Far East J. Theo. Stat. **11**(2) (2003), 113-131

UNIT ROOT AND COINTEGRATION TESTS UNDER HETEROSKEDASTICITY

AKIRA TOKIHISA and SHIGEYUKI HAMORI

Graduate School of Economics, Kobe University
2-1, Rokkodai, Nada-Ku, Kobe, 657, Japan
e-mail: hamori@kobe-u.ac.jp

Abstract

This paper analyzes the influence of the shift of the error variance on the unit root test, the cointegration test. The main findings can be summarized as follows: (1) Heteroskedasticity affects the size of the unit root test; (2) The unit root test based on GLS is affected by the nuisance parameter and thus cannot solve the problem. Data transformation is a good procedure to solve this problem. Monte Carlo experiments support this assertion; (3) The ideas can be applied to the residual-based cointegration test. Monte Carlo experiments show that the null hypothesis of no-cointegration tends to be over-rejected. The size-distortion tends to be improved when data transformation is carried out.

www.pphmj.com

1. Introduction

The unit root test created by Dickey and Fuller [3] and the cointegration test proposed by Engle and Granger [5] have had a great influence on today's econometric analysis. As a result, performing the unit root test and the cointegration test are becoming general practices in

2000 Mathematics Subject Classification: 62M10, 62E20.

Key words and phrases: unit root, cointegration, GLS.

The authors are grateful to Colin McKenzie and the participants of econometric workshop held at Kobe University for useful comments and suggestions.

Received July 17, 2003

© 2003 Pushpa Publishing House

current econometric analyses that use time series data. McKenzie [13] examined the impact of the theoretical literature relating to unit roots and cointegration on empirical analyses, and found the importance of these tests in the empirical literature.

If variables are determined to have unit roots, standard asymptotic theory used for regression analysis cannot be applied, and there are various problems associated with statistical inference. To address such problems, a variety of estimation techniques and testing methods have been developed. (See, for examples, Dickey and Fuller [4], Said and Dickey [15], Phillips and Perron [14], Cochran [1], Hylleberg et al. [10], and Kwiatkowski et al. [12].)

The difference method is often employed when economic variables are non-stationary (have unit roots). However, some variables may not be mutually independent, even though they are non-stationary. In such cases, specification-error may arise in models based on differences. This relationship between non-stationary variables is defined as cointegration, or as a cointegrating relationship. From a statistical standpoint, when a linear combination of non-stationary variables is stationary, then the relation is called *cointegrating* (Engle and Granger [5]). Since the cointegrating relationship implies a long-run equilibrium among economic variables, the economic theory under investigation (e.g., the stability of the money demand function) is often tested based on the cointegration relationship (see, for examples, Friedman and Kuttner [7], Feldstein and Stock [6]).

The present study analyzes the influence of the heteroskedasticity on the unit root test and the cointegration test. Conventional wisdom regarding the unit root test claims that while serial correlation is important, heteroskedasticity in the variance structure is not important. For example, Phillips and Perron [14] derive the mixing condition to satisfy the functional central limit theorem. They show that a z-transformation of a t-type test statistic is invariant to some form of heteroskedasticity. Davidson [2], however, shows that there is nonstandard Brownian motion in the presence of general heteroskedasticity. The following study reanalyzes this problem by taking into consideration a simple form of heteroskedasticity.

www.pphmj.com

The purpose of this article is twofold. The first is to examine the influence of the shift of the error variance on the unit root test (Dickey and Fuller [3]) and the residual-based cointegration test (Engle and Granger [5]). The second is to present a test method when the variance shift exists. The following two conclusions will be confirmed. First, if the variance shift exists, a serious size-distortion occurs in each of the tests. Second, the size-distortion can be corrected by performing an appropriate transformation of the data.

2. Unit Root Test under Heteroskedasticity

Consider the following data generating process (DGP):

$$\Delta y_t = \varepsilon_t + \eta_t DU_t, \quad t = 1, 2, \ldots, T, \tag{1}$$

where

$$DU_t = \begin{cases} 1, & t > T_B, \ (1 < T_B < T), \\ 0, & \text{otherwise}, \end{cases} \tag{2}$$

$$\varepsilon_t \sim iid(0, \sigma_1^2), \quad \eta_t \sim iid(0, \sigma_2^2), \tag{3}$$

and $\{\varepsilon_t\}$ and $\{\eta_t\}$ are independent of each other, and both of them have finite fourth moments. We estimate the model using the following regression equation:

$$y_t = \hat{\rho} y_{t-1} + \hat{e}_t, \tag{4}$$

where $\hat{\rho} = \frac{\sum_{t=1}^{T} y_{t-1} y_t}{\sum_{t=1}^{T} y_{t-1}^2}$, and $\hat{e}_t$ is the residual from the regression. Based on equation (4), the test statistics is given by the t-value of $\hat{\rho}$ as follows:

$$t_{\hat{\rho}} = \frac{\hat{\rho} - 1}{\hat{\sigma}_{\hat{\rho}}}, \tag{5}$$

where $\hat{\sigma}_{\hat{\rho}} = \sqrt{\frac{s_T^2}{\sum_{t=1}^{T} y_{t-1}^2}}$ and $s_T^2 = \frac{\sum_{t=1}^{T} (y_t - \hat{\rho} y_{t-1})^2}{T - 1}$.

www.pphmj.com

Let us derive the limiting distributions of the test statistics, (5). First, from the functional central limit theorem, we have

$$\frac{\sum_{t=1}^{[rT]} \varepsilon_t}{\sigma_1 \sqrt{T}} \xrightarrow{L} B_1(r), \tag{6}$$

and

$$\frac{\sum_{t=1}^{[rT]} \eta_t}{\sigma_2 \sqrt{T}} \xrightarrow{L} B_2(r), \tag{7}$$

where $B_1(\bullet)$ and $B_2(\bullet)$ are independent standard Brownian Motions. Therefore, the following equation is derived from equations, (6) and (7),

$$\frac{\sum_{t=1}^{[rT]} (\varepsilon_t + \eta_t DU_t)}{\sqrt{T}} \xrightarrow{L} \sigma_1 B_1(r) + \sigma_2 [B_2(r) - B_2(\lambda)] du(r) \equiv G(r), \tag{8}$$

where

$$du(r) = \begin{cases} 1, & r > \lambda \equiv T_B / T \\ 0, & \text{otherwise.} \end{cases} \tag{9}$$

It also holds that

$$\frac{\sum_{t=1}^{T} (\varepsilon_t + \eta_t DU_t)^2}{T} \xrightarrow{p} \sigma_1^2 + (1 - \lambda)\sigma_2^2 \equiv \sigma^2. \tag{10}$$

Let $V(r)$ be defined as $V(r) \equiv \sigma^{-1} G(r)$. Then, $\{V(r)\}_{r=0}^{1}$ is a normal random variable with mean zero and covariance $E[V(r)V(s)] = \min\{C(r),$ $C(s)\}$, where $C(r) = \left(\frac{\sigma_1}{\sigma}\right)^2 r + \left(\frac{\sigma_2}{\sigma}\right)^2 (r - \lambda) du(r)$. Since $C(\bullet)$ is continuous on $(0, \lambda)$ and $(\lambda, 1)$ and monotonically increasing on $(0, 1)$, and mapped from $[0, 1]$ to $[0, 1]$, $V(r)$ can be expressed as $V(r) = B(C(r))$, where $\{B(r)\}_{r=0}^{1}$ is a standard Brownian Motion. Here, $\{V(r)\}_{r=0}^{1}$ is called as the variance-transformed Brownian Motion (Davidson [2]).

Note that any increment of $\theta(r)B(r)$ is not independent, where $\theta : [0, 1] \rightarrow \Re$ is any continuous deterministic function (Davidson [2]). On

www.pphmj.com

the other hand, any increment of $\frac{\sum_{t=1}^{[rT]}(\varepsilon_t + \eta_t DU_t)}{\sqrt{T}}$ is asymptotically independent and any increment with respect to r of $B(C(r))$ is independent. Thus, the asymptotic distribution of $\frac{\sum_{t=1}^{[rT]}(\varepsilon_t + \eta_t DU_t)}{\sigma\sqrt{T}}$ is not equal to $r^{-1/2}C^{1/2}(r)B(r)$. Note that the explanation of the limiting distribution in Hamori and Tokihisa [9] is not appropriate.

For the test statistic, it holds that

$$t_{\hat{\rho}-1} \xrightarrow{L} \frac{\frac{1}{2}[\{V(1)\}^2 - 1]}{\sqrt{\int_0^1 \{V(r)\}^2 dr}}. \tag{11}$$

It is clear that $V(1) = B(1)$ since $C(1) = 1$. To rewrite $\int_0^1 \{V(r)\}^2 dr$ in equation (11), we use the change of variable technique. Since $C(\bullet)$ is differentiable on $(0, \lambda)$ and $(\lambda, 1)$, the inverse function of C denoted D and $D'(r) = dD/dr$ exist on $r \in (0, \lambda)$ and $r \in (\lambda, 1)$. Thus

$$\int_0^1 \{V(r)\}^2 dr = \int_0^1 \{B(C(r))\}^2 dr = \int_0^\lambda \{B(r)\}^2 D'(r)dr + \int_\lambda^1 \{B(r)\}^2 D'(r)dr \tag{12}$$

and

$$D'(r) = \begin{cases} 1 + (1-\lambda)\delta^2 & \text{for } r \in (0, \lambda) \\ \dfrac{1 + (1-\lambda)\delta^2}{1+\delta^2} & \text{for } r \in (\lambda, 1) \end{cases}, \tag{13}$$

where $\delta = \frac{\sigma_2}{\sigma_1}$. Using the equation $\int_0^\lambda \{B(r)\}^2 dr + \int_\lambda^1 \{B(r)\}^2 dr = \int_0^1 \{B(r)\}^2 dr$ and (13), the right hand side of (12) is equal to

$$\int_0^1 \{B(r)\}^2 dr + (1-\lambda)\delta^2 \int_0^\lambda \{B(r)\}^2 dr - \lambda \frac{\delta^2}{1+\delta^2} \int_\lambda^1 \{B(r)\}^2 dr. \tag{14}$$

Thus, when an upward shift in variance occurs, the test statistic is

www.pphmj.com

affected by the nuisance parameters (δ and λ), which causes the size-distortion. However, judging from the limiting distributions given by (11), it is not clear whether the upward shift in variance causes over-rejection or under-rejection of the unit root hypothesis. The case of downward shift in variance is similarly analyzed.

To evaluate the effect of heteroskedasticity on the size of the unit root test, we perform Monte Carlo experiments with 5,000 replications in each experiment where the disturbance term is generated from normal random variables. Each experiment is performed under the following settings:

$$\text{DGP}: \Delta y_t = \varepsilon_t, \; t = 1, 2, \ldots, T$$

$$E[\varepsilon_t] = 0, \; Var[\varepsilon_t^2] = \begin{cases} \sigma_1^2 & \text{for } t \le \lambda T, \; \lambda = TB/T, \\ \sigma_2^2 & \text{for } t > \lambda T. \end{cases}$$

$$T = 100, \; 250,$$

$$(\sigma_1^2, \sigma_2^2) = (1, 2), (1, 5), (1, 10), (2, 1), (5, 1), (10, 1),$$

$$\lambda = 0.3, 0.4, 0.5, 0.6, 0.7.$$

Table 1 shows the empirical test size of the t-type test of the unit root based on the raw data series. As is clear from Table 1, the size-distortion becomes larger as the shift in variance becomes larger. In Table 1, for example, when $T = 100$ and $\lambda = 0.7$ at a nominal size of 5% and the combination of variance changes from $(1.0, 2.0)$ to $(1.0, 5.0)$ and $(1.0, 10.0)$, then the empirical test size also increases, going from 6.08 to 14.94 and 21.72. In Table 1, when $T = 100$ and $\lambda = 0.3$ at a nominal size of 5% and the combination of variance changes from $(2.0, 1.0)$ to $(5.0, 1.0)$ and $(10.0, 1.0)$, then the empirical test size also increases, going from 6.06 to 9.56 and 11.60. Similar patterns are observed in other cases. It is interesting to note that the size-distortion tends to be larger for a large value of λ when an upward shift in variance occurs. On the other hand, the size-distortion tends to be larger for a small value of λ when a downward-shift in variance occurs. As is clear from the table, the

www.pphmj.com

size-distortion is more serious for the upward shift in variance than it is for the downward shift in variance.

3. GLS Estimation and Unit Root Test

One may consider the generalized least square (GLS) estimation as one method of dealing with heteroskedasticity. However, if the explanatory variables are non-stationary random variables, the result is not necessarily the same as in the case with stationary explanatory variables. When considering the variance shift at a certain point of time, the t-type unit root test that uses the GLS estimation is influenced by the nuisance parameter.

For the process, given by

$$\Delta y_t = \varepsilon_t, \; t = 1, 2, \ldots, T, \; E[\varepsilon_t^2] = \sigma_1^2 \text{ for } t \le TB, \; 1 < TB < T,$$

$$= \sigma_2^2 \text{ for } t > TB, \tag{15}$$

where $\{\varepsilon_t\}$ is a random variable with mean 0 and independent, we can apply the following GLS regression

$$y_t = \hat{\rho}_{gls} y_{t-1} + \hat{\varepsilon}_t, \; t = 1, \ldots, T. \tag{16}$$

Then, the t-type test statistic given by

$$t_{\hat{\rho}_{gls}-1} = \frac{\sum_{t=1}^{TB} y_{t-1}\varepsilon_t/\sigma_1^2 + \sum_{t=TB+1}^{T} y_{t-1}\varepsilon_t/\sigma_2^2}{\hat{\sigma}\sqrt{\sum_{t=1}^{TB} y_{t-1}^2/\sigma_1^2 + \sum_{t=TB+1}^{T} y_{t-1}^2/\sigma_2^2}}, \tag{17}$$

where

$$\hat{\sigma} = \sqrt{\sigma_1^{-2} T^{-1} \sum_{t=1}^{TB} (y_t - \hat{\rho}_{gls} y_{t-1})^2 + \sigma_2^{-2} T^{-1} \sum_{t=TB+1}^{T} (y_t - \hat{\rho}_{gls} y_{t-1})^2}$$

has the following asymptotic distribution:

www.pphmj.com

$$\frac{\int_0^1 BdB + (\delta - 1)B(\lambda)[B(1) - B(\lambda)]}{\sqrt{\int_0^1 B^2 dr + 2(\delta - 1)B(\lambda)\int_\lambda^1 B(r)dr + (\delta - 1)^2(1 - \lambda)B(\lambda)^2}}, \tag{18}$$

where $\frac{TB}{T} = \lambda$ and $\delta = \frac{\sigma_1}{\sigma_2}$ [see Appendix].

When the variance shift does not occur in the estimation period $[\delta = 1, \lambda = 0, 1]$, each of the second terms of the numerator and the denominator become zero and thus the distribution (16), is equal to the τ-distribution of Dickey and Fuller [3]. Generally speaking, however, the distribution is influenced by the parameter δ. Thus, it is not advisable to apply the GLS regression to perform the unit root test in the presence of the variance shift.

4. Alternative Approach

We present a unit root test that does not depend on the variance of the error term. If series $\{y_t\}$ is a random walk process, its differenced series can be considered to be a random error with mean 0. If the error has a finite variance, the series obtained by the following transformation is the same as the random walk process generated from the random variable with a variance of unity:

www.pphmj.com

$$y_t^* = y_{t-1}^* + \frac{\Delta y_t}{\sigma_t}, \quad t = 1, 2, \ldots, T, \tag{19}$$

where

$$\sigma_t = \sqrt{\mathrm{var}(\Delta y_t)}.$$

Following Dickey and Fuller [3], we regard y_0 as fixed and $y_0^* = y_0$. It is necessary to replace the unknown parameter σ_t by its consistent estimator to perform a feasible transformation of the data. For instance, if the variance shifts only at a certain point in time and that point in time is already known, then the estimator of the variance from OLS regression for each period before and after the change can be used.

Let $\{\varepsilon_t\}$ be an independent process with mean 0 and variance σ_t^2, and let $\{\nu_t\}$ be an independent process with mean 0 and variance 1. If the data generating process (DGP) is given by:

$$\Delta y_t = \varepsilon_t, \tag{20}$$

then we can transform the data as follows:

$$\Delta y_t^* = \nu_t. \tag{21}$$

Thus, the t-type test can be applied for the following regression:

$$y_t^* = \rho y_{t-1}^* + \nu_t. \tag{22}$$

Thus, the null hypothesis and its alternative hypothesis are given as follows:

$$H_0 : \rho = 1 \tag{23}$$

and

$$H_A : \rho < 1. \tag{24}$$

To evaluate the effectiveness of this procedure, we performed Monte Carlo experiments with 5,000 replications in each experiment where the disturbance term was generated from normal random variables. Each experiment are performed under the following settings:

$$\text{DGP} : \Delta y_t = \varepsilon_t,\ t = 1, 2, \ldots, T,$$

$$E[\varepsilon_t] = 0,\ \ Var[\varepsilon_t^2] = \begin{cases} \sigma_1^2 & \text{for } t \le \lambda T,\ \lambda = TB/T, \\ \sigma_2^2 & \text{for } t > \lambda T. \end{cases}$$

$$T = 100,\ 250,$$

$$(\sigma_1^2, \sigma_2^2) = (1, 2), (1, 5), (1, 10), (2, 1), (5, 1), (10, 1),$$

$$\lambda = 0.3, 0.4, 0.5, 0.6, 0.7.$$

Table 2 shows the empirical test size of the t-type test of a unit root using the transformed data. As is clear from Table 2, the empirical test size is improved when the data is transformed. For example, take the case where $T = 100$ and $\lambda = 0.7$ at a nominal size of 5% and the combination

www.pphmj.com

of variance changes from $(1.0, 2.0)$ to $(1.0, 5.0)$ and $(1.0, 10.0)$. Then the corresponding empirical test size is equal to 6.08, 14.94 and 21.72 in Table 1, whereas the corresponding empirical test size is equal to 5.80, 4.86, and 5.48 in Table 2. Similarly, take the case where $T = 100$ and $\lambda = 0.3$ at a nominal size of 5% and the combination of variance changes from $(2.0, 1.0)$ to $(5.0, 1.0)$ and $(10.0, 1.0)$. Then, the corresponding empirical test size is equal to 6.06, 9.56, and 11.60 in Table 1, whereas the corresponding empirical test size is equal to 5.12, 5.24, and 5.42 in Table 2. Similar patterns are observed in other cases.

5. Cointegration Test

This section analyzes the effect of heteroskedasticity on the cointegration test developed by Engle and Granger [5]. The data transformation presented in Section 4 can also be applied to the test of a multivariate cointegration process. However, if cross correlation among error terms exists, a distortion results in the test for the transformation of individual variables like (19). Therefore, a more applicable method is used for the transformed data:

$$y_t^* = y_{t-1}^* + C_t^{-1} \Delta y_t, \quad t = 1, \ldots, T \tag{25}$$

$$y_0^* = y_0 \tag{26}$$

where $\Delta y_t = \varepsilon_t \in \Re^n$, $\{\varepsilon_t\}$ is an independent process with mean 0 and finite nonsingular variance-covariance matrix, and C_t denotes the Cholesky factor of the variance-covariance matrix of ε_t at each point t. $\{C_t^{-1} \Delta y_t\}$ is a process that is independent across time and has mean zero and an identity variance-covariance matrix. The residual-based cointegration test presented by Engle and Granger [5] can be applied to the data transformed as in expressions (25) and (26).

To evaluate the effectiveness of this procedure, we performed Monte Carlo experiments using 5,000 replications in each experiment, where the disturbance term was generated from normal random variables. Each experiment is performed under the following settings:

www.pphmj.com

$$\text{DGP}: \begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{bmatrix} = \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix},$$

$$E(\varepsilon_{1t}^2) = \begin{cases} \sigma_{11}^2 & \text{for } t \le \lambda T, \\ \sigma_{12}^2 & \text{for } t > \lambda T. \end{cases} \quad E(\varepsilon_{2t}^2) = \begin{cases} \sigma_{21}^2 & \text{for } t \le \lambda T, \\ \sigma_{22}^2 & \text{for } t > \lambda T. \end{cases}$$

$$E(\varepsilon_{1t}, \varepsilon_{2t}) = 0,$$

$$T = 100, 250,$$

$$(\sigma_{11}^2, \sigma_{12}^2) = (1, 5), (5, 1), \quad (\sigma_{21}^2, \sigma_{22}^2) = (1, 5), (5, 1),$$

$$\lambda = 0.3, 0.4, 0.5, 0.6, 0.7.$$

Regression: $y_{1t} = \beta y_{2t} + z_t$

$$z_t = \rho z_{t-1} + e_t.$$

Tables 9 and 10, respectively, show the empirical size of the test when the DGP is given by (23). Table 9 shows the results using the raw data and Table 10 shows the results using the transformed data. As is clear from Table 9, the null hypothesis of spurious regression in the sense of Granger and Newbold [8] tends to be over-rejected. The size-distortion tends to be improved when data transformation is carried out.

6. Conclusion

This paper analyzes the effects of heteroskedasticity on the unit root test and the cointegration test. The main findings can be summarized as follows:

(1) Heteroskedasticity affects the size of the standard unit root test. The size-distortion tends to be larger for a large value of λ when an upward shift in variance occurs. The size-distortion tends to be larger for a small value of λ when a downward-shift in variance occurs. The size-distortion is more serious for the upward shift in variance than for the downward shift in variance.

(2) The unit root test based on GLS is affected by the nuisance parameter and thus cannot solve the problem. Data transformation is a

www.pphmj.com

good procedure to solve this problem. Monte Carlo experiments support this assertion.

(3) The ideas above can be applied to the residual-based cointegration test. Monte Carlo experiments show that the null hypothesis of spurious regression in the sense of Granger and Newbold [8] tends to be over-rejected. The size-distortion tends to be improved when data transformation is carried out.

Appendix

The Asymptotic Properties of the Unit Root Test based on the GLS

Let Δy_t be as follows:

$$\Delta y_t = \varepsilon_t, \quad E[\varepsilon_t^2] = \sigma_1^2 \quad \text{for} \quad t \le TB, \ (1 < TB < T),$$

$$= \sigma_2^2 \quad \text{for } t > TB,$$

where $\{\varepsilon_t\}$ is a random variable with mean 0 and independent. Then, the asymptotic distribution of the t-test statistic becomes as follows:

$$t_{\hat{\rho}_{gls}-1} = \frac{\sum_{t=1}^{TB} y_{t-1}\varepsilon_t/\sigma_1^2 + \sum_{t=TB+1}^{T} y_{t-1}\varepsilon_t/\sigma_2^2}{\hat{\sigma}\sqrt{\sum_{t=1}^{TB} y_{t-1}^2/\sigma_1^2 + \sum_{t=TB+1}^{T} y_{t-1}^2/\sigma_2^2}}$$

$$\xrightarrow{L} \frac{\int_0^1 BdB + (\delta - 1)B(\lambda)[B(1) - B(\lambda)]}{\sqrt{\int_0^1 B^2 dr + 2(\delta - 1)B(\lambda)\int_\lambda^1 B(r)dr + (\delta - 1)^2(1 - \lambda)B(\lambda)^2}},$$

where

$$\hat{\sigma} = \sqrt{\sigma_1^{-2}T^{-1}\sum_{t=1}^{TB}(y_t - \hat{\rho}_{gls}y_{t-1})^2 + \sigma_2^{-2}T^{-1}\sum_{t=TB+1}^{T}(y_t - \hat{\rho}_{gls}y_{t-1})^2}$$

and

$$\delta = \frac{\sigma_1}{\sigma_2}.$$

(Proof)

Using the random variable $\{\nu_t\}$ which is independent process with mean 0 and variance 1, we can write as follows:

$$\begin{aligned}\varepsilon_t &= \sigma_1 \nu_t \quad \text{for} \quad t \le TB, \\ &= \sigma_2 \nu_t \quad \text{for} \quad t > TB.\end{aligned}$$

Thus, it holds that

$$T^{-1/2} y_{[rT]} = \sigma_1 T^{-1/2} \sum_{t=1}^{\min\{[rT],\, [\lambda T]\}} \nu_t + \sigma_2 T^{-1/2} \sum_{t=[\lambda T]+1}^{[rT]} \nu_t I([rT] > [\lambda T])$$

$$\xrightarrow{L} \sigma_1 B(\min\{r,\, \lambda\}) + \sigma_2 [B(r) - B(\lambda)] I(r > \lambda).$$

Using this relationship, we have

$$T^{-1} \sum_{t=1}^{TB} y_{t-1}\varepsilon_t / \sigma_1^2 + T^{-1} \sum_{t=TB+1}^{T} y_{t-1}\varepsilon_t / \sigma_2^2$$

$$\xrightarrow{L} \int_0^{\lambda} BdB + \frac{\sigma_1 - \sigma_2}{\sigma_2} B(\lambda) \left(\int_{\lambda}^{1} dB \right) + \int_{\lambda}^{1} BdB$$

$$= \int_0^1 BdB + (\delta - 1) B(\lambda) [B(1) - B(\lambda)]$$

and

$$T^{-2} \sum_{t=1}^{TB} y_{t-1}^2 / \sigma_1^2 + T^{-2} \sum_{t=TB+1}^{T} y_{t-1}^2 / \sigma_2^2$$

$$\xrightarrow{L} \int_0^{\lambda} B^2 dr + \int_{\lambda}^{1} \left\{ \frac{\sigma_1 - \sigma_2}{\sigma_2} B(\lambda) + B(r) \right\}^2 dr$$

$$= \int_0^{\lambda} B^2 dr + \int_{\lambda}^{1} B^2 dr + 2 \left(\frac{\sigma_1 - \sigma_2}{\sigma_2} \right) B(\lambda) \int_{\lambda}^{1} Bdr + \left(\frac{\sigma_1 - \sigma_2}{\sigma_2} \right)^2 B(\lambda)^2 \int_{\lambda}^{1} dr$$

$$= \int_0^1 B^2 dr + 2(\delta - 1) B(\lambda) \int_{\lambda}^{1} Bdr + (\delta - 1)^2 B(\lambda)^2 (1 - \lambda).$$

www.pphmj.com

Also, we have:

$$\hat{\sigma}^2 = \sigma_1^{-2} T^{-1} \sum_{t=1}^{TB} (y_t - \hat{\rho}_{gls} y_{t-1})^2 + \sigma_2^{-2} T^{-1} \sum_{t=TB+1}^{T} (y_t - \hat{\rho}_{gls} y_{t-1})^2$$

$$= \sigma_1^{-2} T^{-1} \sum_{t=1}^{TB} \Delta y_t^2 + \sigma_2^{-2} T^{-1} \sum_{t=TB+1}^{T} \Delta y_t^2 + o_p(1) = \sigma_1^{-2} T^{-1} \sum_{t=1}^{TB} \varepsilon_t^2$$

$$+ \sigma_2^{-2} T^{-1} \sum_{t=TB+1}^{T} \varepsilon_t^2 + o_p(1)$$

$$= T^{-1} \sum_{t=1}^{TB} \nu_t^2 + T^{-1} \sum_{t=TB+1}^{T} \nu_t^2 + o_p(1)$$

$$= T^{-1} \sum_{t=1}^{T} \nu_t^2 + o_p(1) \xrightarrow{p} E(\nu_t^2) = 1.$$

Thus, we have the above-mentioned limiting distribution.

www.pphmj.com

Table 1. Empirical size of the unit root test

$$\text{DGP}: \Delta y_t = \varepsilon_t,$$

where

$$E[\varepsilon_t] = 0, \; Var[\varepsilon_t^2] = \begin{cases} \sigma_1^2 & \text{for } t \le \lambda T, \\ \sigma_2^2 & \text{for } t > \lambda T. \end{cases}$$

$$\text{Regression: } y_t = \rho y_{t-1} + \nu_t$$

Size	Sample	σ_1^2	σ_2^2	$\lambda = 0.3$	$\lambda = 0.4$	$\lambda = 0.5$	$\lambda = 0.6$	$\lambda = 0.7$
5%	100	1.0	2.0	6.38	6.30	6.44	6.50	6.08
		1.0	5.0	8.92	9.24	12.42	12.84	14.94
		1.0	10.0	8.78	12.30	14.40	17.16	21.72
		2.0	1.0	6.06	5.98	6.16	6.62	5.78
		5.0	1.0	9.56	8.94	7.60	6.12	6.14
		10.0	1.0	11.60	9.60	7.96	7.12	6.86
	250	1.0	2.0	6.48	6.86	6.52	6.90	6.32
		1.0	5.0	7.78	9.00	11.80	12.24	12.58
		1.0	10.0	9.20	12.18	14.54	17.88	21.94
		2.0	1.0	5.24	6.38	6.16	5.38	5.28
		5.0	1.0	6.38	8.42	7.24	6.90	5.62
		10.0	1.0	11.74	9.78	7.94	7.16	5.66
10%	100	1.0	2.0	12.52	12.22	12.38	11.94	12.72
		1.0	5.0	15.38	16.12	19.84	20.96	22.84
		1.0	10.0	15.66	19.50	21.48	25.28	30.04
		2.0	1.0	11.48	10.90	11.18	12.08	10.88
		5.0	1.0	15.20	14.08	13.18	11.10	12.04
		10.0	1.0	17.46	14.38	13.92	11.98	11.58
	250	1.0	2.0	12.30	12.62	11.78	13.22	12.46
		1.0	5.0	14.82	15.86	18.54	20.24	19.70
		1.0	10.0	15.62	19.04	22.32	26.06	29.62
		2.0	1.0	10.84	11.70	11.36	10.82	10.66
		5.0	1.0	14.82	13.84	12.64	12.34	10.86
		10.0	1.0	16.66	15.54	13.26	12.60	10.86

The number of replication is 5000.

$\{\varepsilon_t\}$ is drawn from a normal distribution.

www.pphmj.com

Table 2. Empirical size of the unit root test with data transformation

$$\text{DGP}: \Delta y_t = \varepsilon_t,\ t = 1, 2, \ldots, T,$$

where

$$E[\varepsilon_t] = 0,\ \ Var[\varepsilon_t^2] = \begin{cases} \sigma_1^2 & \text{for } t \le \lambda T, \\ \sigma_2^2 & \text{for } t > \lambda T. \end{cases}$$

$$\text{Regression: } y_t^* = \rho y_{t-1}^* + \nu_t.$$

Size	Sample	σ_1^2	σ_2^2	$\lambda = 0.3$	$\lambda = 0.4$	$\lambda = 0.5$	$\lambda = 0.6$	$\lambda = 0.7$
5%	100	1.0	2.0	5.52	5.76	5.10	5.42	5.80
		1.0	5.0	5.22	4.74	4.82	5.20	4.86
		1.0	10.0	5.48	5.40	5.56	4.72	5.48
		2.0	1.0	5.12	5.12	4.78	5.12	4.76
		5.0	1.0	5.24	5.00	5.58	4.86	4.18
		10.0	1.0	5.42	5.78	5.16	5.58	5.30
	250	1.0	2.0	4.28	5.14	4.64	4.98	4.82
		1.0	5.0	5.16	4.92	4.68	5.10	4.78
		1.0	10.0	4.94	4.62	4.60	4.72	4.90
		2.0	1.0	4.78	4.86	5.60	4.76	5.24
		5.0	1.0	5.28	4.90	4.98	4.98	4.54
		10.0	1.0	5.18	5.52	4.88	4.72	5.12
10%	100	1.0	2.0	10.76	10.92	10.78	10.88	10.58
		1.0	5.0	10.70	10.22	10.08	11.00	10.10
		1.0	10.0	10.42	11.10	11.48	9.90	10.96
		2.0	1.0	10.44	10.06	10.58	10.24	9.56
		5.0	1.0	10.40	10.38	10.44	10.08	9.14
		10.0	1.0	10.70	11.64	10.96	11.02	10.64
	250	1.0	2.0	9.48	10.64	9.84	10.48	9.80
		1.0	5.0	10.46	10.00	9.28	10.54	9.92
		1.0	10.0	9.72	9.88	10.04	10.22	10.64
		2.0	1.0	10.00	9.72	10.68	10.18	9.94
		5.0	1.0	10.72	9.90	10.28	10.22	10.74
		10.0	1.0	10.48	10.06	10.10	9.70	10.76

The number of replication is 5000.

$\{\varepsilon_t\}$ is drawn from a normal distribution.

www.pphmj.com

Table 3. Empirical size of the cointegration test

$$\text{DGP}: \begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{bmatrix} = \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix},$$

where

$$E(\varepsilon_{1t}^2) = \begin{cases} \sigma_{11}^2 & \text{for } t \leq \lambda T, \\ \sigma_{12}^2 & \text{for } t > \lambda T. \end{cases} \quad E(\varepsilon_{2t}^2) = \begin{cases} \sigma_{21}^2 & \text{for } t \leq \lambda T, \\ \sigma_{22}^2 & \text{for } t > \lambda T. \end{cases}$$

$$E(\varepsilon_{1t}\,\varepsilon_{2t}) = 0,\ E(\varepsilon_{1t}) = E(\varepsilon_{2t}) = 0.$$

$$\text{Regression: } y_{1t} = \beta y_{2t} + z_t$$

$$z_t = \rho z_{t-1} + e_t$$

Size	Sample	σ_1^2		σ_2^2		$\lambda = 0.3$	$\lambda = 0.4$	$\lambda = 0.5$	$\lambda = 0.6$	$\lambda = 0.7$
		σ_{11}^2	σ_{12}^2	σ_{21}^2	σ_{22}^2					
5%	100	1	5	1	5	6.76	7.58	8.18	7.66	7.90
		5	1	5	1	11.12	10.26	10.42	9.32	7.96
		1	5	5	1	5.76	6.64	6.86	6.18	6.46
		5	1	1	5	6.28	5.54	5.68	5.80	5.40
	250	1	5	1	5	7.06	7.70	7.56	7.36	7.76
		5	1	5	1	10.28	10.60	9.20	8.12	7.30
		1	5	5	1	5.98	5.54	5.70	6.30	5.92
		5	1	1	5	5.60	5.54	5.26	5.02	5.64
10%	100	1	5	1	5	13.32	14.44	14.26	13.78	13.58
		5	1	5	1	17.56	18.30	16.50	17.16	15.24
		1	5	5	1	12.20	12.08	11.98	12.44	11.92
		5	1	1	5	10.52	11.08	10.40	10.98	10.84
	250	1	5	1	5	13.52	13.76	13.84	14.42	13.78
		5	1	5	1	17.90	17.48	16.46	15.14	14.78
		1	5	5	1	11.54	12.08	10.60	11.82	12.28
		5	1	1	5	10.24	9.82	10.24	10.32	10.40

The number of replication is 5000.

$\{\varepsilon_{1t}\}$ and $\{\varepsilon_{2t}\}$ are drawn from a normal distribution.

www.pphmj.com

Table 4. Empirical size of the cointegration test with data transformation

$$\text{DGP}: \begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{bmatrix} = \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix},$$

where

$$E(\varepsilon_{1t}^2) = \begin{cases} \sigma_{11}^2 & \text{for } t \le \lambda T, \\ \sigma_{12}^2 & \text{for } t > \lambda T. \end{cases} \quad E(\varepsilon_{2t}^2) = \begin{cases} \sigma_{21}^2 & \text{for } t \le \lambda T, \\ \sigma_{22}^2 & \text{for } t > \lambda T. \end{cases}$$

$$E(\varepsilon_{1t}\,\varepsilon_{2t}) = 0,\ E(\varepsilon_{1t}) = E(\varepsilon_{2t}) = 0$$

$$\text{Regression}:\ y_{1t}^* = \beta y_{2t}^* + z_t$$

$$z_t = \rho z_{t-1} + e_t$$

Size	Sample	y_{1t}		y_{2t}		$\lambda = 0.3$	$\lambda = 0.4$	$\lambda = 0.5$	$\lambda = 0.6$	$\lambda = 0.7$
		σ_{11}^2	σ_{12}^2	σ_{21}^2	σ_{22}^2					
5%	100	1	5	1	5	6.36	6.40	6.20	6.84	6.90
		5	1	5	1	6.60	6.94	6.78	7.08	6.14
		1	5	5	1	7.76	6.92	6.76	6.40	5.98
		5	1	1	5	4.68	4.62	4.34	3.88	3.88
	250	1	5	1	5	5.92	6.02	5.34	5.52	5.72
		5	1	5	1	6.10	6.16	6.68	6.74	6.18
		1	5	5	1	6.50	5.90	6.18	5.60	5.32
		5	1	1	5	4.50	4.68	4.08	4.08	3.44
10%	100	1	5	1	5	12.04	11.74	11.58	12.16	10.92
		5	1	5	1	13.42	14.36	14.30	14.28	12.40
		1	5	5	1	13.72	13.00	12.88	13.16	12.30
		5	1	1	5	9.82	8.98	8.50	8.68	7.86
	250	1	5	1	5	11.76	11.42	11.18	11.00	11.50
		5	1	5	1	11.54	13.04	12.08	12.42	12.18
		1	5	5	1	12.92	12.42	12.64	11.22	10.80
		5	1	1	5	8.86	7.62	7.62	7.24	7.32

The number of replication is 5000.

$\{\varepsilon_{1t}\}$ and $\{\varepsilon_{2t}\}$ are drawn from a normal distribution.

www.pphmj.com

References

[1] J. H. Cochran, How big is the random walk in GNP?, J. Political Economy 96 (1988), 893-920.

[2] J. E. H. Davidson, Stochastic Limit Theory, 2nd ed., Oxford University Press, 1995.

[3] D. A. Dickey and W. A. Fuller, Distribution of the estimators for autoregressive time series with a unit root, J. Amer. Stat. Assoc. 74 (1979), 427-431.

[4] D. A. Dickey and W. A. Fuller, Likelihood ratio statistics for autoregressive time series with a unit root, Econometrica 49 (1981), 1057-1072.

[5] R. F. Engle and C. W. J. Granger, Co-integration and error correction: representation, estimation and testing, Econometrica 55 (1987), 251-276.

[6] M. Feldstein and J. H. Stock, The use of monetary aggregate to target nominal GDP, NBER Working Paper, No. 4304, 1993.

[7] B. M. Friedman and K. N. Kuttner, Money, income, prices, and interest rates, American Economic Review 82 (1992), 472-492.

[8] C. W. J. Granger and P. Newbold, Spurious regressions in econometrics, J. Econometrics 2 (1974), 111-120.

[9] S. Hamori and A. Tokihisa, Testing for a unit root in the presence of a variance shift, Economics Letters 57 (1997), 245-253.

[10] S. Hylleberg, R. F. Engle, C. W. J. Granger and B. S. Yoo, Seasonal integration and cointegration, J. Econometrics 44 (1990), 215-238.

[11] K. Kim and P. Schmidt, Unit root tests with conditional heteroskedasticity, J. Econometrics 59 (1993), 287-300.

[12] D. Kwiatkowski, P. C. B. Phillips, P. Schmidt and Y. Shin, Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?, J. Econometrics 54 (1992), 159-178.

[13] C. McKenzie, Unit roots and cointegration analysis: The impact on empirical analysis in economics, Japanese Economic Review 48 (1997), 18-28.

[14] P. C. B. Phillips and P. Perron, Testing for a unit root in time series regression, Biometrica 75 (1988), 335-346.

[15] S. E. Said and D. A. Dickey, Hypothesis testing in ARIMA (p, l, q) models, J. Amer. Stat. Assoc. 80 (1985), 369-374.

www.pphmj.com

■