\sigma_{0v}} & \frac{1}{T^2}\sum_{t=1}^{T}(x_{it}^*)^2 \end{array} \right] \\ = L^{'}\left(\frac{1}{T^2}\sum_{t=1}^{T}\left[\begin{array}{c} y_{it} \\ x_{it} \end{array} \right] \left[\begin{array}{cc} y_{it} & x_{it} \end{array} \right] \right) L \\ \Rightarrow \int_0^1 \mathbf{W}_i^*(r)\left[\mathbf{W}_i^*(r)\right]^{'} dr \\ = \left[\begin{array}{ll} \int_0^1 V_i^*(r)^2 dr & \int_0^1 V_i^*(r) W_i^*(r) dr \\ \int_0^1 W_i^*(r) V_i^*(r) dr & \int_0^1 W_i^*(r)^2 dr \end{array} \right]. \end{array} \tag{42}$$

It follows that

$$\xi_{4iT} \quad \Rightarrow \quad \int_{0i}^1 W_i^*(r)^2 dr - \left\{ \int_0^1 W_i^*(r) dr \right\}^2 .$$

proving (b). Recall that

$$\zeta_{3iT} = \frac{1}{T^2}\sum_{t=1}^{T}\frac{(x_{it}^* - \overline{x}_i^*)\, y_{it}^*}{\sigma_{0v}} = \frac{1}{T^2}\sum_{t=1}^{T} x_{it}^* \frac{y_{it}^*}{\sigma_{0v}} - \left(T^{-1/2}\overline{x}_i^*\right)\left(T^{-1/2}\frac{\overline{y}_i^*}{\sigma_{0v}}\right).$$

Using (41) and (42), we have

$$\frac{1}{T^2}\sum_{t=1}^{T} x_{it}^* \frac{y_{it}^*}{\sigma_{0v}} \quad \Rightarrow \quad \int_0^1 W_i^*(r) V_i^*(r) dr,$$

$$T^{-1/2}\overline{x}_i^* \Rightarrow \int_0^1 W_i^*(r) dr,$$

and

$$T^{-1/2}\frac{\overline{y}_i}{\sigma_{0v}} \Rightarrow \int_0^1 V_i^*(r) dr.$$

Hence,

$$\zeta_{3iT} \Rightarrow \int_0^1 W_i^*(r) V_i^*(r) dr - \int_0^1 W_i^*(r) dr \int_0^1 V_i^*(r) dr.$$

This proves (a).

$(c) - (k)$ are immediate from the Lemma 1. ■

D Proof of Theorem 2

Proof. Using (a) in Theorem 1 we can show that

$$\widehat{\beta}^* \xrightarrow{p} 0. \tag{43}$$

Recall that

$$\widehat{\beta} \quad = \quad \widehat{\beta}^* \sigma_{0\varepsilon}^{-1} + \sigma_{0\varepsilon}^{-2}\sigma_{0u\varepsilon}.$$

Then we obtain

$$\widehat{\beta} \xrightarrow{p} \frac{\sigma_{0u\varepsilon}}{\sigma_{0\varepsilon}^2}, \tag{44}$$

proving (a). Using (b) in Theorem 1 we get

$$\sqrt{N}\widehat{\beta}^* \Rightarrow N\left(0, \frac{2\sigma_{0v}^2}{5}\right).$$

Then

$$\sqrt{N}\left(\widehat{\beta} - \frac{\sigma_{0u\varepsilon}}{\sigma_{0\varepsilon}^2}\right) \Rightarrow N\left(0, \frac{2\sigma_{0v}^2}{5\sigma_{0\varepsilon}^2}\right) \tag{45}$$

establishes (b). Now $T^{-1/2}t_{\beta^*} \Rightarrow N\left(0, \frac{2}{5}\right)$ from Theorem 1 and:

$$\begin{aligned}
&\tfrac{1}{\sigma_{0v}^2}\tfrac{1}{T}s^2 \\
&= \tfrac{1}{N}\textstyle\sum_{i=1}^N \tfrac{1}{T^2}\sum_{t=1}^T \tfrac{1}{\sigma_{0v}^2}\widehat{e}_{it}^2 \\
&= \tfrac{1}{N}\textstyle\sum_{i=1}^N \tfrac{1}{T^2}\sum_{t=1}^T \tfrac{1}{\sigma_{0v}^2}\widehat{e}_{it}^{*2} \\
&= \tfrac{1}{N}\textstyle\sum_{i=1}^N \left[\tfrac{1}{T^2}\sum_{t=1}^T \left(\tfrac{y_{it}^* - \overline{y}_i^*}{\sigma_{0v}}\right)^2 - \left(\widehat{\beta}^*\right)^2 \tfrac{1}{\sigma_{0v}^2}\tfrac{1}{T^2}\left(x_{it}^* - \overline{x}_i^*\right)^2\right] \\
&\xrightarrow{p} \tfrac{1}{6},
\end{aligned} \tag{46}$$

where

$$\widehat{e}_{it}^* = y_{it}^* - \widehat{\alpha}_i^* - \widehat{\beta}^* x_{it}^*,$$

$$\frac{1}{N}\sum_{i=1}^N \sum_{t=1}^T \frac{1}{T^2}\left(x_{it}^* - \overline{x}_i^*\right)^2 = O_p(1),$$

and

$$\widehat{\beta}^* = o_p(1).$$

Observe that

$$\begin{aligned}
T^{-1/2}t_{\beta} &= \frac{T^{-1/2}\widehat{\beta}\sqrt{\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \overline{x}_i)^2}}{s} \\
&= \frac{\sqrt{N}\widehat{\beta}\frac{1}{\sigma_{0v}}\sqrt{\frac{1}{N}\sum_{i=1}^N \frac{1}{T^2}\sum_{t=1}^T (x_{it} - \overline{x}_i)^2}}{\sqrt{\frac{1}{\sigma_{0v}^2}\frac{1}{T}s^2}} \\
&= \frac{\sqrt{N}\left(\widehat{\beta}^*\sigma_{0\varepsilon}^{-1} + \sigma_{0\varepsilon}^{-2}\sigma_{0u\varepsilon}\right)\frac{\sigma_{0\varepsilon}}{\sigma_{0v}}\sqrt{\frac{1}{N}\sum_{i=1}^N \frac{1}{T^2}\sum_{t=1}^T \left(x_{it}^* - \overline{x}_i^*\right)^2}}{\sqrt{\frac{1}{\sigma_{0v}^2}\frac{1}{T}s^2}} \\
&= \frac{\frac{\sqrt{N}}{\sigma_{0v}}\left(\widehat{\beta}^* + \frac{\sigma_{0u\varepsilon}}{\sigma_{0\varepsilon}}\right)\sqrt{\frac{1}{N}\sum_{i=1}^N \frac{1}{T^2}\sum_{t=1}^T \left(x_{it}^* - \overline{x}_i^*\right)^2}}{\sqrt{\frac{1}{\sigma_{0v}^2}\frac{1}{T}s^2}}.
\end{aligned}$$

It follows that

$$
\begin{aligned}
& T^{-1/2}t_{\beta} - \frac{T^{-1/2}\sigma_{0\varepsilon}^{-2}\sigma_{0u\varepsilon}\sqrt{\sum_{i=1}^{N}\sum_{t=1}^{T}(x_{it}-\overline{x}_{i})^{2}}}{s} \\
&= \frac{T^{-1/2}\widehat{\beta}\sqrt{\sum_{i=1}^{N}\sum_{t=1}^{T}(x_{it}-\overline{x}_{i})^{2}}-T^{-1/2}\sigma_{0\varepsilon}^{-2}\sigma_{0u\varepsilon}\sqrt{\sum_{i=1}^{N}\sum_{t=1}^{T}(x_{it}-\overline{x}_{i})^{2}}}{s} \\
&= \frac{T^{-1/2}\left(\widehat{\beta}-\sigma_{0\varepsilon}^{-2}\sigma_{0u\varepsilon}\right)\sqrt{\sum_{i=1}^{N}\sum_{t=1}^{T}(x_{it}-\overline{x}_{i})^{2}}}{s} \\
&= \frac{T^{-1/2}\left(\widehat{\beta}^{*}\sigma_{0\varepsilon}^{-1}\right)\sqrt{\sum_{i=1}^{N}\sum_{t=1}^{T}(x_{it}-\overline{x}_{i})^{2}}}{s} \\
&= \frac{\frac{\sqrt{N}}{\sigma_{0v}}\left(\widehat{\beta}^{*}\right)\sqrt{\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T^{2}}\sum_{t=1}^{T}\left(x_{it}^{*}-\overline{x}_{i}^{*}\right)^{2}}}{\sqrt{\frac{1}{\sigma_{0v}^{2}}\frac{1}{T}s^{2}}} \\
&\Rightarrow \frac{N\left(0,\frac{2}{5}\right)\cdot\sqrt{\frac{1}{6}}}{\sqrt{\frac{1}{6}}} \\
&= N\left(0,\tfrac{2}{5}\right).
\end{aligned}
$$

proving (c). It is easy to show that

$$
\begin{aligned}
& \tfrac{1}{N}\sum_{i=1}^{N}\tfrac{1}{T^{2}}\sum_{t=1}^{T}\left(x_{it}-\overline{x}_{i}\right)^{2} \\
&= \sigma_{0\varepsilon}^{2}\tfrac{1}{N}\sum_{i=1}^{N}\tfrac{1}{T^{2}}\sum_{t=1}^{T}\left(x_{it}^{*}-\overline{x}_{i}^{*}\right)^{2} \\
&= \sigma_{0\varepsilon}^{2}\tfrac{1}{N}\sum_{i=1}^{N}\zeta_{4iT} \\
&\xrightarrow{p} \tfrac{\sigma_{0\varepsilon}^{2}}{6},
\end{aligned}
\tag{47}
$$

and

$$
\begin{aligned}
& \tfrac{1}{N}\sum_{1}^{N}\tfrac{1}{T^{2}}\sum_{t=1}^{T}\left(y_{it}-\overline{y}_{i}\right)^{2} \\
&= \tfrac{1}{N}\sum_{i=1}^{N}\tfrac{1}{T^{2}}\sum_{t=1}^{T}\left(y_{it}^{*}+\tfrac{\sigma_{0u\varepsilon}}{\sigma_{0\varepsilon}}x_{it}^{*}-\overline{y}_{it}^{*}-\tfrac{\sigma_{0u\varepsilon}}{\sigma_{0\varepsilon}}\overline{x}_{it}^{*}\right)^{2} \\
&= \tfrac{1}{N}\sum_{i=1}^{N}\tfrac{1}{T^{2}}\sum_{t=1}^{T}\left[\left(y_{it}^{*}-\overline{y}_{it}^{*}\right)^{2}+\left(\tfrac{\sigma_{0u\varepsilon}}{\sigma_{0\varepsilon}}\right)^{2}\left(x_{it}^{*}-\overline{x}_{it}^{*}\right)^{2}-2\tfrac{\sigma_{0u\varepsilon}}{\sigma_{0\varepsilon}}\left(y_{it}^{*}-\overline{y}_{it}^{*}\right)\left(x_{it}^{*}-\overline{x}_{it}^{*}\right)\right] \\
&\xrightarrow{p} \tfrac{\sigma_{0v}^{2}}{6}+\left(\tfrac{\sigma_{0u\varepsilon}}{\sigma_{0\varepsilon}}\right)^{2}\tfrac{1}{6}.
\end{aligned}
\tag{48}
$$

To prove (d), we use (47) and (48) to obtain

$$
\begin{aligned}
R^{2} &= \frac{\sum_{i=1}^{N}\sum_{t=1}^{T}(\widehat{y}_{it}-\overline{y}_{i})^{2}}{\sum_{i=1}^{N}\sum_{t=1}^{T}(y_{it}-\overline{y}_{i})^{2}} \\
&= \frac{\widehat{\beta}^{2}\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T^{2}}\sum_{t=1}^{T}(x_{it}-\overline{x}_{i})^{2}}{\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T^{2}}\sum_{t=1}^{T}(y_{it}-\overline{y}_{i})^{2}} \\
&\xrightarrow{p} \frac{\left(\frac{\sigma_{0u\varepsilon}}{\sigma_{0\varepsilon}^{2}}\right)^{2}\frac{\sigma_{0\varepsilon}^{2}}{6}}{\frac{\sigma_{0v}^{2}}{6}+\left(\frac{\sigma_{0u\varepsilon}}{\sigma_{0\varepsilon}}\right)^{2}\frac{1}{6}} \\
&= \frac{\sigma_{0u\varepsilon}^{2}}{\sigma_{0v}^{2}\sigma_{0\varepsilon}^{2}+\sigma_{0u\varepsilon}^{2}}.
\end{aligned}
$$

Again using (12) gives

$$\begin{aligned}
&DW \\
&= \frac{\sum_{i=1}^{N}\sum_{t=2}^{T}\left(\widehat{e}_{it}-\widehat{e}_{it-1}\right)^2}{\sum_{i=1}^{N}\sum_{t=1}^{T}\widehat{e}_{it}^2} \\
&= \frac{\sum_{i=1}^{N}\sum_{t=2}^{T}\left(\widehat{e}_{it}^*-\widehat{e}_{it-1}^*\right)^2}{\sum_{i=1}^{N}\sum_{t=1}^{T}\widehat{e}_{it}^{*2}} \\
&= \frac{1}{T}\frac{\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T}\sum_{t=2}^{T}\left(u_{it}^*-\widehat{\beta}\varepsilon_{it}^*\right)^2}{\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T^2}\sum_{t=1}^{T}\widehat{e}_{it}^{*2}}.
\end{aligned}$$

Now,

$$\begin{aligned}
&\tfrac{1}{N}\textstyle\sum_{i=1}^{N}\tfrac{1}{T}\sum_{t=2}^{T}\left(u_{it}^*-\widehat{\beta}^*\varepsilon_{it}^*\right)^2 \\
&= \tfrac{1}{N}\textstyle\sum_{i=1}^{N}\left\{\tfrac{1}{T}\sum_{t=2}^{T}u_{it}^{*2}-\widehat{\beta}^*\tfrac{1}{T}\sum_{t=2}^{T}u_{it}^*\varepsilon_{it}^*+\tfrac{1}{T}\left(\widehat{\beta}^*\right)^2\sum_{t=2}^{T}\varepsilon_{it}^{*2}\right\} \\
&= \tfrac{1}{N}\textstyle\sum_{i=1}^{N}\left\{\tfrac{1}{T}\sum_{t=2}^{T}u_{it}^{*2}\right\}-\widehat{\beta}^*\tfrac{1}{N}\sum_{i=1}^{N}\left\{\tfrac{1}{T}\sum_{t=2}^{T}u_{it}^*\varepsilon_{it}^*\right\}+\widehat{\beta}^{*2}\tfrac{1}{N}\sum_{i=1}^{N}\tfrac{1}{T}\sum_{t=2}^{T}\varepsilon_{it}^{*2} \\
&\xrightarrow{p}\sigma_{0v}^2
\end{aligned}$$

because

$$\frac{1}{N}\sum_{i=1}^{N}\left\{\frac{1}{T}\sum_{t=2}^{T}u_{it}^{*2}\right\}\xrightarrow{p}\sigma_{0v}^2,$$

$$\widehat{\beta}^*=o_p(1),$$

$$\frac{1}{N}\sum_{i=1}^{N}\left\{\frac{1}{T}\sum_{t=2}^{T}u_{it}^*\varepsilon_{it}^*\right\}=o_p(1),$$

and

$$\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T}\sum_{t=2}^{T}\varepsilon_{it}^{*2}=O_p(1).$$

From (46),

$$\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T^2}\sum_{t=1}^{T}\widehat{e}_{it}^{*2}\xrightarrow{p}\frac{\sigma_{0v}^2}{6}. \tag{49}$$

Thus,

$$DW=O_p(T^{-1})$$

or

$$DW\xrightarrow{p}0. \tag{50}$$

It also follows that

$$TDW\xrightarrow{p}6. \tag{51}$$

Therefore, we establish parts (e) and (f). ■

E Proof of Theorem 3

Proof. First, we note that

$$\begin{aligned} \widehat{e}^*_{it-1} &= y^*_{it-1} - \widehat{\alpha}^*_i - \widehat{\beta}^* x^*_{it-1} \\ &= y^*_{it-1} - \overline{y}^*_i - \widehat{\beta}^* \left(x^*_{it-1} - \overline{x}^*_i \right) \end{aligned}$$

and

$$\triangle \widehat{e}^*_{it} = u^*_{it} - \widehat{\beta}^* \varepsilon^*_{it},$$

where

$$u^*_{it} = u_{it} - \frac{\sigma_{0u\varepsilon}}{\sigma^2_{0\varepsilon}} \varepsilon_{it}$$

and

$$\varepsilon^*_{it} = \frac{\varepsilon_{it}}{\sigma_{0\varepsilon}}.$$

It follows that

$$\begin{aligned} \xi_{6NT} &= \tfrac{1}{N} \textstyle\sum_{i=1}^N \left\{ \tfrac{1}{T^2} \sum_{t=2}^T \widehat{e}^{*2}_{it-1} \right\} \\ &= \tfrac{1}{N} \textstyle\sum_{i=1}^N \left\{ \tfrac{1}{T^2} \sum_{t=2}^T \left(y^*_{it-1} - \overline{y}^*_i \right)^2 \right\} - \widehat{\beta}^{*2} \tfrac{1}{N} \sum_{i=1}^N \left\{ \tfrac{1}{T^2} \sum_{t=2}^T \left(x^*_{it-1} - \overline{x}^*_i \right)^2 \right\} \\ &= \tfrac{1}{N} \textstyle\sum_{i=1}^N \left\{ \tfrac{1}{T^2} \sum_{t=2}^T \left(y^*_{it-1} - \overline{y}^*_i \right)^2 \right\} + o_p(1), \end{aligned} \tag{52}$$

since

$$\frac{1}{N} \sum_{i=1}^N \left\{ \frac{1}{T^2} \sum_{t=2}^T \left(x^*_{it-1} - \overline{x}^*_i \right)^2 \right\} = O_p(1)$$

and

$$\widehat{\beta}^{*2} = o_p(1).$$

We also note that

$$\frac{1}{T^2} \sum_{t=2}^T \left(y^*_{it-1} - \overline{y}^*_i \right)^2 \Rightarrow \sigma^2_{0v} \left[\int_0^1 V_i^{*2}(r) dr - \left\{ \int_0^1 V_i^*(r) dr \right\}^2 \right] \equiv \zeta_{6i}, \tag{53}$$

$$E \left[\zeta_{6i} \right] = \frac{\sigma^2_{0v}}{6} \equiv \mu_6,$$

and

$$Var \left[\zeta_{6i} \right] = \frac{\sigma^4_{0v}}{45} \equiv \sigma^2_6.$$

Thus,

$$\frac{1}{N} \sum_{i=1}^N \left\{ \frac{1}{T^2} \sum_{t=2}^T \left(y^*_{it} - \overline{y}^*_i \right)^2 \right\} \xrightarrow{p} \frac{\sigma^2_{0v}}{6} \equiv \mu_6 \tag{54}$$

and

$$\zeta_{6NT} \xrightarrow{p} \frac{\sigma_{0v}^{2}}{6} \equiv \mu_{6}$$

by the law of large numbers for triangular arrays and equation (52). Moreover,

$$\begin{aligned}
&\sqrt{N}\tfrac{1}{N}\textstyle\sum_{i=1}^{N}\tfrac{1}{T}\sum_{t=2}^{T}\widehat{e}_{it-1}^{*}\bigtriangleup\widehat{e}_{it}^{*}\\
&=\sqrt{N}\tfrac{1}{N}\textstyle\sum_{i=1}^{N}\tfrac{1}{T}\sum_{t=2}^{T}\left[y_{it-1}^{*}-\overline{y}_{i}^{*}-\widehat{\beta}^{*}\left(x_{it-1}^{*}-\overline{x}_{i}^{*}\right)\right]\left[u_{it}^{*}-\widehat{\beta}^{*}\varepsilon_{it}^{*}\right]\\
&=\sqrt{N}\tfrac{1}{N}\textstyle\sum_{i=1}^{N}\tfrac{1}{T}\sum_{t=2}^{T}\left[\left(y_{it-1}^{*}-\overline{y}_{i}^{*}\right)u_{it}^{*}-\widehat{\beta}^{*}\left(x_{it-1}^{*}-\overline{x}_{i}^{*}\right)u_{it}^{*}-\widehat{\beta}^{*}\left(y_{it-1}^{*}-\overline{y}_{i}^{*}\right)\varepsilon_{it}^{*}+\widehat{\beta}^{*2}\left(x_{it-1}^{*}-\overline{x}_{i}^{*}\right)\right]\varepsilon_{it}^{*}\\
&=\sqrt{N}\tfrac{1}{N}\textstyle\sum_{i=1}^{N}\tfrac{1}{T}\sum_{t=2}^{T}\left(y_{it-1}^{*}-\overline{y}_{i}^{*}\right)u_{it}^{*}+o_{p}(1),
\end{aligned}$$

since

$$\frac{1}{N}\sum_{i=1}^{N}\left\{\frac{1}{T}\sum_{t=2}^{T}\left(x_{it-1}^{*}-\overline{x}_{i}^{*}\right)\varepsilon_{it}^{*}\right\}\xrightarrow{p}-\frac{1}{2},$$

$$\frac{1}{N}\sum_{i=1}^{N}\left\{\frac{1}{T}\sum_{t=2}^{T}\left(y_{it-1}^{*}-\overline{y}_{i}^{*}\right)\varepsilon_{it}^{*}\right\}=o_{p}(1),$$

and

$$\frac{1}{N}\sum_{i=1}^{N}\left\{\frac{1}{T}\sum_{t=2}^{T}\left(x_{it-1}^{*}-\overline{x}_{i}^{*}\right)u_{it}^{*}\right\}=o_{p}(1)$$

by the law of large numbers for triangular arrays. Hence,

$$\begin{aligned}
\sqrt{N}T\left(\widehat{\rho}-1\right) &= \frac{\sqrt{N}\xi_{5NT}}{\xi_{6NT}}\\
&= \frac{\sqrt{N}\frac{1}{N}\sum_{i=1}^{N}\left\{\frac{1}{T}\sum_{t=2}^{T}\left(y_{it-1}^{*}-\overline{y}_{i}^{*}\right)u_{it}^{*}\right\}}{\frac{1}{N}\sum_{i=1}^{N}\left\{\frac{1}{T^{2}}\sum_{t=2}^{T}\left(y_{it-1}^{*}-\overline{y}_{i}^{*}\right)^{2}\right\}}+o_{p}(1).
\end{aligned}\tag{55}$$

Next, we derive the asymptotic distribution of

$$\begin{aligned}
\frac{\sqrt{N}\frac{1}{N}\sum_{i=1}^{N}\left\{\frac{1}{T}\sum_{t=2}^{T}\left(y_{it-1}^{*}-\overline{y}_{i}^{*}\right)u_{it}^{*}\right\}}{\frac{1}{N}\sum_{i=1}^{N}\left\{\frac{1}{T^{2}}\sum_{t=2}^{T}\left(y_{it-1}^{*}-\overline{y}_{i}^{*}\right)^{2}\right\}} &= \frac{\sqrt{N}\xi_{5NT}^{'}}{\xi_{6NT}}\\
&= \frac{\sqrt{N}\frac{1}{N}\sum_{i=1}^{N}\zeta_{5iT}^{'}}{\frac{1}{N}\sum_{i=1}^{N}\zeta_{6iT}},
\end{aligned}$$

where

$$\zeta_{5iT}^{'}=\frac{1}{T}\sum_{t=2}^{T}\left(y_{it-1}^{*}-\overline{y}_{i}^{*}\right)u_{it}^{*}$$

and

$$\xi_{5NT}^{'}=\frac{1}{N}\sum_{i=1}^{N}\zeta_{5iT}^{'}.$$

But

$$\tfrac{1}{T}\textstyle\sum_{t=2}^{T}\left(y_{it-1}^{*}-\overline{y}_{i}^{*}\right)u_{it}^{*} \quad = \quad T^{-1}\sum_{t=2}^{T}y_{it-1}^{*}u_{it}^{*}-\left(T^{-3/2}\sum_{t=2}^{T}y_{it}^{*}\right)\left(T^{-1/2}\sum_{t=2}^{T}u_{it}^{*}\right),$$

$$T^{-1}\sum_{t=2}^{T} y^*_{it-1}u^*_{it} \quad\Rightarrow\quad \tfrac{\sigma^2_{0v}}{2}\left[V_i^*(1)^2 - \tfrac{\sigma^2_v}{\sigma^2_{0v}}\right],$$

$$T^{-3/2}\sum_{t=2}^{T} y^*_{it} \Rightarrow \sigma_{0v}\int_0^1 V_i^*(r)dr,$$

and

$$T^{-1/2}\sum_{t=2}^{T} u^*_{it} \Rightarrow \sigma_{0v}V_i^*(1).$$

Clearly,

$$\frac{1}{T}\sum_{t=2}^{T}\left(y^*_{it-1} - \overline{y}^*_i\right)u^*_{it} \Rightarrow \frac{\sigma^2_{0v}}{2}\left[V_i^*(1)^2 - \frac{\sigma^2_v}{\sigma^2_{0v}}\right] - \sigma^2_{0v}V_i^*(1)\int_0^1 V_i^*(r)dr \equiv \zeta_{5i}, \tag{56}$$

$$E\left[\zeta_{5i}\right] = -\frac{\sigma^2_v}{2} = \mu_5,$$

and

$$Var\left[\zeta_{5i}\right] = \frac{\sigma^4_{0v}}{12} = \sigma^2_5,$$

since

$$E\left[V_i^*(1)\int_0^1 V_i^*(r)dr\right] = \frac{1}{2}$$

and

$$E\left[V_i^*(1)^2\right] = 1.$$

Let $\mu_{5T} \equiv E\left[\zeta'_{5iT}\right] = \mu_5 + O(T^{-1})$ and $\mu_{6T} \equiv E\left[\zeta_{6iT}\right] = \mu_6 + O(T^{-1})$. The next step is to find an appropriate normalization of $\left(\frac{\xi'_{5NT}}{\xi_{6NT}} - \frac{\mu_{5T}}{\mu_{6T}}\right)$ to make sure it converges to a proper random variable. For this purpose, we note the following standardization

$$\frac{\sqrt{N}\xi'_{5NT}}{\xi_{6NT}} - \frac{\sqrt{N}\mu_{5T}}{\mu_{6T}} = \frac{\sqrt{N}\left(\zeta'_{\xi NT} - \mu_{1T}\right)}{\xi_{6NT}} + \xi'_{5NT}\sqrt{N}\left[\frac{1}{\xi_{6NT}} - \frac{1}{\mu_{6T}}\right]. \tag{57}$$

Since $\{\xi'_{5iT}\}$ satisfies the conditions of the central limit theorem for triangular arrays, we can readily see that the $\sqrt{N}\left(\xi'_{5NT} - \mu_{5T}\right)$ will converge to a normal variable with an appropriate normalization and the ξ_{6NT} will converge to μ_6 in probability. Using the Slutsky theorem, we obtain

$$\frac{\sqrt{N}\left(\xi'_{5NT} - \mu_{5T}\right)}{\xi_{6NT}} \Rightarrow N\left(0, \frac{1}{\mu_6^2}\sigma_5^2\right)$$

as $N \to \infty$ and $T \to \infty$. Similarly,

$$\xi'_{5NT}\sqrt{N}\left[\frac{1}{\xi_{6NT}} - \frac{1}{\mu_{6T}}\right] \Rightarrow N\left(0, \frac{\mu_5^2}{\mu_6^4}\sigma_6^2\right), \tag{58}$$

since it can be shown that $cov\left(\zeta_{5iT}^{'}, \zeta_{6iT}\right) = O(T^{-1})$ and $\sqrt{N}\left[\xi_{5NT} - \mu_{1T}\right]$ and $\sqrt{N}\left[\xi_{6NT} - \mu_{2T}\right]$ are asymptotically uncorrelated. We therefore conclude

$$\sqrt{N}\left[\frac{\sqrt{N}\xi_{5NT}^{'}}{\xi_{6NT}} - \frac{\mu_{5T}}{\mu_{6T}}\right] \Rightarrow N\left(0, \frac{1}{\mu_6^2}\sigma_5^2 + \frac{\mu_5^2}{\mu_6^4}\sigma_6^2\right), \tag{59}$$

where

$$\begin{aligned} \frac{\sigma_5^2}{\mu_6^2} &= \frac{\frac{\sigma_{0v}^4}{12}}{\left(\frac{\sigma_{0v}^2}{6}\right)^2} \\ &= 3 \end{aligned}$$

and

$$\begin{aligned} \frac{\mu_5^2}{\mu_6^4}\sigma_6^2 &= \frac{\left(\frac{\sigma_v^2}{2}\right)^2}{\left(\frac{\sigma_{0v}^2}{6}\right)^4}\frac{\sigma_{0v}^4}{45} \\ &= \frac{7.2\sigma_v^4}{\sigma_{0v}^4}. \end{aligned}$$

Thus,

$$\sqrt{N}T\left(\widehat{\rho} - 1\right) - \frac{\sqrt{N}\mu_{5T}}{\mu_{6T}} \Rightarrow N\left(0, 3 + \frac{7.2\sigma_v^4}{\sigma_{0v}^4}\right). \tag{60}$$

The t-statistic is

$$\begin{aligned} t_\rho &= \frac{\left(\widehat{\rho}-1\right)\sqrt{\sum_{i=1}^{N}\sum_{t=2}^{T}\widehat{e}_{it-1}^{*2}}}{s_e} \\ &= \frac{\sqrt{N}T\left(\widehat{\rho}-1\right)\sqrt{\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T^2}\sum_{t=2}^{T}\widehat{e}_{it-1}^{*2}}}{s_e} \\ &= \frac{\sqrt{N}T\left(\widehat{\rho}-1\right)\sqrt{\xi_{6NT}}}{s_e} \\ &= \frac{\frac{\sqrt{N}\xi_{5NT}}{\zeta_{6NT}}\sqrt{\xi_{6NT}}}{s_e} \\ &= \frac{\sqrt{N}\xi_{5NT}}{s_e\sqrt{\xi_{6NT}}} \\ &= \frac{\sqrt{N}\xi_{5NT}^{'}}{s_e\sqrt{\xi_{6NT}}} + o(1), \end{aligned} \tag{61}$$

where $s_e^2 = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=2}^{T}\left(\widehat{e}_{it}^{*} - \widehat{\rho}\widehat{e}_{it-1}^{*}\right)^2$. It follows that

$$\begin{aligned} s_e^2 &= \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=2}^{T}\left\{\left(u_{it}^{*} - \widehat{\beta}^{*}\varepsilon_{it}^{*}\right)\right\}^2 + o_p(1) \\ &= \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=2}^{T}\left(u_{it}^{*2} - 2\widehat{\beta}^{*}u_{it}^{*}\varepsilon_{it}^{*} + \widehat{\beta}^{*2}\varepsilon_{it}^{*2}\right) + o_p(1) \\ &\xrightarrow{p} \sigma_v^2. \end{aligned}$$

That is, s_e^2 is a consistent estimator of σ_v^2. Next, we derive the asymptotic distribution of

$$\frac{\sqrt{N}\xi_{5NT}^{'}}{s_e\sqrt{\xi_{6NT}}}.$$

We use the same approach for the asymptotic distribution of $\frac{\sqrt{N}\xi_{5NT}^{'}}{s\sqrt{\xi_{6NT}}}$ as we used for $\sqrt{N}T\left(\widehat{\rho}-1\right)$. We make the following normalization:

$$\begin{aligned} \frac{\sqrt{N}\xi_{5NT}^{'}}{s_e\sqrt{\xi_{6NT}}}-\frac{\sqrt{N}\mu_{5T}}{s_e\sqrt{\mu_{6T}}} &= \frac{\sqrt{N}\xi_{5NT}^{'}}{s_e\sqrt{\xi_{6NT}}}-\frac{\sqrt{N}\mu_{5T}}{s_e\sqrt{\mu_{6T}}} \\ &= \frac{\sqrt{N}\left[\xi_{5NT}^{'}-\mu_{5T}\right]}{s_e\sqrt{\xi_{6NT}}}+\frac{\mu_{5T}}{s_e}\sqrt{N}\left[\sqrt{\frac{1}{\xi_{6NT}}}-\sqrt{\frac{1}{\mu_{6T}}}\right]. \end{aligned} \tag{62}$$

Since

$$\frac{\sqrt{N}\left[\xi_{5NT}^{'}-\mu_{5T}\right]}{s_e\sqrt{\xi_{6NT}}}\Rightarrow N\left(0,\frac{\sigma_5^2}{\mu_6\sigma_v^2}\right),$$

$$\frac{\mu_{5T}}{s_e}\sqrt{N}\left[\sqrt{\frac{1}{\xi_{6NT}}}-\sqrt{\frac{1}{\mu_{6T}}}\right]\Rightarrow N\left(0,\frac{1}{4}\frac{\mu_5^2\sigma_6^2}{\mu_6^3\sigma_v^2}\right),$$

and $\sqrt{N}\left[\xi_{5NT}^{'}-\mu_{5T}\right]$ and $\sqrt{N}\left[\xi_{6NT}-\mu_{6T}\right]$ are asymptotically uncorrelated, we note that

$$\frac{\sqrt{N}\xi_{5NT}^{'}}{s_e\sqrt{\xi_{6NT}}}-\frac{\sqrt{N}\mu_{5T}}{s_e\sqrt{\mu_{6T}}}\Rightarrow N\left(0,\frac{\sigma_5^2}{\mu_6\sigma_v^2}+\frac{1}{4}\frac{\mu_5^2\sigma_2^2}{\mu_6^3\sigma_v^2}\right)$$

where

$$\begin{aligned} \frac{\sigma_5^2}{\mu_6\sigma_v^2} &= \frac{\frac{\sigma_{0v}^4}{12}}{\frac{\sigma_{0v}^2}{6}\sigma_v^2} \\ &= \frac{\sigma_{0v}^2}{2\sigma_v^2} \end{aligned}$$

and

$$\begin{aligned} \frac{1}{4}\frac{\mu_5^2\sigma_6^2}{\mu_6^3\sigma_v^2} &= \frac{1}{4}\frac{\left(\frac{\sigma_v^2}{2}\right)^2\frac{\sigma_{0v}^4}{45}}{\left(\frac{\sigma_{0v}^2}{6}\right)^3\sigma_v^2} \\ &= \frac{3\sigma^2v}{10\sigma_{0v}^2}. \end{aligned}$$

Thus,

$$t_\rho-\frac{\sqrt{N}\mu_{5T}}{s_e\sqrt{\mu_{6T}}}\Rightarrow N\left(0,\frac{\sigma_{0v}^2}{2\sigma_v^2}+\frac{3\sigma_v^2}{10\sigma_{0v}^2}\right).$$

Therefore, we established Theorem 3. ■

F The Proof of Theorem 4

Proof. Observe that

$$\begin{aligned}
\xi_{8NT} &= \tfrac{1}{N}\sum_{i=1}^{N}\zeta_{8iT} \\
&= \tfrac{1}{N}\sum_{i=1}^{N}\tfrac{1}{T^2}\left(e_i^{*\prime}Q_i^{*}e_i^{*}\right) \\
&= \tfrac{1}{N}\sum_{i=1}^{N}\tfrac{1}{T^2}\left(e_i^{*\prime}e_i^{*}\right)+o_p(1) \\
&= \tfrac{1}{N}\sum_{i=1}^{N}\tfrac{1}{T^2}\sum_{t=2}^{T}\left(\widehat{e}_{it-1}^{*2}\right)+o_p(1) \\
&= \tfrac{1}{N}\textstyle\sum_{i=1}^{N}\left\{\tfrac{1}{T^2}\sum_{t=2}^{T}\left(y_{it}^{*}-\overline{y}_i^{*}\right)^2\right\}+o_p(1).
\end{aligned}$$

We also note that

$$\frac{1}{T^2}\sum_{t=2}^{T}\left(y_{it}^{*}-\overline{y}_i^{*}\right)^2 \Rightarrow \sigma_{0v}^2\left[\int_0^1 V_i^{*2}(r)dr-\left\{\int_0^1 V_i^{*}(r)dr\right\}^2\right]\equiv\zeta_{8i}, \tag{63}$$

$$E\left[\zeta_{8i}\right]=\frac{\sigma_{0v}^2}{6}\equiv\mu_8,$$

and

$$Var\left[\zeta_{8i}\right]=\frac{\sigma_{0v}^4}{45}\equiv\sigma_8^2.$$

Under the null hypothesis that $\rho=1$, expression (16) can be written as

$$d(L)\bigtriangleup\widehat{e}_{it}^{*}=v_{itp}, \tag{64}$$

where $d(L)=\left(1-\varphi_1L-\varphi_2L^2-\cdots\varphi_pL^p\right)$ and L is the backshift operator. Moreover,

$$\begin{aligned}
\sqrt{N}\xi_{7NT} &= \sqrt{N}\tfrac{1}{N}\sum_{i=1}^{N}\zeta_{7NT} \\
&= \sqrt{N}\tfrac{1}{N}\sum_{i=1}^{N}\tfrac{1}{T}\left(e_i^{*\prime}Q_i^{*}v_i\right) \\
&= \sqrt{N}\tfrac{1}{N}\sum_{i=1}^{N}\tfrac{1}{T}\left(e_i^{*\prime}v_i\right)+o_p(1) \\
&= \sqrt{N}\tfrac{1}{N}\sum_{i=1}^{N}\tfrac{1}{T}\sum_{t=2}^{T}\left(\widehat{e}_{it-1}^{*}v_{itp}\right)+o_p(1) \\
&= \sqrt{N}\tfrac{1}{N}\textstyle\sum_{i=1}^{N}\left\{\tfrac{1}{T}\sum_{t=2}^{T}\left(y_{it}^{*}-\overline{y}_i^{*}\right)d(L)u_{it}^{*}\right\}+o_p(1).
\end{aligned}$$

Since

$$\begin{aligned}
&\sqrt{N}\tfrac{1}{N}\sum_{i=1}^{N}\tfrac{1}{T}\sum_{t=1}^{T}\widehat{e}_{it-1}^{*}v_{it} \\
&=\sqrt{N}\tfrac{1}{N}\sum_{i=1}^{N}\tfrac{1}{T}\left[\sum_{t=2}^{T}\widehat{e}_{it-1}^{*}\left(1-\varphi_1L-\varphi_2L^2-\cdots-\varphi_pL^p\right)\bigtriangleup\widehat{e}_{it}^{*}\right] \\
&=\sqrt{N}\tfrac{1}{N}\sum_{i=1}^{N}\tfrac{1}{T}\left[\sum_{t=2}^{T}\widehat{e}_{it-1}^{*}\bigtriangleup\widehat{e}_{it}^{*}-\varphi_1\widehat{e}_{it-1}^{*}\bigtriangleup\widehat{e}_{it-1}^{*}-\varphi_2\widehat{e}_{it-1}^{*}\bigtriangleup\widehat{e}_{it-2}^{*}-\cdots-\varphi_p\widehat{e}_{it-1}^{*}\bigtriangleup\widehat{e}_{it-p}^{*}\right] \\
&=\tfrac{1}{\sqrt{NT}}\textstyle\sum_{i=1}^{N}\sum_{t=2}^{T}\left[\left(y_{it-1}^{*}-\overline{y}_i^{*}\right)\left(u_{it}-\varphi_1u_{it-1}-\varphi_2u_{it-2}-\cdots-\varphi_pu_{it-p}\right)\right]+o_p(1),
\end{aligned}$$

it follows that

$$\begin{aligned}\sqrt{N}\xi_{7NT} &= \sqrt{N}\tfrac{1}{N}\sum_{i=1}^{N}\tfrac{1}{T}\sum_{t=2}^{T}\widehat{e}^{*}_{it-1}v_{it}+o_p\left(1\right)\\ &= \tfrac{1}{\sqrt{N}T}\textstyle\sum_{i=1}^{N}\sum_{t=2}^{T}\left(y^{*}_{it-1}-\overline{y}^{*}_{i}\right)\left(u_{it}-\varphi_1 u_{it-1}-\varphi_2 u_{it-2}-\cdots-\varphi_p u_{it-p}\right)+o_p(1).\end{aligned}$$

Hence,

$$\begin{aligned}t_{ADF} &= \frac{\sqrt{N}\xi_{7NT}}{s_v\sqrt{\xi_{8NT}}}\\ &= \frac{\sqrt{N}\frac{1}{N}\sum\limits_{i=1}^{N}\left\{\frac{1}{T}\sum\limits_{t=2}^{T}(y_{it-1}-\overline{y}_i)d(L)u^{*}_{it}\right\}}{s_v\sqrt{\xi_{8NT}}}+o_p\left(1\right)\\ &= \frac{\sqrt{N}\xi^{'}_{7NT}}{s_v\sqrt{\xi_{8NT}}}+o_p\left(1\right),\end{aligned}$$

where

$$\zeta^{'}_{7iT}=\frac{1}{T}\sum_{t=2}^{T}\left(y^{*}_{it-1}-\overline{y}^{*}_{i}\right)d\left(L\right)u^{*}_{it}$$

and

$$\xi^{'}_{7NT}=\frac{1}{N}\sum_{i=1}^{N}\zeta^{'}_{7iT}.$$

Assume that the sequence $\{u^{*}_{it}\}$ satisfies condition $(C2)$ in Phillips and Ouliaris (1990) so that

$$\frac{1}{T}\sum_{t=2}^{[rT]}d\left(L\right)u^{*}_{it}\Rightarrow d\left(1\right)\sigma_{0v}V^{*}_{i}\left(r\right).$$

Consequently,

$$\frac{1}{T}\sum_{t=2}^{T}y^{*}_{it-1}d\left(L\right)u^{*}_{it}\Rightarrow d\left(1\right)\sigma^{2}_{0v}\int_0^1 V^{*}_{i}\left(r\right)dV^{*}_{i}\left(r\right)$$

and

$$\frac{1}{T}\sum_{t=2}^{T}\overline{y}^{*}_{i}d\left(L\right)u^{*}_{it}\Rightarrow d\left(1\right)\sigma^{2}_{0v}V^{*}_{i}\left(1\right)\int_0^1 V^{*}_{i}\left(r\right)dr.$$

Clearly,

$$\frac{1}{T}\sum_{t=2}^{T}\left(y^{*}_{it-1}-\overline{y}^{*}_{i}\right)d\left(L\right)u^{*}_{it}\Rightarrow d\left(1\right)\sigma^{2}_{0v}\int_0^1 V^{*}_{i}\left(r\right)dV^{*}_{i}\left(r\right)-d\left(1\right)\sigma^{2}_{0v}V^{*}_{i}\left(1\right)\int_0^1 V^{*}_{i}\left(r\right)dr\equiv\zeta_{7i}$$

with

$$E\left(\zeta_{7i}\right)=-d\left(1\right)\frac{\sigma^{2}_{0v}}{2}$$

and

$$Var\left(\zeta_{7i}\right)=d^{2}\left(1\right)\frac{\sigma^{4}_{0v}}{12}.$$

Let

$$s_v^2 = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=2}^{T} \left\{ \left(\widehat{e}_{it}^* - \widehat{\rho}^* \widehat{e}_{it-1} \right) - \left(\widehat{\varphi}_1 L + \widehat{\varphi}_2 L^2 + \cdots + \widehat{\varphi}_p L^p \right) \triangle \widehat{e}_{it}^* \right\}^2 .$$

Observe that $\{\widehat{e}_{it}^*\}$ are linear combinations of u_{it}^* and ε_{it}^*. Those linear combinations are in general ARMA processes. Therefore, we need to have $p \to \infty$ to capture the true structure of v_{itp}. As a result of $p \to \infty$, $\widehat{\rho} \to 1$ and the estimated coefficients $\{\widehat{\varphi}_i\}$ may well represent $\{\varphi_i\}$:

$$\begin{aligned} s_v^2 &= \tfrac{1}{NT} \sum_{i=1}^{N} \sum_{t=2}^{T} \left\{ \left(1 - \varphi_1 L - \varphi_2 L^2 - \cdots - \varphi_p L^p\right) \left(u_{it}^* - \widehat{\beta}^* \varepsilon_{it}^* \right) \right\}^2 + o_p(1) \\ &= \tfrac{1}{NT} \sum_{i=1}^{N} \sum_{t=2}^{T} d^2 \left(L\right) \left(u_{it}^{*2} - 2\widehat{\beta}^* u_{it}^* \varepsilon_{it}^* + \widehat{\beta}^{*2} \varepsilon_{it}^{*2} \right) + o_p(1) \\ &\xrightarrow{p} d^2 \left(1\right) \sigma_v^2 \equiv \sigma^2 . \end{aligned}$$

We are now in a position to derive the asymptotic distribution of

$$\frac{\sqrt{N} \xi_{7NT}^{'}}{s_v \sqrt{\xi_{8NT}}} .$$

Define

$$\mu_{7T} \equiv E\left(\zeta_{7iT}^{'} \right) = E\left(\zeta_{7i}^{'} \right) + o_p\left(1\right)$$

and

$$\mu_{8T} \equiv E\left(\zeta_{8iT} \right) = E\left(\zeta_{8i} \right) + o_p\left(1\right) .$$

Then, we make the following normalization:

$$\begin{aligned} t_{ADF} - \tfrac{\sqrt{N} \mu_{7T}}{s_v \sqrt{\mu_{8T}}} &= \tfrac{\sqrt{N} \xi_{7NT}^{'}}{s_v \sqrt{\xi_{8NT}}} - \tfrac{\sqrt{N} \mu_{7T}}{s_v \sqrt{\mu_{8T}}} \\ &= \tfrac{\sqrt{N} \left(\xi_{7NT} - \mu_{7T} \right)}{s_v \sqrt{\xi_{8NT}}} + \tfrac{\mu_{7T}}{s_v} \sqrt{N} \left(\sqrt{\tfrac{1}{\zeta_{\xi NT}}} - \sqrt{\tfrac{1}{\mu_{8T}}} \right) \end{aligned}$$

Since

$$\frac{\sqrt{N} \left(\xi_{7NT} - \mu_{7T} \right)}{s_v \sqrt{\xi_{8NT}}} \Rightarrow N\left(0, \frac{d^2\left(1\right) \frac{\sigma_{0v}^4}{12}}{\sigma^2 \frac{\sigma_{0v}^2}{6}} \right) = N\left(0, \frac{d^2\left(1\right) \sigma_{0v}^2}{2\sigma^2} \right)$$

$$\frac{\mu_{7T}}{s_v} \sqrt{N} \left(\sqrt{\frac{1}{\xi_{8NT}}} - \sqrt{\frac{1}{\mu_{8T}}} \right) \Rightarrow N\left(0, \frac{1}{4} \frac{\left(-d\left(1\right) \frac{\sigma_v^2}{2} \right)^2 \frac{\sigma_{0v}^4}{45}}{\sigma^2 \left(\frac{\sigma_{0v}^2}{6} \right)^3} \right) = N\left(0, \frac{3d^2\left(1\right) \sigma_v^4}{10\sigma^2 \sigma_{0v}^2} \right)$$

and $\sqrt{N}\left(\xi_{7NT} - \mu_{7T} \right)$ and $\sqrt{N}\left(\xi_{8NT} - \mu_{8T} \right)$ are asymptotically uncorrelated., we conclude that

$$t_{ADF} - \frac{\sqrt{N} \mu_{7T}}{s_v \sqrt{\mu_{8T}}} \Rightarrow N\left(0, \frac{d^2\left(1\right) \sigma_{0v}^2}{2d^2\left(1\right) \sigma_v^2} + \frac{3d^2\left(1\right) \sigma_v^4}{10\left(d^2\left(1\right) \sigma_v^2 \right) \sigma_{0v}^2} \right)$$

$$t_{ADF} - \frac{\sqrt{N} \mu_{7T}}{s_v \sqrt{\mu_{8T}}} \Rightarrow N\left(0, \frac{\sigma_{0v}^2}{2\sigma_v^2} + \frac{3\sigma_v^2}{10\sigma_{0v}^2} \right)$$

as $N \to \infty$ and $T \to \infty$. ■

References

[1] Banerjee, A., Dolado, J., Galbraith, J. W., and Hendry, D. F. (1993), *Co-integration, Error Correction, and the Econometric Analysis of Non-stationary Data*, Oxford University Press, New York.

[2] Billingsley, P. (1986), *Probability and Measure*, John Wiley, New York.

[3] Breitung, J., and Meyer, W. (1994), "Testing for Unit Roots in Panel Data: Are Wages on Different Bargaining Levels Cointegrated?" *Applied Economics*, 26, 353-361.

[4] Granger, C. W. J., and Newbold, P. (1974), "Spurious Regressions in Econometrics," *Journal of Econometrics*, 2, 111-120.

[5] Haug, A. A. (1996), "Tests for Cointegration: A Monte Carlo Comparison," *Journal of Econometrics*, 71, 89-115.

[6] Im, K., Pesaran, H., and Shin, Y. (1995), "Testing for Unit Roots in Heterogeneous Panels," Manuscript, University of Cambridge.

[7] Kao, C., and Chen, B. (1995a), "On the Estimation of a Cointegrated Regression in Panel Data when the Cross-Section and Time-Series Dimensions are Comparable," Manuscript, Department of Economics, Syracuse University.

[8] Kao, C., and Chen, B. (1995b), "Residual-Based Tests for Cointegration in Panel Data when the Cross-Section and Time-Series Dimensions are Comparable," Manuscript, Department of Economics, Syracuse University.

[9] Levin, A., and Lin, C.-F. (1992), "Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties," Discussion Paper 92-93, Department of Economics, UC-San Diego.

[10] Levin, A., and Lin, C.-F. (1993), "Unit Root Tests in Panel Data: New Results," Discussion Paper, Department of Economics, UC-San Diego.

[11] Park, J., and Ogaki, M. (1991), "Seemingly Unrelated Canonical Cointegrating Regressions," The Rochester Center for Economic Research, Working Paper No. 280.

[12] Pedroni, P. (1995), "Panel Cointegration: Asymptotic and Finite Sample Properties of Pooled Time Series Tests with an Application to the PPP Hypothesis," Indiana University Working Papers in Economics, No. 95-013.

[13] Pesaran, H., and Smith, R. (1995), "Estimating Long-Run Relationships from Dynamic Heterogenous Panels," forthcoming in *Journal of Econometrics.*

[14] Phillips, P. C. B. (1986), "Understanding Spurious Regressions in Econometrics," *Journal of Econometrics*, 33, 311-340.

[15] Phillips, P. C. B. (1987), "Time Series Regression with a Unit Root," *Econometrica*, 55, 277-302.

[16] Phillips, P. C. B., and Ouliaris, S. (1990), "Asymptotic Properties of Residual Based Test for Cointegration," *Econometrica*, 58, 165-193.

[17] Quah, D. (1994), "Exploiting Cross Section Variation for Unit Root Inference in Dynamic Data," *Economics Letters*, 44, 9-19.

Table 1: LSDV Regression with N = T

N(T)	$\overline{\beta}$	$SSD(\widehat{\beta})$	$ESE(\widehat{\beta})$	$\overline{t_\beta}$	$SSD(t_\beta)$	$P(\lvert t_\beta\rvert > 1.96)$	$\overline{R^2}$	$\overline{DW}$
10	-.0028	.2865	.1495	-.0185	1.9603	.315	.0384	.6184
20	.0024	.2039	.0727	.0339	2.8375	.4889	.02	.3057
30	.0012	.1652	.0479	.0248	3.4693	.5723	.0133	.2033
40	.0006	.1423	.0358	.016	4.0065	.6166	.01	.1519
50	.0001	.1265	.0286	.006	4.4392	.6598	.0079	.1212
60	.0004	.1160	.0237	.0207	4.8987	.6894	.0066	.1009
70	-.0006	.1072	.0203	-.0365	5.3014	.7096	.0057	.0864
80	-.0001	.1002	.0178	-.0057	5.6369	.7264	.0049	.0755

Note:
(a) $\overline{\beta} = \frac{1}{M}\sum_{i=1}^{M}\widehat{\beta}_i$
(b) $\overline{t_\beta} = \frac{1}{M}\sum_{i=1}^{M} t_{\beta_i}$
(c) $\overline{R^2} = \frac{1}{M}\sum_{i=1}^{M} R_i^2$
(d) $\overline{DW} = \frac{1}{M}\sum_{i=1}^{M} DW_i$
(e) The number of replications = 10,000.

Table 2: LSDV Regression with N = 30

T	$\overline{\beta}$	$SSD(\widehat{\beta})$	$ESE(\widehat{\beta})$	$\overline{t_\beta}$	$SSD(t_\beta)$	$P(\lvert t_\beta\rvert > 1.96)$	$\overline{R^2}$	$\overline{DW}$
10	.0006	.1669	.0861	.0104	1.9546	.3130	.0136	.5686
20	.0022	.1660	.0593	.0397	2.8192	.4858	.0134	.2991
30	.0012	.1652	.0479	.0248	3.4693	.5723	.0133	.2033
40	.0004	.1639	.0413	.0131	3.9855	.6218	.0131	.1538
50	-.0001	.1651	.0368	-.0074	4.5078	.6651	.0133	.1238
60	.0013	.1643	.0336	.0381	4.9167	.6889	.0131	.1033
70	-.0012	.1647	.0310	-.0419	5.3318	.7103	.0132	.0890
100	.0026	.1634	.0259	.1072	6.3427	.7548	.0130	.0626
150	.0002	.1647	.0211	.08022	7.8426	.8022	.0132	.0418

Note:
(a) $\overline{\beta} = \frac{1}{M}\sum_{i=1}^{M}\widehat{\beta}_i$
(b) $\overline{t_\beta} = \frac{1}{M}\sum_{i=1}^{M} t_{\beta_i}$
(c) $\overline{R^2} = \frac{1}{M}\sum_{i=1}^{M} R_i^2$
(d) $\overline{DW} = \frac{1}{M}\sum_{i=1}^{M} DW_i$
(e) The number of replications = 10,000.

Table 3: LSDV Regression with T = 30

N	$\overline{\beta}$	$SSD(\widehat{\beta})$	$ESE(\widehat{\beta})$	$\overline{t_\beta}$	$SSD(t_\beta)$	$P(\lvert t_\beta\rvert > 1.96)$	$\overline{R^2}$	$\overline{DW}$
10	.0029	.2848	.0832	.0261	3.5000	.5732	.0378	.2232
20	.0037	.2005	.0588	.0567	3.4577	.5788	.0195	.2080
30	.0012	.1652	.0479	.0248	3.4693	.5723	.0133	.2033
40	.0009	.1425	.0415	.0195	3.4532	.5692	.0099	.2007
50	.0001	.1265	.0371	.0259	3.4306	.5699	.0109	.1993
60	.0017	.1163	.0339	.0444	3.4478	.5635	.0067	.1983
70	.0020	.1080	.0313	.0588	3.4613	.5745	.0058	.1976

Note:

(a) $\overline{\beta} = \frac{1}{M}\sum_{i=1}^{M}\widehat{\beta}_i$

(b) $\overline{t_\beta} = \frac{1}{M}\sum_{i=1}^{M} t_{\beta_i}$

(c) $\overline{R^2} = \frac{1}{M}\sum_{i=1}^{M} R_i^2$

(d) $\overline{DW} = \frac{1}{M}\sum_{i=1}^{M} DW_i$

(e) The number of replications = 10,000.

Table 4: The Empirical Size at 5 Percent for $a_1 = 1$, $\sigma = 1$, $\rho = 1$, $\pi = 0.0$, and $\theta = 0$

			$T = 10$					$T = 25$		
N	DF_ρ	DF_t	DF_ρ^*	DF_t^*	ADF	DF_ρ	DF_t	DF_ρ^*	DF_t^*	ADF
1	.350	.282	.034	.075	.139	.435	.239	.157	.098	.137
2	.257	.233	.028	.079	.138	.307	.180	.130	.092	.125
5	.182	.191	.021	.084	.167	.195	.139	.088	.089	.115
10	.149	.179	.018	.096	.210	.142	.114	.068	.079	.113
15	.142	.179	.021	.099	.272	.135	.118	.059	.080	.123
20	.144	.193	.023	.115	.315	.121	.109	.054	.078	.123
25	.149	.212	.024	.121	.367	.115	.110	.052	.078	.129
50	.178	.270	.041	.192	.575	.119	.122	.053	.089	.161
75	.208	.320	.068	.255	.722	.117	.126	.054	.095	.185
100	.242	.372	.098	.314	.829	.121	.139	.061	.106	.214
150	.318	.472	.173	.433	.939	.132	.154	.070	.123	.265
200	.378	.551	.269	.535	.979	.142	.170	.079	.138	.308
250	.446	.623	.372	.630	.993	.153	.187	.091	.156	.356
300	.513	.682	.473	.708	.998	.171	.212	.105	.177	.407
			$T = 50$					$T = 100$		
1	.461	.218	.354	.145	.167	.467	.213	.429	.179	.184
2	.322	.159	.249	.115	.135	.319	.152	.291	.133	.139
5	.202	.118	.153	.093	.107	.209	.115	.187	.103	.113
10	.149	.101	.106	.079	.092	.150	.092	.129	.082	.093
15	.127	.091	.087	.075	.090	.127	.083	.105	.078	.085
20	.121	.093	.084	.076	.093	.116	.084	.098	.077	.087
25	.113	.091	.077	.074	.094	.108	.078	.090	.071	.084
50	.105	.090	.072	.075	.098	.095	.076	.079	.071	.081
75	.093	.085	.065	.075	.107	.087	.073	.073	.067	.079
100	.092	.093	.064	.077	.114	.085	.070	.071	.062	.081
150	.094	.094	.069	.084	.127	.084	.075	.069	.067	.085
200	.098	.098	.069	.086	.141	.085	.076	.068	.068	.091
250	.098	.104	.072	.091	.149	.079	.078	.069	.071	.096
300	.104	.113	.075	.096	.166	.079	.076	.065	.069	.095

Note:

(a) The number of replications =10,000.

(b) All rejection frequencies are based on the asymptotic critical value 1.645.

Table 5: The Power at 5 Percent for $a_1 = 1$, $\sigma = 1$, $\rho = .90$, $\pi = 0$, and $\theta = 0$

	$T = 10$					$T = 25$				
N	DF_ρ	DF_t	DF_ρ^*	DF_t^*	ADF	DF_ρ	DF_t	DF_ρ^*	DF_t^*	ADF
1	.359	.283	.028	.068	.137	.496	.255	.135	.084	.139
2	.296	.252	.022	.073	.136	.451	.234	.154	.091	.145
5	.270	.249	.017	.077	.184	.465	.261	.185	.115	.176
10	.286	.277	.014	.090	.248	.569	.341	.255	.163	.225
15	.329	.317	.017	.100	.323	.674	.421	.334	.205	.276
20	.375	.358	.018	.112	.373	.756	.488	.406	.248	.321
25	.419	.405	.019	.121	.435	.823	.559	.484	.291	.368
50	.626	.581	.027	.195	.665	.966	.799	.774	.507	.566
75	.773	.714	.040	.264	.811	.994	.914	.913	.674	.715
100	.870	.809	.054	.325	.899	1.00	.967	.972	.793	.822
150	.958	.919	.093	.464	.974	1.00	.996	.998	.922	.934
200	.988	.964	.143	.583	.993	1.00	1.00	1.00	.977	.977
250	.998	.988	.195	.679	.998	1.00	1.00	1.00	.995	.992
300	.999	.995	.266	.758	.000	1.00	1.00	1.00	.998	.997
	$T = 50$					$T = 100$				
1	.622	.293	.458	.179	.197	.845	.471	.781	.391	.355
2	.569	.341	.255	.163	.225	.939	.667	.898	.601	.518
5	.798	.471	.657	.340	.324	.999	.949	.994	.922	.843
10	.936	.698	.853	.561	.496	1.00	1.00	1.00	.998	.985
15	.983	.841	.944	.724	.630	1.00	1.00	1.00	1.00	.998
20	.997	.922	.982	.835	.735	1.00	1.00	1.00	1.00	1.00
25	.999	.995	.995	.905	.818	1.00	1.00	1.00	1.00	1.00
50	1.00	1.00	1.00	.996	.977	1.00	1.00	1.00	1.00	1.00
75	1.00	1.00	1.00	1.00	.997	1.00	1.00	1.00	1.00	1.00
100	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
150	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
200	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
250	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
300	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Note:
(a) The number of replications =10,000.
(b) All rejection frequencies are based on the asymptotic critical value 1.645.

Table 6: The Empirical Size of 5 Percent of Tests with the Null Hypothesis of No Cointegration

	$\sigma = .25$			$\sigma = 1$			$\sigma = 4$		
θ	-0.5	0	0.5	-0.5	0	0.5	-0.5	0	0.5
$\pi = -0.8$									
DF_ρ	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DF_t	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DF_ρ^*	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DF_t^*	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
ADF	.814	.038	.062	1.00	.985	.998	.993	.672	.466
$\pi = 0$									
DF_ρ	.100	.102	.102	.104	.105	.104	.104	.104	.101
DF_t	.089	.089	.089	.087	.090	.088	.089	.088	.088
DF_ρ^*	.069	.069	.069	.073	.072	.073	.073	.070	.075
DF_t^*	.072	.072	.072	.077	.090	.088	.074	.076	.073
ADF	.097	.096	.096	.099	.098	.100	.099	.099	.100
$\pi = 0.8$									
DF_ρ	.005	.000	.000	.145	.020	.067	.287	.086	.306
DF_t	.014	.000	.000	.113	.028	.064	.206	.078	.215
DF_ρ^*	.287	.095	.041	.208	.163	.273	.101	.076	.089
DF_t^*	.161	.070	.048	.149	.116	.174	.098	.078	.088
ADF	.519	.615	.416	.157	.317	.269	.049	.113	.044

Note:

(a) The number of replications = 10,000.

(b) N = 50 and T = 50.

(c) All rejection frequencies are based on the asymptotic critical value 1.645.

(d) $a_1 = 1$.

Table 7: The Power of 5 Percent of Tests with the Null Hypothesis of No Cointegration and the Alternative Hypothesis of Cointegration ($\rho = 0.90$)

	$\sigma = .25$			$\sigma = 1$			$\sigma = 4$		
θ	-0.5	0	0.5	-0.5	0	0.5	-0.5	0	0.5
$\pi = -0.8$									
DF_ρ	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DF_t	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DF_ρ^*	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DF_t^*	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
ADF	.198	.021	.012	.882	.781	.877	.986	.991	.999
$\pi = 0$									
DF_ρ	.100	.487	.215	.998	1.00	1.00	1.00	1.00	1.00
DF_t	.044	..298	.165	.907	1.00	1.00	1.00	1.00	1.00
DF_ρ^*	.127	.424	.293	.997	1.00	1.00	1.00	1.00	1.00
DF_t^*	.054	.251	.206	.812	.996	1.00	.999	1.00	1.00
ADF	.095	.259	.224	.813	.977	.997	.988	.998	.998
$\pi = 0.8$									
DF_ρ	.665	.369	.111	1.00	1.00	1.00	1.00	1.00	1.00
DF_t	.324	.186	.076	1.00	1.00	1.00	1.00	1.00	1.00
DF_ρ^*	.998	.995	.973	1.00	1.00	1.00	1.00	1.00	1.00
DF_t^*	.888	.855	.738	1.00	1.00	1.00	1.00	1.00	1.00
ADF	.844	.993	.996	.987	1.00	1.00	.982	.999	.988

Note:
(a) The number of replications = 10,000.
(b) N = 50 and T = 50.
(c) All rejection frequencies are based on asymptotic critical value: 1.645.
(d) $a_1 = 1$.

Table 8: The Size-adjusted Power of 5 Percent of Tests with the Null Hypothesis of No Cointegration and the Alternative Hypothesis of Cointegration ($\rho = 0.85$)

	$\sigma = .25$			$\sigma = 1$			$\sigma = 4$		
θ	-0.5	0	0.5	-0.5	0	0.5	-0.5	0	0.5
$\pi = -0.8$									
DF_ρ	.001	.027	.012	.000	.000	.000	.000	.001	.010
DF_t	.002	.026	.011	.000	.000	.000	.000	.000	.003
DF_ρ^*	.000	.012	.004	.000	.000	.000	.000	.000	.000
DF_t^*	.000	.014	.106	.000	.000	.000	.000	.000	.000
ADF	.000	.026	.004	.000	.000	.000	.073	.925	1.00
$\pi = 0$									
DF_ρ	.297	.732	.508	1.00	1.00	1.00	1.00	1.00	1.00
DF_t	.128	.446	.334	1.00	1.00	1.00	1.00	1.00	1.00
DF_ρ^*	.388	.751	.728	1.00	1.00	1.00	1.00	1.00	1.00
DF_t^*	.128	.458	.464	1.00	1.00	1.00	1.00	1.00	1.00
ADF	.128	.299	.336	.991	.999	1.00	1.00	1.00	1.00
$\pi = 0.8$									
DF_ρ	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DF_t	.998	.997	.993	1.00	1.00	1.00	1.00	1.00	1.00
DF_ρ^*	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DF_t^*	.990	.995	.992	1.00	1.00	1.00	1.00	1.00	1.00
ADF	.582	.974	.996	1.00	1.00	1.00	1.00	1.00	.988

Note:

(a) The number of replications = 10,000.

(b) N = 50 and T = 50.

(c) All rejection frequencies are based on the 5 percent percentile of the empirical distribution.

(d) $a_1 = 1$.