CAMBRIDGE
UNIVERSITY PRESS

A RESIDUAL-BASED TEST OF THE NULL OF COINTEGRATION AGAINST THE ALTERNATIVE OF NO COINTEGRATION

YONGCHEOL SHIN
University of Cambridge

This paper proposes a residual-based test of the null of cointegration using a structural single equation model. It is shown that the limiting distribution of the test statistic for cointegration can be made free of nuisance parameters when the cointegrating relation is efficiently estimated. The limiting distributions are given in terms of a mixture of a Brownian bridge and vector Brownian motion. It is also shown that this test is consistent. Critical values are given for standard, demeaned, and detrended cases. Combining results from our test for cointegration with results from the Phillips-Ouliaris test for no cointegration, we find that there is evidence of cointegration between real consumption and real disposable income over the postwar period.

1. INTRODUCTION

There has recently been a surge of interest in the problem of testing for cointegration among economic time series. More generally, it is thought to be important, for both economic and statistical reasons, to be able to determine whether there is a stable long-run relationship between multiple economic series, even though each series is considered to be an $I(1)$ process. See Campbell and Perron [4] for further discussion.

However, most studies address the question of testing the null hypothesis of *no cointegration* rather than *cointegration*, and there have been very few attempts to test the cointegration hypothesis directly [5,10,27,34]. Park, Ouliaris, and Choi [18] and Park [16] consider tests of the null of cointegration against the alternative of no cointegration, but their tests are rather *ad hoc*. Since our primary interest is the hypothesis of cointegration, it is often argued that cointegration would be a more natural choice of the null hypoth-

91

esis. Yet no simple and straightforward residual-based statistical test of co-integration proceeds along these lines.

This paper develops a direct residual-based test for cointegration using a structural single equation model. The test is also shown to be an LM test and involves procedures that are designed to detect the presence of stationarity in the residuals of cointegrating regressions among the levels of economic time series. This procedure represents a modification of the methodology proposed by Kwiatkowski, Phillips, Schmidt, and Shin (hereafter KPSS) [11] who develop a test for stationarity in the univariate case. KPSS use the components model

$$y_t = \alpha + \delta t + X_t, \qquad X_t = \gamma_t + v_t, \qquad \gamma_t = \gamma_{t-1} + u_t,$$

where v_t are stationary and u_t are i.i.d. Then they test the null hypothesis that X_t has no random walk error component ($\sigma_u^2 = 0$). In this paper, we consider the cointegrating regression

$$y_t = \alpha + \delta t + Z_t'\beta + X_t,$$

where y_t and Z_t are scalar and m-vector $I(1)$ variables, and we develop appropriate procedures for testing the null hypothesis that X_t has no random walk error component. The basic difference between KPSS and this paper is just that $I(1)$ regressors in the cointegrating regression are added to the components model. Therefore, our tests can be viewed as a multivariate extension of the KPSS stationarity tests, just as the above-mentioned cointegration tests are multivariate extensions of unit root tests. Since our null hypothesis is cointegration rather than no cointegration, our cointegration test does not suffer from the "conceptual pitfalls" indicated by Phillips and Ouliaris [27].

It is well known that the limiting distribution of the least-squares estimators of the cointegrating vector is in general nonstandard and biased [22,24]. The distribution of cointegration test statistics based on the OLS estimator involves various nuisance parameters even asymptotically, and this poses a serious obstacle to inference. Most existing cointegration tests do not consider the issue of efficient estimation of the cointegrating vector. Recently, there have been many studies on the efficient estimation of the cointegrating vector [17,22,23,25,26,33,35]. Efficient estimation also simplifies the inference because it removes the nuisance parameters from the limiting distribution.

We will derive the limiting distribution of the test statistics for cointegration using an efficient estimator of the cointegrating vector, which will be shown not to involve any nuisance parameter dependency. Generally, the appropriately designed and transformed test statistics for cointegration should have the same limiting distribution even if we use different types of efficient estimators [17,23,25,26,33,35]. It will be shown that the limiting distribution

of the test statistic for cointegration involves a combination of a Brownian bridge and a functional of Brownian motion and also depends on the compound normal distribution (see [22]). Note that this is different from the limiting distribution of the test statistic for no cointegration, which depends on a functional of Brownian motion only and contains spurious regression distribution (see [20]).

Recently, Hansen [7] has proposed LM tests for parameter stability in the context of cointegrating regression models using the fully modified estimator of Phillips and Hansen [25]. His L_c test statistics in particular are similar to ours. He allows every coefficient to be a random walk and then tests the joint hypothesis that the variance of each random walk coefficient is zero. Under this null, the relationship *is* cointegrated, so his test is a test of the null of cointegration. However, his alternative is not the most natural one for a cointegration test, because under his alternative X_t is not $I(1)$. Our test fits his framework if all the coefficients except the intercept in the cointegrating relation are assumed to be constant, so only stability of the intercept is tested. See also Quintos and Phillips [31] and Tanaka [36].

We apply our cointegration test to an aggregate consumption function and find that there is evidence of cointegration between real consumption expenditure and real disposable income over the postwar period.

The plan of the paper is as follows. The preliminary results and the relevant asymptotic theory are presented in Sections 2 and 3. Comparisons with other cointegration tests are given in Section 4. The results of the application are discussed in Section 5. Discussions and concluding remarks are given in Section 6. An Appendix contains proofs of the paper's results.

For notational convenience we use "→" to signify weak convergence and "≡" to signify equality in distribution. Continuous stochastic processes such as the Brownian motion $B(r)$ on $[0,1]$ are simply written as B. We also write integrals with respect to Lebesgue measure such as $\int_0^1 B(r)\, dr$ simply as $\int_0^1 B$, and denote $\sum_{t=1}^T$ simply as $\sum$.

2. PRELIMINARY RESULTS

To derive a residual-based test for cointegration, we consider a single equation specification. There are three cases: the cointegrating regression without intercept and trend, with intercept only, and with intercept and trend.

$$y_t = Z_t'\beta + X_t, \tag{1}$$

$$y_t = \alpha_\mu + Z_t'\beta_\mu + X_t, \tag{2}$$

$$y_t = \alpha_\tau + \delta_\tau t + Z_t'\beta_\tau + X_t, \tag{3}$$

where in each case $X_t = \gamma_t + v_{1t}$, $\gamma_t = \gamma_{t-1} + u_t$, and $\Delta Z_t = v_{2t}$. Here u_t is i.i.d. $(0, \sigma_u^2)$, so γ_t is a random walk. Our null hypothesis of cointegration

is $\sigma_u^2 = 0$. We assume that u_t is independent of v_{1t}, which is not restrictive under the null but is restrictive under the alternative. The assumption that $\gamma_0 = 0$ entails no loss of generality so long as the regression includes an intercept, as in (2) and (3). The scalar v_{1t} and m-vector v_{2t} are stationary so that y_t and Z_t are scalar and m-vector $I(1)$ processes, respectively. Assume that $v_t = (v_{1t}, v_{2t}')'$ satisfies a multivariate invariance principle; the random sequence $\{v_t\}$ is assumed to be strictly stationary and ergodic with zero mean, finite variance, and spectral density matrix $f_{vv}(\lambda)$. See Park and Phillips [19]. Define the long-run covariance matrix of v_t as

$$\Omega = \lim_{T\to\infty} \text{var}(T^{-1/2}\Sigma v_t) \equiv \begin{pmatrix} \omega_{11} & \Omega_{21}' \\ \Omega_{21} & \Omega_{22} \end{pmatrix} \begin{matrix} 1 \\ m \end{matrix} = 2\pi f_{vv}(0). \tag{4}$$

We also define $\Sigma \equiv E(v_t v_t')$, $\Sigma_{21} \equiv E(v_{2t} v_{1t})$, $\Lambda \equiv \sum_{s=1}^{\infty} E(v_{t-s} v_t')$, $\Lambda_{21} \equiv \sum_{s=1}^{\infty} E(v_{2t-s} v_{1t})$, and $\Delta_{21} = \Sigma_{21} + \Lambda_{21}$. Note that Ω is the long-run covariance matrix of $(v_{1t}, \Delta Z_t')'$, which is different from the long-run covariance matrix of $(\Delta y_t, \Delta Z_t')'$ as defined in Phillips and Ouliaris [27]. Here cointegration does not generally lead to the singularity of Ω. Define the long-run variance of v_{1t} conditional on $\{v_{2t}\}$ as $\omega_{1\cdot 2} = \omega_{11} - \Omega_{21}' \Omega_{22}^{-1} \Omega_{21}$. We mainly deal with the case of "regular" cointegration (as defined in Park [17]), which excludes multicointegration as defined by Granger and Lee [6]. Since $\omega_{1\cdot 2}$ is always positive in this case, our test of the null hypothesis of cointegration does not have the conceptual pitfalls identified by Phillips and Ouliaris. Ω_{22} is assumed to be positive definite so that there are no cointegrating relationships among the regressors, Z_t.

We now construct the stochastic process B_T by $B_T = T^{-1/2} \sum_{j=1}^{[Tr]} v_j$, where $[Tr]$ is the integer part of tr. Under the above conditions, B_T converges weakly to B as $T \to \infty$, where B denotes a vector Brownian motion with covariance matrix Ω. We partition B as $B \equiv (B_1, B_2')'$ conformably with $(v_{1t}, v_{2t}')'$. B_1 and B_2 are not generally independent unless $\Omega_{21} = 0$. Under additional regularity conditions (see Lemma 2.1 of Park and Phillips [19], Phillips [21], and Theorem 4.1 of Hansen [9]), the following preliminaries hold[1]:

$$T^{-3/2}\Sigma Z_t \to \int_0^1 B_2, \qquad T^{-2}\Sigma Z_t Z_t' \to \int_0^1 B_2 B_2', \qquad T^{-5/2}\Sigma t Z_t \to \int_0^1 r B_2,$$

$$T^{-3/2}\Sigma t v_{1t} \to \int_0^1 r dB_1 \quad \text{and} \quad T^{-1}\Sigma Z_t v_{1t} \to \int_0^1 B_2 dB_1 + \Delta_{21}.$$

We will derive a residual-based test of the null of cointegration, which is a direct extension of the LM test of univariate stationarity in KPSS [11]. Our test statistic for cointegration is both the one-sided LM test statistic and the locally best invariant test statistic for the hypothesis $\sigma_u^2 = 0$, under the

stronger assumptions that the regressors are strictly exogenous, the error v_{1t} is i.i.d. normal, and u_t is i.i.d. normal. Nyblom [14] also shows that the statistic is an approximate LM statistic even if v_{1t} is nonnormal and u_t is a martingale difference sequence.[2] KPSS use the components model

$$y_t = \alpha + \delta t + X_t, \qquad X_t = \gamma_t + v_t, \qquad \gamma_t = \gamma_{t-1} + u_t, \tag{5}$$

and then test the null hypothesis that X_t has no random walk error component.

Now we consider the cointegration test, in which the null is simply $\sigma_u^2 = 0$ so that $\gamma_t = 0$ and X_t is $I(0)$ under the null. Following KPSS, let $\hat{X}_t$, $\hat{X}_{\mu t}$, and $\hat{X}_{\tau t}$ be the OLS residuals from the cointegrating regressions (1), (2), and (3), and define S_t, $S_{\mu t}$, and $S_{\tau t}$ as the partial sum processes of these residuals. Let $s^2(\ell)$, $s_\mu^2(\ell)$, and $s_\tau^2(\ell)$ be consistent semiparametric estimators of the long-run variance of the regression error v_{1t} (that is, of ω_{11}) under the null (see [1,2,8,13,28] for a discussion of possible estimators). Then the test statistics for cointegration in (1), (2), and (3) are derived as

$$CI \equiv T^{-2} \Sigma S_t^2 / s^2(\ell), \qquad CI_\mu \equiv T^{-2} \Sigma S_{\mu t}^2 / s_\mu^2(\ell),$$

$$\text{and} \quad CI_\tau \equiv T^{-2} \Sigma S_{\tau t}^2 / s_\tau^2(\ell) \tag{6}$$

It is well known that the single-equation OLS estimators generally involve second-order bias terms due to the presence of Δ_{21}, the correlation between v_{1t} and v_{2t}.[3] Although the cointegrating vector β ($m \times 1$) based on the OLS estimation is superconsistent, it is inefficient (see [22]). In addition, inference is complicated because of the dependence of the limiting distribution of the estimated cointegration vector on nuisance parameters. Therefore, it is clear that the limiting distribution of the test statistics for cointegration based on the OLS residuals involves a function of the nuisance parameters ω_{11}, Ω_{22}, and Δ_{21}. To avoid this problem, either we need a strict exogeneity assumption (in Theorem 1) or we need efficient estimation (in Theorem 2).

THEOREM 1. *Assume that $\Omega_{21} = 0$; that is, Z_t is strictly exogenous with respect to v_{1t}. Then, the test statistics for cointegration, CI, CI_μ, CI_τ have the following limiting distributions:*

$$CI \to \int_0^1 Q^2, \qquad CI_\mu \to \int_0^1 Q_\mu^2, \qquad \text{and} \quad CI_\tau \to \int_0^1 Q_\tau^2,$$

where

$$Q = W_1 - \left(\int_0^r W_2' \right) \left(\int_0^1 W_2 W_2' \right)^{-1} \left(\int_0^1 W_2 \, dW_1 \right),$$

$$Q_\mu = V_1 - \left(\int_0^r \overline{W}_2' \right) \left(\int_0^1 \overline{W}_2 \overline{W}_2' \right)^{-1} \left(\int_0^1 \overline{W}_2 \, dW_1 \right),$$

and

$$Q_\tau = V_1^{(2)} - \left(\int_0^r W_2^{*\prime} \right) \left(\int_0^1 W_2^* W_2^{*\prime} \right)^{-1} \left(\int_0^1 W_2^* \, dW_1 \right). \qquad \blacksquare$$

W_1 and W_2 are *independent* scalar and m-vector *standard* Brownian motion. $\overline{W}_2 = W_2 - \int_0^1 W_2$ is an m-vector standard demeaned Brownian motion, $W_2^* = W_2 + (6r - 4)\int_0^1 W_2 + (-12r + 6)\int_0^1 rW_2$ is an m-vector standard demeaned and detrended Brownian motion, $V_1 = W_1 - rW_1(1)$ is a standard Brownian bridge, and $V_1^{(2)} = W_1 + (2r - 3r^2)W_1(1) + (-6r + 6r^2)\int_0^1 W_1$ is a standard second-level Brownian bridge. See KPSS [11] for further discussion and references.

Theorem 1 shows that the statistics for cointegration based on the OLS estimation can be made to be free of nuisance parameter with the assumption of strict exogeneity, because Q, Q_μ, and Q_τ depend only on the dimension of $Z_t(m)$ and different functionals of standard Brownian motion. Note that if we include level and/or time trend as the regressors, the limiting distribution of the cointegration test statistic is a combination of a standard Brownian bridge (a standard second-level Brownian bridge), which is constructed from the cointegrating regression error, and a functional of an m-vector standard demeaned (detrended) Brownian motion, which is constructed from the m-vector integrated regressors.

3. ASYMPTOTIC THEORY: A MODIFIED SINGLE EQUATION MODEL

Generally, the exogeneity assumption given in Theorem 1 is too restrictive in time series modeling. The cointegration tests developed in the previous section are not expected to be robust to the problem of endogenous regressors, because the limiting distributions of those statistics would then involve nuisance parameters. We now show that the test statistics for cointegration based on efficient estimation of the regression coefficients do not involve any nuisance parameter asymptotically. In this section, we use the linear leads and lags OLS estimator as defined by Saikkonen [33] to prove that the limiting distributions of the cointegration test statistics based on efficient estimation are the same as in Theorem 1. In general, when v_t is serially correlated, it is not sufficient only to consider the contemporaneous correlation between v_{1t} and v_{2t}. Therefore, the approach given in the last section is modified by using not only present but also past and future values of ΔZ_t as additional regressors. By the stationarity of v_t, we would expect that values of ΔZ_t in the very remote past and future can only have a negligible impact on y_t. The following additional assumptions are now required:

Condition 1. The spectral density matrix $f_{vv}(\lambda)$ is bounded away from zero.

$$f_{vv}(\lambda) \geq aI_T, \qquad \lambda \in [0, \pi], \qquad a > 0. \tag{7}$$

Condition 2. The covariance function of v_t is absolutely summable.

$$\sum_{j=-\infty}^{\infty} \|\Gamma(j)\| < \infty, \tag{8}$$

where $|\Gamma(j)| = E(v_t v'_{t+j})$ and $\|\cdot\|$ is the standard Euclidean norm. When Conditions 1 and 2 hold, $v_{1t} = \sum_{j=-\infty}^{\infty} v'_{2t-j} \pi_j + \varepsilon_t$, where $\sum_{j=-\infty}^{\infty} \|\pi_j\| < \infty$ and ε_t is a stationary process such that $E(v_{2t}\varepsilon_{t+j}) = 0$ for $j = 0, \pm 1, \pm 2, \ldots$. Furthermore, $f_{\varepsilon\varepsilon}(\lambda) = f_{v_1 v_1}(\lambda) - f_{v_1 v_2}(\lambda) f_{v_2 v_2}^{-1}(\lambda) f_{v_2 v_1}(\lambda)$, which implies that $2\pi f_{\varepsilon\varepsilon}(0) = \omega_{1\cdot2}$. Then, equation (1) can be transformed into $y_t = Z'_t \beta + \sum_{j=-K}^{K} \Delta Z'_{t-j} \pi_j + \varepsilon_t^*$, where $\varepsilon_t^* = \varepsilon_t + \sum_{|j|>K}^{\infty} v'_{2t-j} \pi_j$. Since the sequence $\{\pi_j\}$ is absolutely summable, $\pi_j \approx 0$ for $|j| > K$, K large enough. For simplicity, we use the same truncation value for both leads and lags of ΔZ_t. If $\pi_j = 0$ for $|j| > K$, we have $\varepsilon_t^* = \varepsilon_t$. Then ε_t^* is strictly exogenous with respect to v_{2t} so that the endogeneity problem in simple least-squares estimation can be eliminated. However, we generally cannot assume that $\pi_j = 0$ for $|j| > K$ with K fixed; instead, we choose K such that, as $T \to \infty$, and $K \to \infty$,

$$K^3/T \to 0, \quad \text{and} \quad T^{1/2} \sum_{|j|>K}^{\infty} \|\pi_j\| \to 0. \tag{9}$$

See Saikkonen [33] for further details. The order of K can also be chosen using model selection criterion such as AIC or BIC. For further discussion of this matter, see Phillips and Ploberger [29], in which they also suggest a new *consistent* model selection criteria "PIC," which allows for automatic order selection of the stochastic regressors (and also the degree of the deterministic trend) and is designed to accommodate nonstationary series. However, the assumption given in (9) is sufficient to develop the asymptotic distribution of the test statistic for cointegration in this section. The same kind of extensions can also be applied to (2) and (3). Therefore, for a chosen lag truncation (K), we consider the modified least-squares regression equations:

$$y_t = Z'_t \tilde{\beta} + \sum_{j=-K}^{K} \Delta Z'_{t-j} \tilde{\pi}_j + \tilde{\varepsilon}_t^*, \tag{10}$$

$$y_t = \tilde{\alpha}_\mu + Z'_t \tilde{\beta}_\mu + \sum_{j=-K}^{K} \Delta Z'_{t-j} \tilde{\pi}_{\mu j} + \tilde{\varepsilon}_{\mu t}^*, \tag{11}$$

$$y_t = \tilde{\alpha}_\tau + \tilde{\delta}_\tau t + Z'_t \tilde{\beta}_\tau + \sum_{j=-K}^{K} \Delta Z'_{t-j} \tilde{\pi}_{\tau j} + \tilde{\varepsilon}_{\tau t}^*. \tag{12}$$

We now construct the stochastic process B_T^* by $B_T^* = T^{-1/2} \sum_{j=1}^{[Tr]} w_j$, where $w_t = (\varepsilon_t, v'_{2t})'$. B_T^* converges weakly to B^* as $T \to \infty$, where B^* denotes a vector Brownian motion with covariance matrix Ω^*, which is block-diagonal; that is, $\Omega^* = \mathrm{diag}(\omega_{1\cdot2}, \Omega_{22})$. We partition B^* as $B^* \equiv (B_{1\cdot2}, B'_2)'$ conform-

ably with $(\varepsilon_t, v'_{2t})'$, where $B_{1\cdot2} \equiv B_1 - \Omega'_{21}\Omega_{22}^{-1}B_2$. Note that $B_{1\cdot2}$ is independent of the m-vector Brownian motion, B_2. See [22,33] for the form of $B_{1\cdot2}$.

LEMMA 1. *Let $\tilde{\beta}$ and $\tilde{\pi}_j$, $\tilde{\alpha}_\mu$, $\tilde{\beta}_\mu$, and $\tilde{\pi}_{\mu j}$, and $\tilde{\alpha}_\tau$, $\tilde{\delta}_\tau$, $\tilde{\beta}_\tau$, and $\tilde{\pi}_{\tau j}$ be the OLS estimators obtained from* (10), (11), *and* (12). *Then,*

$$T(\tilde{\beta} - \beta) \to \left(\int_0^1 B_2 B_2'\right)^{-1}\left(\int_0^1 B_2\, dB_{1\cdot2}\right),$$

$$T(\tilde{\beta}_\mu - \beta_\mu) \to \left(\int_0^1 \bar{B}_2 \bar{B}_2'\right)^{-1}\left(\int_0^1 \bar{B}_2\, dB_{1\cdot2}\right),$$

$$T(\tilde{\beta}_\tau - \beta_\tau) \to \left(\int_0^1 B_2^* B_2^{*'}\right)^{-1}\left(\int_0^1 B_2^*\, dB_{1\cdot2}\right),$$

$$T^{1/2}(\tilde{\alpha}_\mu - \alpha_\mu) \to B_{1\cdot2}(1) - \left(\int_0^1 B_2'\right)\left(\int_0^1 \bar{B}_2 \bar{B}_2'\right)^{-1}\left(\int_0^1 \bar{B}_2\, dB_{1\cdot2}\right),$$

$$T^{1/2}(\tilde{\alpha}_\tau - \alpha_\tau) \to 4B_{1\cdot2}(1) - 6\int_0^1 r\, dB_{1\cdot2}$$

$$+ \left(-4\int_0^1 B_2' + 6\int_0^1 rB_2'\right)\left(\int_0^1 B_2^* B_2^{*'}\right)^{-1}\left(\int_0^1 B_2^*\, dB_{1\cdot2}\right),$$

$$T^{3/2}(\tilde{\delta}_\tau - \delta_\tau) \to -6B_{1\cdot2}(1) + 12\int_0^1 r\, dB_{1\cdot2}$$

$$+ \left(6\int_0^1 B_2' - 12\int_0^1 rB_2'\right)\left(\int_0^1 B_2^* B_2^{*'}\right)^{-1}\left(\int_0^1 B_2^*\, dB_{1\cdot2}\right),$$

$$\left(\frac{T}{K}\right)^{1/2}\sum_{j=-K}^K (\tilde{\pi}_j - \pi_j) = O_p(1), \qquad \left(\frac{T}{K}\right)^{1/2}\sum_{j=-K}^K (\tilde{\pi}_{\mu j} - \pi_{\mu j}) = O_p(1),$$

and

$$\left(\frac{T}{K}\right)^{1/2}\sum_{j=-K}^K (\tilde{\pi}_{\tau j} - \pi_{\tau j}) = O_P(1). \qquad \blacksquare$$

Here $\bar{B}_2 = B_2 - \int_0^1 B_2$ is an m-vector demeaned Brownian motion, and $B_2^* = B_2 + (6r - 4)\int_0^1 B_2 + (-12r + 6)\int_0^1 rB_2$ is an m-vector demeaned and detrended Brownian motion. Following Stock and Watson [35], we may call this the dynamic OLS estimator. The estimates of the cointegrating vectors in (10), (11), and (12) are not only superconsistent but also efficient. Note that these asymptotics fall within the LAMN (locally asymptotically mixed normal) family (see Phillips [22]).

Based on the above results, let $\tilde{\varepsilon}_j^*$, $\tilde{\varepsilon}_{\mu j}^*$, and $\tilde{\varepsilon}_{\tau j}^*$ be the correct OLS residuals obtained from (10), (11), and (12), and $\tilde{S}_t = \sum_{j=1}^t \tilde{\varepsilon}_j^*$, $\tilde{S}_{\mu t} = \sum_{j=1}^t \tilde{\varepsilon}_{\mu j}^*$,

and $\tilde{S}_{\tau t} = \sum_{j=1}^{t} \tilde{\varepsilon}_{\tau j}^*$. Let $\tilde{s}^2(\ell)$, $\tilde{s}_\mu^2(\ell)$, and $\tilde{s}_\tau^2(\ell)$ be semiparametric consistent estimators of the long-run variance of ε_t in (10), (11), and (12), based on $\tilde{\varepsilon}_j^*$, $\tilde{\varepsilon}_{\mu j}^*$, and $\tilde{\varepsilon}_{\tau j}^*$, respectively. Then the modified test statistics for cointegration are defined as

$$C \equiv T^{-2} \Sigma \tilde{S}_t^2 / \tilde{s}^2(\ell), \qquad C_\mu \equiv T^{-2} \Sigma \tilde{S}_{\mu t}^2 / \tilde{s}_\mu^2(\ell), \qquad \text{and}$$

$$C_\tau \equiv T^{-2} \Sigma \tilde{S}_{\tau t}^2 / \tilde{s}_\tau^2(\ell). \tag{13}$$

THEOREM 2. *The limiting distributions of the modified test statistics for cointegration, C, C_μ, and C_τ are the same as in Theorem 1.*   ∎

Although the asymptotic results in Theorem 2 are obtained using the dynamic OLS estimators, it is important to note that these results are not affected, at least asymptotically, if we use instead different types of efficient estimators [17,23,25,26], because the limiting distribution of these estimators is the same. For example, if we estimate the cointegrating regression by the fully modified procedure of Phillips and Hansen [25], then we can use the residuals (obtained from the appropriately modified regression) to construct the test statistics.[4]

Additionally, there are a number of other important points to bear in mind. If there is cointegration in the demeaned specification given in (11), this may correspond to "deterministic cointegration," which implies that the same cointegrating vector eliminates deterministic trends as well as stochastic trends. But if the linear stationary combinations of $I(1)$ variables have nonzero linear trends as given in (12), this corresponds to "stochastic cointegration." For definitions of deterministic and stochastic cointegration, see Ogaki and Park [15].

Critical values for C, C_μ, and C_τ are given in Table 1 with $m = 1$ to 5. Critical values are calculated via a Monte Carlo simulation, using a sample size of 2000, and the random number generator GASDEV/RAN3 of Press, Flannery, Teukolsky, and Vetterling [30]. When $m = 1$, 2, and 3 we use 50,000 replications. Otherwise, we use 20,000 replications.

4. COMPARISON WITH OTHER COINTEGRATION TESTS

Phillips and Ouliaris [27] provide residual-based tests for the presence of no cointegration in multiple time series. Although their tests are similar to our tests in the sense that they are based on the residuals of the cointegrating regression, their tests are residual-based *unit root* tests. We now show that the limiting distribution of our statistic under the null of no cointegration ($\sigma_u^2 > 0$) is based on the same basic functional of Brownian motion as that of the Phillips and Ouliaris test statistic, although the final form of the limiting distributions of the two test statistics is quite different.

Table 1. Critical values for the cointegration test statistics

Fractile	Number of Regressors (m)				
	1	2	3	4	5
Standard (C)					
0.010	0.027	0.023	0.021	0.018	0.016
0.025	0.034	0.029	0.025	0.022	0.020
0.050	0.043	0.035	0.030	0.026	0.023
0.100	0.057	0.046	0.038	0.033	0.029
0.200	0.083	0.065	0.053	0.045	0.039
0.300	0.113	0.087	0.070	0.058	0.050
0.400	0.150	0.115	0.090	0.074	0.063
0.500	0.199	0.150	0.117	0.096	0.081
0.600	0.267	0.199	0.154	0.125	0.104
0.700	0.368	0.271	0.209	0.167	0.139
0.800	0.527	0.391	0.295	0.236	0.198
0.900	0.841	0.624	0.475	0.374	0.307
0.950	1.199	0.895	0.682	0.537	0.433
0.975	1.601	1.190	0.926	0.715	0.580
0.990	2.126	1.623	1.305	1.003	0.781
Demeaned (C_μ)					
0.010	0.020	0.017	0.015	0.014	0.013
0.025	0.024	0.021	0.018	0.016	0.015
0.050	0.029	0.024	0.021	0.019	0.017
0.100	0.035	0.029	0.025	0.022	0.019
0.200	0.046	0.037	0.031	0.027	0.024
0.300	0.057	0.045	0.037	0.031	0.027
0.400	0.069	0.053	0.043	0.036	0.031
0.500	0.083	0.063	0.050	0.042	0.046
0.600	0.101	0.074	0.059	0.048	0.041
0.700	0.125	0.090	0.070	0.057	0.047
0.800	0.161	0.115	0.088	0.069	0.057
0.900	0.231	0.163	0.121	0.094	0.075
0.950	0.314	0.221	0.159	0.121	0.097
0.975	0.407	0.285	0.203	0.153	0.120
0.990	0.533	0.380	0.271	0.208	0.158
Detrended (C_τ)					
0.010	0.015	0.014	0.012	0.011	0.011
0.025	0.017	0.016	0.014	0.013	0.012
0.050	0.020	0.018	0.016	0.015	0.014
0.100	0.024	0.021	0.019	0.017	0.016
0.200	0.030	0.026	0.023	0.021	0.019
0.300	0.035	0.030	0.027	0.024	0.021
0.400	0.040	0.035	0.030	0.027	0.024
0.500	0.046	0.040	0.034	0.030	0.027

TABLE 1 continued.

Fractile	Number of Regressors (m)				
	1	2	3	4	5
Detrended (C_τ) *continued*					
0.600	0.053	0.045	0.039	0.034	0.030
0.700	0.062	0.052	0.045	0.039	0.034
0.800	0.075	0.063	0.054	0.046	0.040
0.900	0.097	0.081	0.069	0.056	0.050
0.950	0.121	0.101	0.085	0.073	0.061
0.975	0.147	0.122	0.102	0.088	0.072
0.990	0.184	0.150	0.126	0.109	0.087

THEOREM 3. *Under the alternative hypothesis of no cointegration* ($\sigma_u^2 > 0$), *the modified test statistics* C, C_μ, *and* C_τ *(normalized by* ℓ/T) *have the following limiting distributions:*

$$(\ell/T)C \rightarrow \int_0^1 \left(\int_0^a Q_{PO}\right)^2 da/L \int_0^1 Q_{PO}^2,$$

$$(\ell/T)C_\mu \rightarrow \int_0^1 \left(\int_0^a Q_{PO\mu}\right)^2 da/L \int_0^1 Q_{PO\mu}^2, \quad and$$

$$(\ell/T)C_\tau \rightarrow \int_0^1 \left(\int_0^a Q_{PO\tau}\right)^2 da/L \int_0^1 Q_{PO\tau}^2,$$

where

$$Q_{PO} = W_1 - W_2'\left(\int_0^1 W_2 W_2'\right)^{-1}\left(\int_0^1 W_2 W_1\right),$$

$$Q_{PO\mu} = \overline{W}_1 - \overline{W}_2'\left(\int_0^1 \overline{W}_2 \overline{W}_2'\right)^{-1}\left(\int_0^1 \overline{W}_2 W_1\right), \quad and$$

$$Q_{PO\tau} = W_1^* - W_2^{*\prime}\left(\int_0^1 W_2^* W_2^{*\prime}\right)^{-1}\left(\int_0^1 W_2^* W_1\right). \qquad \blacksquare$$

Here the constant L is defined by $L = \int_{-1}^1 k(s)\, ds$, where $k(s)$ represents the weight function used in $\tilde{s}^2(\ell)$ (see the Appendix). Theorem 3 shows that our cointegration test statistics C, C_μ, and C_τ are consistent; that is, they diverge at a rate of (T/ℓ) under the alternative. However, it should be noted that our cointegration test statistics are critically dependent on the choice of

the lag truncation parameter ℓ, and that the behavior of ℓ is critical for the test to have good power.

Note that the limiting distributions of the cointegration test statistics under the alternative are also free of nuisance parameters, because the scale effect from the variance $\sigma_u^2 > 0$ cancels out. This is quite similar to the results obtained in KPSS [11]. Generally, Q (which is the basic functional of the limiting distribution of the test statistic for cointegration) is different from Q_{PO} (which is the basic functional of the limiting distribution of the test statistic for no cointegration) in two ways. First, Q_{PO} involves a spurious regression distribution $\left(\int_0^1 W_2 W_2'\right)^{-1}\left(\int_0^1 W_2 W_1\right)$, while Q has a compound normal distribution $\left(\int_0^1 W_2 W_2'\right)^{-1}\left(\int_0^1 W_2 \, dW_{1 \cdot 2}\right)$. See Phillips [22]. Second, Q_{PO} involves the functional of integrated regressors, W_2, while Q involves the functional of the partial sum process of the integrated regressors, $\int_0^r W_2$. Since the primary interest is the hypothesis of cointegration, we may conclude that our test is often a more natural choice. The same kind of arguments can also be made against most existing cointegration tests based on the null of no cointegration.

Recently, Hansen [7] derives the limiting distribution of the LM test statistic for parameter stability in the context of cointegrating regression models. Following Hansen, rewrite (2) as

$$y_t = A_1 + A_2 Z_t + X_t = A_1 + A_2 Z_t + \gamma_t + v_{1t}, \tag{14}$$

which can also be written as

$$y_t = A_{1t} + A_2 Z_t + v_t \quad \text{with} \quad A_{1t} = A_1 + \gamma_t. \tag{15}$$

This shows that the alternative hypothesis of a random walk only in the intercept is identical to "no cointegration," so that the test statistic in this case is a test of the null of cointegration against the alternative of no cointegration. In other words, our proposed test fits with Hansen's framework if all coefficients except the intercept term are assumed to be constant, so only stability of intercept is tested. However, Hansen's L_c statistic is not designed as a direct test for cointegration, because it actually tests the stability of *all* coefficients, not just the intercept term. As noted by Hansen, a rejection of the null of constant parameters does not imply the particular alternative the test is designed to detect. In particular, X_t is not an $I(1)$ process under his alternative.

Quintos and Phillips [31] derive similar LM tests for parameter constancy in cointegrating regressions using the single-equation varying coefficient regression. Although their test statistic is a test of the null of cointegration and has the advantage of detecting (specific) cointegration failure caused by subset of parameters, their alternative is not $I(1)$ either (see also Tanaka [36]). Therefore, these tests may not be as powerful as our test against the alternative of no cointegration.

5. EMPIRICAL APPLICATIONS: AGGREGATE CONSUMPTION FUNCTION

Three points are worth noting before we apply our cointegration test. First, we should pretest to see whether all dependent and independent variables are $I(0)$ or $I(1)$. We use both the KPSS stationarity test and the augmented Dickey–Fuller unit root test to check this property. Second, efficient estimation should be used to allow for correlation between the regression errors and first-differenced regressors. Here we use the dynamic OLS method, and we choose $K = 5$ (which is approximately equal to $T^{1/3}$ in our application). This choice is also consistent with simulation results of Stock and Watson [35]. Finally, we use semiparametric corrections to remove persistent serial correlation of the residual process and therefore the long-run variance of the cointegrating regression residual is estimated using the Bartlett window. We choose $\ell = 10$ as the appropriate choice for the lag truncation parameter, based on the consideration that the residual from the cointegrating regression is generally very persistent, and based on the results of the KPSS Monte Carlo simulation which suggest that this choice is a compromise between the large size distortions that we expect for smaller number of lags and the low power that we expect for larger number of lags in the context of the univariate stationarity tests.

Since the empirical results of applying cointegration tests are critically dependent on the choice of ℓ and K, especially on ℓ, applied economists should pay attention to the central importance of these choices. General treatments of these choices are given in [29] for automatic choice of K using consistent model selection criteria, and in [1,2] for a data-dependent choice of ℓ. In our example of an aggregate consumption function, however, the empirical results of applying our cointegration tests are not very sensitive to the choice of K after the value of ℓ is selected. (These results are not reported but are available upon request.) We also note that the use of the plug-in bandwidth parameter recommended in [1] always gives a very large value of ℓ when there is heavy autocorrelation (e.g., $\ell = 41$ is chosen for the Bartlett window when the estimate of the AR(1) parameter is 0.9 and $T = 178$, which is very plausible empirically), in which case the null of cointegration is rarely rejected. Unfortunately, it can be easily shown that the test statistic for cointegration using a prewhitened kernel estimator of the long-run variance with the plug-in bandwidth parameter recommended in [2] is *not* consistent against the alternative of no cointegration. Therefore, we may conclude that our choices of ℓ and K are relatively reasonable.

We now test for a stable long-run consumption function using data obtained from Citibase Data for 1947:1-1991:2. GC is nominal aggregate quarterly U.S. consumption expenditure; GCN is nominal aggregate quarterly U.S. nondurable consumption expenditure; GCS is nominal aggregate quar-

terly U.S. service consumption expenditure; GYD is nominal total disposal income; GYD82 is real total disposable income in 1982 dollars; and GPOP is total population. The price deflator (P) is obtained by dividing GYD by GYD82, and is used to transform the variables (except for GPOP) into real units.

We consider two types of consumption data sets. First, we consider the consumption function using variables measured in total units; therefore, we use real total consumption expenditure (GC/P), real NDS consumption expenditure (GCN/P + GCS/P), and real disposable income (GYD82). Second, we consider the consumption function using variables measured in per capita (PC) units; we use PC real consumption expenditure (real total consumption expenditure/GPOP), PC real NDS consumption expenditure (real NDS consumption expenditure/GPOP), and PC real disposable income (GYD82/GPOP). All consumption and income variables, after construction as just described, are then measured in logarithms.

In Table 2, the results of applying the KPSS stationarity test and the augmented Dickey–Fuller unit root test to the above variables are given. It is found that real total consumption expenditure, real NDS consumption, real disposable income, PC real NDS consumption expenditure, and PC real disposable income are $I(1)$ processes, possibly with drift, because for each we reject the stationarity hypothesis but not the unit root hypothesis. For PC real consumption expenditure, it is not clear whether these series are trend stationary or follow an $I(1)$ process with drift, because we do not reject either trend stationarity or the unit root hypothesis; but since the null of trend

TABLE 2. The KPSS stationarity test and the ADF unit root test

Country	Stationarity Test[a]		Unit Root Test[b]	
	$\hat{\eta}_\mu$	$\hat{\eta}_\tau$	τ_μ	τ_τ
Real total consumption	1.7187*[d]	0.2636*	−1.1403	−1.4389
Real total NDS consumption	1.7194*	0.2262*	−1.0173	−1.2637
Real disposable income	1.7173*	0.3356*	−1.8501	−0.5050
Per capita real consumption	1.7142*	0.1420[c]	−0.3270	−2.4982
Per capita real NDS consumption	1.7171*	0.1646*	0.0349	−2.8482
Per capita real disposable income	1.7150*	0.1953*	−0.8654	−1.3256

[a]We use the KPSS stationarity test with $\ell = 10$. Upper tail 5% critical values for level stationarity ($\hat{\eta}_\mu$) and trend stationarity ($\hat{\eta}_\tau$) tests are 0.461 and 0.146.
[b]We use the ADF unit root test statistics, τ_μ and τ_τ with five augmentations. Lower tail 5% critical values for τ_μ and τ_τ are −2.86 and −3.41.
[c]We reject the trend stationarity at the 10% level.
[d]*implies the rejection of the null hypothesis at the 5% level.

stationarity is rejected at the 10% level, we may conclude that PC real consumption expenditure is close to an $I(1)$ process.

In Table 3 we present the results of applying our cointegration test and the Phillips–Ouliaris no-cointegration test to the consumption functions. We use demeaned and detrended equations (11) and (12) because it is reasonable to include intercept and/or trend in multiple time series regression. Since the concept of deterministic cointegration is stronger than the concept of stochastic cointegration, it is sensible that we first test for the presence of stochastic cointegration and then test for the presence of deterministic cointegration sequentially. There is strong evidence of stochastic cointegration between real total consumption (real NDS consumption) expenditure and real disposable income, because we do not reject the null hypothesis of cointegration but we do reject the null of no cointegration in the detrended specification. However, there is no clear evidence of stochastic cointegration between PC real consumption expenditure and PC real disposable income, and there is strong evidence of no cointegration between PC real NDS consumption expenditure and PC real disposal income.

Next we check for the presence of deterministic cointegration using the demeaned specification. There may be weak evidence of deterministic cointegration between real total consumption expenditure and real disposable income — although we reject both hypotheses, the null of cointegration is not rejected at the 2.5% level. On the other hand, it is not clear whether there is deterministic cointegration between real NDS consumption expenditure and real disposable income, because we fail to reject both hypotheses. Finally, when we use the data measured in per capita terms, the results are not clear either, so that there is no evidence of deterministic cointegration in this case.

We may conclude that there is *weak evidence* of deterministic cointegration between real total consumption expenditure and real disposable income over the postwar time period. On the other hand, there is *strong evidence* of stochastic cointegration between real NDS consumption expenditure and real disposable income. However, we do not find any evidence of cointegration for the consumption function using the data measured in per capita units. One may note in Table 3 that including a deterministic trend in the consumption function reduces the marginal propensity to consume by a considerable amount. This is probably evidence against the correctness of the specification. More formal testing procedures such as the Wald test for the restriction on the coefficients on intercept and/or trend could be used to arrive at more formal conclusions [33].

6. DISCUSSION AND CONCLUDING REMARKS

We have derived the limiting distribution of a residual-based test for cointegration using a structural single equation model and tabulated its critical val-

TABLE 3. Tests for cointegration and tests for no cointegration

	α	δ	β	C_μ or $C_\tau{}^a$	Z_a Test[b]	Z_t Test[b]
Between real total consumption and real disposable income						
Demeaned	−0.0728		0.9964	0.3231^c	−33.4700*	−4.3031*
	$(0.0252)^d$		(0.0032)			
Detrended	2.0700	0.0028	0.6686	0.0869	−48.8490*	−5.3642*
	(0.1417)	(0.0002)	(0.0216)			
Between real NDS consumption and real disposable income						
Demeaned	−0.1971		0.9948	0.2896	−17.0490	−2.9648
	(0.0257)		(0.0033)			
Detrended	2.3492	0.0033	0.6052	0.0379	−33.1480*	−4.7760*
	(0.0988)	(0.0001)	(0.0151)			
Between per capita real total consumption and per capita real disposable income						
Demeaned	−0.0790		0.9922	0.3161^c	−27.5650*	−4.2766*
	(0.0104)		(0.0048)			
Detrended	0.3569	0.0015	0.7096	0.2699*	−27.3770*	−4.2598*
	(0.0566)	(0.0002)	(0.0365)			
Between per capita real NDS consumption and per capita real disposable income						
Demeaned	−0.2222		0.9932	0.2860	−16.8750	−2.9510
	(0.0104)		(0.0048)			
Detrended	0.4105	0.0022	0.5830	0.2788*	−18.6490	−3.9250*
	(0.0425)	(0.0001)	(0.0274)			

[a]We use $\ell = 10$ and $K = 5$ when testing for cointegration. Upper 5% critical values for demeaned and detrended cases are 0.314 and 0.121.

[b]We use the Phillips and Ouliaris test statistics Z_a and Z_t to test for no cointegration, which are based on the simple OLS regression with $\ell = 10$. Lower tail 5% critical values for Z_a (demeaned, and detrended) tests are −20.4935, and −27.0866. Lower tail 5% critical values for Z_t (demeaned, and detrended) tests are −3.3654, and −3.8.

[c]We do not reject the null of cointegration at the 2.5% level.

[d]Numbers in ( ) indicate the OLS standard error of the coefficient.

ues via a Monte Carlo simulation. The limiting distribution does not involve any nuisance parameter dependency, because the test is based on efficient estimation of the regression coefficients.

Existing tests for cointegration are actually tests of the null hypothesis of no cointegration. Combining our tests of cointegration with existing tests of no cointegration may help to lead to more definite conclusions than either set of tests separately. For example, there is strong evidence of cointegration if we reject the null of no cointegration and fail to reject the null of cointegration. Similarly, there is strong evidence of no cointegration if we reject the null of cointegration and fail to reject the null of no cointegration. If neither hypothesis can be rejected, the data are not sufficiently informative to decide between cointegration and no cointegration. Finally, if both hypotheses are rejected, doubt is cast upon the validity and usefulness of the model, and more complicated alternatives (e.g., fractional integration) might need to be considered.

We apply our cointegration test to a bivariate empirical example of an aggregate consumption function. To get more comprehensive results, we combine our results with the results obtained using the Phillips–Ouliaris test statistics for no cointegration. We find that there is evidence of cointegration between real total consumption (real NDS consumption) expenditure and real disposable income over the postwar period.

If there are more than two regressors in any meaningful economic relationship (e.g., a money demand function), our assumption that there is not a cointegrating relationship among the regressors needs to be checked. Thus, after we pretest whether the dependent variable and all regressors are $I(1)$, we need to check whether or not there is cointegration among the regressors. In the case that there is cointegration among the regressors, we conjecture that the limiting distribution of the test statistic for cointegration is not fundamentally affected; that is, it depends only on the rank of the covariance matrix of the regressors, Ω_{22} (see Wooldridge [37]). This is the case in which there is more than one cointegrating vector among the dependent variable and the regressors. Therefore, possible future research could be in the direction of extending our results to find the system-based tests for cointegration. This could be a useful addition to the Johansen tests [10], which are basically a multivariate extension of unit root tests.

The results of this paper are mainly asymptotic. It has been shown that both parametric and semiparametric corrections or any combination can generally be used to deal with serial correlation of the residuals and the endogeneity of the regressors. In this paper, we suggest a conservative choice of the number of lags used in semiparametrically estimating the long-run variance of the residual of the cointegrating regression and of the number of leads and lags of first-differenced regressors to be used in parametrically estimating the cointegrating regression. However, the finite sample performance of our cointegration tests using different efficient estimators of the

cointegrating vector, and using different lag windows and different choices of ℓ and K (probably selected in a data-dependent way) is still unknown. Considering the fact that this choice matters in empirical applications using economic data with typical sample sizes (100 to 200), much care should be taken. Further research will be needed.

NOTES

1. This result is simply assumed in [19]. One of the referees kindly informs me that a proof for linear processes is shown in [21] and that a proof for mixing process is shown in [9].

2. One of the referees makes this point. For the derivation of an LM test statistic in the form given in (6) see [11] and [12]. Our LM test statistic for cointegration is represented as $\Sigma S_t^2/\hat{\sigma}_{v_1}^2$ (apart from an appropriate normalization), where $\hat{\sigma}_{v_1}^2$ is a consistent estimate of the error variance (the sum of squared residuals, divided by T), under the assumptions that the regressors are strictly exogenous, the error v_{1t} is i.i.d. $N(0, \sigma_{v_1}^2)$, and u_t is i.i.d. $N(0, \sigma_u^2)$. However, the i.i.d. assumption of v_{1t} as well as the assumption of strictly exogenous regressors are unrealistic. Therefore, we will consider the asymptotic distribution of the statistics under weaker assumptions in text.

3. The bias is also due to Ω_{21}, since when $\Omega_{21} \neq 0$, $\int_0^1 B_2 \, dB_1$ is skewed and $\left(\int_0^1 B_2 B_2'\right)^{-1} \times \left(\int_0^1 B_2 \, dB_1\right)$ is not mean zero, which is also pointed out by the referee.

4. Following Phillips and Hansen [25], first transform y_t into $y_t^+ = y_t - \hat{\Omega}_{21}'\hat{\Omega}_{22}^{-1}\Delta Z_t$, and run the regression: $y_t^+ = Z_t'\hat{\beta}_{FM} + \hat{\varepsilon}_t^+$, where $\hat{\beta}_{FM}$ is the fully modified estimator. See [25] for further details and notations. Then, we obtain the residuals of this regression and the consistent estimate of the long-run variance of the residuals to construct the statistics as given in (6). As mentioned, the statistic of the null hypothesis of cointegration using the fully modified procedure should have the same limiting distribution given in Theorem 2.

REFERENCES

1. Andrews, D.W.K. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* 59 (1991): 817–858.
2. Andrews, D.W.K. & J.C. Monahan. An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. *Econometrica* 60 (1992): 953–966.
3. Berk, K.N. Consistent autoregressive spectral estimates. *Annals of Statistics* 2 (1974): 489–502.
4. Campbell, J.Y. & P. Perron. Pitfalls and opportunities: What macroeconomists should know about unit roots. Unpublished manuscript prepared for NBER Macroeconomic Conference, 1991.
5. Engle, R.F. & C.W.J. Granger. Cointegration and error correction representation: Estimation and testing. *Econometrica* 55 (1987): 251–276.
6. Granger, C.W.J. & T.-H. Lee. Multicointegration. *Advances in Econometrics* 8 (1990): 71–84.
7. Hansen, B. Tests for parameter instability in regressions with I(1) processes. *Journal of Business and Economic Statistics* 10 (1992): 321–335.
8. Hansen, B. Consistent covariance matrix estimation for dependent heterogeneous processes. *Econometrica* 60 (1992): 967–972.
9. Hansen, B. Convergence to stochastic integrals for dependent heterogeneous processes. *Econometric Theory* 8 (1992): 489–500.
10. Johansen, S. Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica* 59 (1991): 1551–1580.
11. Kwiatkowski, D., P.C.B. Phillips, P. Schmidt & 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? *Journal of Econometrics* 54 (1992): 159–178.

12. Kwiatkowski, D., P.C.B. Phillips & P. Schmidt. Testing for stationarity in the components representation of a time series. *Econometric Theory* 8 (1992): 586–591.
13. Newey, W.K. & K.D. West. A simple, positive semidefinite, heteroskedasticity and auto-correlated consistent covariance matrix. *Econometrica* 55 (1987): 703–708.
14. Nyblom, J. Testing for the constancy of parameters over time. *Journal of American Statistical Association* 84 (1989): 223–230.
15. Ogaki, M. & J.Y. Park. A cointegration approach to estimating preference parameters. Rochester Center for Economic Research Working Paper No. 209, University of Rochester, 1989.
16. Park, J.Y. Testing for unit roots by variable addition. In T.B. Fomby and G.F. Rhodes (eds.), *Advances in Econometrics: Cointegration, Spurious Regressions, and Unit Roots.* Greenwich: JAI Press, 1990.
17. Park, J.Y. Canonical cointegrating regressions. *Econometrica* 60 (1992): 119–143.
18. Park, J.Y., S. Ouliaris & B. Choi. A new approach to testing for a unit root. Unpublished manuscript. Cornell University, 1988.
19. Park, J.Y. & P.C.B. Phillips. Statistical inference in regressions with integrated processes: Part 1. *Econometric Theory* 4 (1988): 468–498.
20. Phillips, P.C.B. Understanding spurious regressions in econometrics. *Journal of Econometrics* 33 (1986): 311–340.
21. Phillips, P.C.B. Weak convergence of sample covariance matrices to stochastic integrals via martingale approximations. *Econometric Theory* 4 (1988): 528–533.
22. Phillips, P.C.B. Optimal inference in cointegrated systems. *Econometrica* 59 (1991): 283–306.
23. Phillips, P.C.B. Spectral regression for cointegrated time series. In W. Barnett (ed.), *Nonparametric and Semiparametric Methods in Economics and Statistics.* New York: Cambridge University Press, 1991.
24. Phillips, P.C.B. & S.N. Durlauf. Multiple time series regression with integrated processes. *Review of Economic Studies* 53 (1986): 473–496.
25. Phillips, P.C.B. & B. Hansen. Statistical inference in instrumental variables regression with I(1) processes. *Review of Economic Studies* 57 (1990): 99–125.
26. Phillips, P.C.B. & M. Loretan. Estimating long run economic equilibria. *Review of Economic Studies* 58 (1991): 407–436.
27. Phillips, P.C.B. & S. Ouliaris. Asymptotic properties of residual based tests for cointegration. *Econometrica* 58 (1990): 165–193.
28. Phillips, P.C.B. & P. Perron. Testing for a unit root in time series regression. *Biometrika* 75 (1988): 335–346.
29. Phillips, P.C.B. & W. Ploberger. Posterior odds testing for a unit root with data-based model selection. Cowles Foundation Discussion Paper No. 1017, Yale University, 1992.
30. Press, W.H., B.P. Flannery, S.A. Teukolsky & W.T. Vetterling. *Numerical Recipes: The Art of Statistical Computing.* New York: Cambridge University Press, 1986.
31. Quintos, C.E. & P.C.B. Phillips. Parameter constancy in cointegrating regressions. Unpublished manuscript. Yale University, 1992.
32. Said, S.E. & D.A. Dickey. Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika* 71 (1984): 599–607.
33. Saikkonen, P. Asymptotically efficient estimation of cointegration regressions. *Econometric Theory* 7 (1991): 1–21.
34. Stock, J.H. & M.W. Watson. Testing for common trends. *Journal of American Statistical Association* 83 (1988): 1097–1107.
35. Stock, J.H. & M.W. Watson. A simple estimator of cointegrating vectors in higher order integrated systems. Forthcoming in *Econometrica*.
36. Tanaka, K. An alternative approach to the asymptotic theory of spurious regression, cointegration, and near cointegration. *Econometric Theory* 9 (1993): 36–61.
37. Wooldridge, J. Notes on regressions with difference stationary data. Unpublished manuscript, Massachusetts Institute of Technology, 1991.

APPENDIX

We sketch the proofs of lemma and theorems. For more technical details, see [11,19,25,26,33,35]. For notational convenience we denote $\int_0^1 B(r) \, dr$ as $\int B$ and $\sum_{t=1}^T Z_t$ (or $\sum_{t=K+1}^{T-K} Z_t$) as ΣZ_t.

Proof of Theorem 1. We transform (1), (2), and (3) in matrix form: $y_t = Z_t' b + X_t$, $y_t = Z_{\mu t}' b_\mu + X_t$, and $y_t = Z_{\tau t}' b_\tau + X_t$, where $Z_{\mu t} = (1, Z_t')'$ and $Z_{\tau t} = (1, t, Z_t')'$. Define the scale matrices: $D \equiv T^{-1} I_m$, $D_\mu \equiv \text{diag}(T^{-1/2}, T^{-1} I_m)$, and $D_\tau \equiv \text{diag}(T^{-1/2}, T^{-3/2}, T^{-1} I_m)$. Let $\hat{b} = \hat{\beta}$, $\hat{b}_\mu = (\hat{\alpha}_\mu, \hat{\beta}_\mu')'$, and $\hat{b}_\tau = (\hat{\alpha}_\tau, \hat{\delta}_\tau, \hat{\beta}_\tau')'$ be the OLS estimators, respectively. Using the preliminaries given on page 95 of main text, we can show

$$D^{-1}(\hat{b} - b) \to \left(\int B_2 B_2' \right)^{-1} \left(\int B_2 \, dB_1 + \Delta_{21} \right)$$

$$D_\mu^{-1}(\hat{b}_\mu - b_\mu) \to \left(\int \bar{B}_2 \bar{B}_2' \right)^{-1} \left(\begin{array}{c} \left(\int \bar{B}_2 \bar{B}_2' \right) B_1(1) - \int B_2' \left(\int B_2 \, dB_1 + \Delta_{21} \right) \\ \int \bar{B}_2 \, dB_1 + \Delta_{21} \end{array} \right)$$

$$D_\tau^{-1}(\hat{b}_\tau - b_\tau) \to \left(\int B_2^* B_2^{*\prime} \right)^{-1} \begin{bmatrix} \left(\int B_2^* B_2^{*\prime} \right) \left\{ 4 B_1(1) - 6 \int r \, dB_1 \right\} \\ + \left(-4 \int B_2' - 6 \int r B_2' \right) \left(\int B_2^* \, dB_1 + \Delta_{21} \right) \\ \\ \left(\int B_2^* B_2^{*\prime} \right) \left\{ -6 B_1(1) + 12 \int r \, dB_1 \right\} \\ + \left(6 \int B_2' - 12 \int r B_2' \right) \left(\int B_2^* \, dB_1 + \Delta_{21} \right) \\ \\ \int B_2^* \, dB_1 + \Delta_{21} \end{bmatrix}.$$

Using the above results, we derive the asymptotic results for the partial sum process of the OLS residuals. Note that $X_t = v_{1t}$ under the null. Then, for equation (1),

$$T^{-1/2} S_{[Tr]} = T^{-1/2} \sum_{j=1}^{[Tr]} \hat{X}_j = T^{-1/2} \sum_{j=1}^{[Tr]} v_{1j} - T^{-3/2} \sum_{j=1}^{[Tr]} Z_j' T(\hat{\beta} - \beta) \to Q_B$$

$$\equiv B_1 - \left(\int_0^r B_2' \right) \left(\int B_2 B_2' \right)^{-1} \left(\int B_2 \, dB_1 + \Delta_{21} \right).$$

To develop a limiting distribution of our cointegration test statistic, we need a consistent estimator of the long-run variance of v_{1t}. We can use any heteroskedasticity and autocorrelated consistent covariance estimator, which is generally estimated by

$$s^2(\ell) \equiv T^{-1}\Sigma \hat{X}_t^2 + \left(\frac{2}{T}\right)\sum_{s=1}^{\ell} w(s,\ell) \sum_{t=s+1}^{T} \hat{X}_t \hat{X}_{t-s},$$

where $w(s,\ell)$ is a real-valued kernel and ℓ is a bandwidth parameter. See [1,2,8,13,28] for a discussion of possible estimators. Especially, Hansen [8] has proved (in his Theorems 2 and 3) that $s^2(\ell) \to \omega_{11}$ under general regularity conditions.

We prove main results. Since Ω_{21} is assumed to be zero, B_1 and B_2 are independent. Then,

$$Q_B = \omega_{11}^{1/2} W_1 - \left(\int_0^r W_2'\right)\Omega_{22}^{1/2}\Omega_{22}^{-1/2}\left(\int W_2 W_2'\right)^{-1}\Omega_{22}^{-1/2}\Omega_{22}^{1/2}\omega_{11}^{1/2}\left(\int W_2\, dW_1\right)$$

$$= \omega_{11}^{1/2} Q.$$

Therefore, combining the above results, we get

$$CI \equiv T^{-2}\Sigma S_t^2/s^2(\ell) \to \omega_{11} \int \frac{Q^2}{\omega_{11}} = \int Q^2.$$

Similarly, we can show for (2) and (3)

$$T^{-1/2}S_{\mu[\mathrm{Tr}]} = T^{-1/2}\sum_{j=1}^{[\mathrm{Tr}]} v_{1j} - \left(\frac{[\mathrm{Tr}]}{T}\right)T^{1/2}(\hat{\alpha}_\mu - \alpha_\mu) - T^{-3/2}\sum_{j=1}^{[\mathrm{Tr}]} Z_j' T(\hat{\beta}_\mu - \beta_\mu) \to Q_{B\mu}$$

and

$$T^{-1/2}S_{\tau[\mathrm{Tr}]} = T^{-1/2}\sum_{j=1}^{[\mathrm{Tr}]} v_{1j} - \left(\frac{[\mathrm{Tr}]}{T}\right)T^{1/2}(\hat{\alpha}_\tau - \alpha_\tau) - T^{-2}\sum_{j=1}^{[\mathrm{Tr}]} jT^{3/2}(\hat{\delta}_\tau - \delta_\tau)$$

$$- T^{-3/2}\sum_{j=1}^{[\mathrm{Tr}]} Z_j' T(\hat{\beta}_\tau - \beta_\tau) \to Q_{B\tau},$$

where

$$Q_{B\mu} \equiv B_1 - rB_1(1) - \left(\int_0^r \bar{B}_2'\right)\left(\int \bar{B}_2 \bar{B}_2'\right)^{-1}\left(\int \bar{B}_2\, dB_1 + \Delta_{21}\right)$$

and

$$Q_{B\tau} \equiv B_1 - (2r - 3r^2)B_1(1) + (-6r + 6r^2)\int B_1$$

$$- \left(\int_0^r B_2^{*\prime}\right)\left(\int B_2^* B_2^{*\prime}\right)^{-1}\left(\int B_2^*\, dB_1 + \Delta_{21}\right).$$

Note that the result of consistency is also valid for $s_\mu^2(\ell)$ and $s_\tau^2(\ell)$. See Theorem 3 in Hansen [8]. Therefore, combining these results with the assumption of strict exogeneity, we get the results for CI_μ and CI_τ statistics. ∎

Proof of Lemma 1. We transform (10), (11), and (12) into matrix form: $y_t = Z_t^{*\prime} b^* + \varepsilon_t^*$, $y_t = Z_{\mu t}^{*\prime} b_\mu^* + \varepsilon_t^*$, and $y_t = Z_{\tau t}^{*\prime} b_\tau^* + \varepsilon_t^*$, where $Z_t^* = (Z_t', \Delta Z_{t-K}', \ldots, \Delta Z_{t+K}')'$, $Z_{\mu t}^* = (1, Z_t', \Delta Z_{t-K}', \ldots, \Delta Z_{t+K}')'$, and $Z_{\tau t}^* = (1, t, Z_t', \Delta Z_{t-K}', \ldots, \Delta Z_{t+K}')'$. Let $\tilde{b} = (\tilde{\beta}', \tilde{\pi}')'$, $\tilde{b}_\mu = (\tilde{\alpha}_\mu, \tilde{\beta}_\mu', \tilde{\pi}_\mu')'$, and $\tilde{b}_\tau = (\tilde{\alpha}_\tau, \tilde{\delta}_\tau, \tilde{\beta}_\tau', \tilde{\pi}_\tau')'$ be the OLS estimators,

respectively. Define the scale matrices: $D^* \equiv \text{diag}(T^{-1}I_m, T^{-1/2}I_m, \ldots, T^{-1/2}I_m)$, $D_\mu^* \equiv \text{diag}(T^{-1/2}, T^{-1}I_m, T^{-1/2}I_m, \ldots, T^{-1/2}I_m)$, and $D_\tau^* \equiv \text{diag}(T^{-1/2}, T^{-3/2}, T^{-1}I_m, T^{-1/2}I_m, \ldots, T^{-1/2}I_m)$. Note that we now have the data from $K + 1$ to $T - K$. Now the number of observations are $T - 2K$, but we will use T instead of $T - 2k$ without loss of generality. Using (8) and (9), it can be shown that $\sum_{|j|>K}^{\infty} v'_{2t-j}\pi_j = o_p(T^{-1/2})$, which is also proved by Lemma A5 of [33]. Then, following the analysis of [3,32,33] (especially, see Lemma 5.1 of [32] and Lemma A4 of [33]), we can show that

$$D^{*-1}(\tilde{b} - b) = (D^*\Sigma Z_t^* Z_t^{*\prime} D^*)^{-1}(D^*\Sigma Z_t^* \varepsilon_t^*) \to R^{-1}(D^*\Sigma Z_t^* \varepsilon_t),$$

$$D_\mu^{*-1}(\tilde{b}_\mu - b_\mu) = (D_\mu^*\Sigma Z_{\mu t}^* Z_{\mu t}^{*\prime} D_\mu)^{-1}(D_\mu^*\Sigma Z_{\mu t}^* \varepsilon_t^*) \to R_\mu^{-1}(D_\mu^*\Sigma Z_{\mu t}^* \varepsilon_t)$$

and

$$D_\tau^{*-1}(\tilde{b}_\tau - b_\tau) = (D_\tau^*\Sigma Z_{\tau t}^* Z_{\tau t}^{*\prime} D_\tau^*)^{-1}(D_\tau^*\Sigma Z_{\tau t}^* \varepsilon_t^*) \to R_\tau^{-1}(D_\tau^*\Sigma Z_{\tau t}^* \varepsilon_t),$$

where

$$R = \text{diag}\{T^{-2}\Sigma Z_t Z_t', E(U_t U_t')\} \text{ with } U_t = (\Delta Z_{t-K}', \ldots, \Delta Z_{t+K}')'$$

and

$$R_\mu = \begin{bmatrix} 1 & T^{-3/2}\Sigma Z_t' & 0 \\ T^{-3/2}\Sigma Z_t & T^{-2}\Sigma Z_t Z_t' & 0 \\ 0 & 0 & E(U_t U_t') \end{bmatrix}$$

and

$$R_\tau = \begin{bmatrix} 1 & T^{-2}\Sigma t & T^{-3/2}\Sigma Z_t' & 0 \\ T^{-2}\Sigma t & T^{-3}\Sigma t^2 & T^{-5/2}\Sigma t Z_t' & 0 \\ T^{-3/2}\Sigma Z_t & T^{-5/2}\Sigma t Z_t & T^{-2}\Sigma Z_t Z_t' & 0 \\ 0 & 0 & 0 & E(U_t U_t') \end{bmatrix}.$$

After solving and rearranging, we have the asymptotic results of Lemma 1. (Note that we need an additional assumption that v_{2t} has finite fourth moments. See also conditions given in Theorem 2 in [3].) The order in probability for $\sum_{i=-K}^{K}(\tilde{\pi}_j - \pi_j)$ is given in the appendix of Saikkonen [33]. ∎

Proof of Theorem 2. Using Lemma 1 and following the analysis of [3,32,33] again, we can show for (10) that

$$T^{-1/2}\tilde{S}_{[Tr]} = T^{-1/2}\sum_{j=1}^{[Tr]}\varepsilon_j^* - T^{-1/2}\sum_{j=1}^{[Tr]}Z_j'(\tilde{\beta} - \beta) - T^{-1/2}\sum_{j=1}^{[Tr]}\sum_{i=-K}^{K}\Delta Z_{j-i}'(\tilde{\pi}_i - \pi_i)$$

$$= T^{-1/2}\sum_{j=1}^{[Tr]}\varepsilon_j - T^{-1/2}\sum_{j=1}^{[Tr]}\left(\sum_{|i|>K}^{\infty}v_{2t-i}'\pi_i\right)$$

$$- T^{-3/2}\sum_{j=1}^{[Tr]}Z_j'T(\tilde{\beta} - \beta) - T^{-1/2}\sum_{j=1}^{[Tr]}\sum_{i=-K}^{K}\Delta Z_{j-1}'(\tilde{\pi}_i - \pi_i)$$

The first component converges weakly to $B_{1\cdot2}$, and the third to $\left(\int_0^r B_2'\right)\left(\int_0^1 B_2 B_2'\right)^{-1} \times$ $\left(\int_0^1 B_2\, dB_{1\cdot2}\right)$ by Lemma 1, so what is needed to be shown is that the final two terms converge in probability to zero, uniformly in r. Indeed,

$$E\sup_{r\le1}\left|T^{-1/2}\sum_{j=1}^{[\mathrm{Tr}]}\left(\sum_{|i|>K}v_{2t-i}'\pi_i\right)\right| \le E\sup_{r\le1}T^{-1/2}\sum_{j=1}^{[\mathrm{Tr}]}\sum_{|i|>K}|v_{2t-i}'|\,|\pi_i|$$

$$= T^{-1/2}\sum_{j=1}^{[\mathrm{Tr}]}\sum_{|i|>K}E|v_{2t-i}'|\,|\pi_i|$$

$$\le \sup_t E|v_{2t}|T^{1/2}\sum_{|i|>K}^\infty|\pi_i| \to 0,$$

by (9), and hence

$$\sup_{r\le1}\left|T^{-1/2}\sum_{j=1}^{[\mathrm{Tr}]}\left(\sum_{|i|>K}v_{2t-i}'\pi_i\right)\right| \to 0,$$

by Markov's inequality. In addition,

$$\sup_{r\le1}\left|T^{-1/2}\sum_{j=1}^{[\mathrm{Tr}]}\sum_{i=-K}^K\Delta Z_{j-i}'(\tilde\pi_i-\pi_i)\right| = \sup_{r\le1}\left|T^{-1/2}\sum_{i=-K}^K\sum_{j=1}^{[\mathrm{Tr}]}\Delta Z_{j-i}'(\tilde\pi_i-\pi_i)\right|$$

$$= \sup_{r\le1}\left|T^{-1/2}\sum_{i=-K}^K(Z_{[\mathrm{tr}]-i}-Z_{1-i})'(\tilde\pi_i-\pi_i)\right|$$

$$\le 2\cdot T^{-1/2}\sum_{t\le T}^{\sup}|Z_t'|\left|\sum_{i=-K}^K(\tilde\pi_i-\pi_i)\right|$$

$$= O_p(1)O_p(K^{1/2}/T^{1/2}) = o_p(1).$$

Therefore, we can show

$$T^{-1/2}\tilde S_{[\mathrm{Tr}]} \to Q_{1\cdot2} \equiv B_{1\cdot2}-\left(\int_0^r B_2'\right)\left(\int_0^1 B_2 B_2'\right)^{-1}\left(\int_0^1 B_2\, dB_{1\cdot2}\right).$$

Next, since $B_{1\cdot2}\equiv\omega_{1\cdot2}^{1/2}W_1$, and $B_{1\cdot2}$ and B_2 are independent,

$$Q_{1\cdot2} = \omega_{1\cdot2}^{1/2}W_1-\left(\int_0^r W_2'\right)\Omega_{22}^{1/2}\Omega_{22}^{-1/2}\left(\int W_2 W_2'\right)^{-1}\cdot\Omega_{22}^{-1/2}\Omega_{22}^{1/2}\omega_{1\cdot2}^{1/2}(W_2\, dW_1)$$

$$= \omega_{1\cdot2}^{1/2}Q.$$

Now the long-run variance of the residual is estimated by $\tilde s^2(\ell) = T^{-1}\Sigma\tilde\varepsilon_t^{*2} + (2/T)\sum_{s=1}^\ell w(s,\ell)\sum_{t=s+1}^T\tilde\varepsilon_t^*\tilde\varepsilon_{t-s}^*$. Using the fact that $\sum_{|j|>K}^\infty v_{2t-j}'\pi_j = O_p(T^{-1/2})$ andTheorem 3 in [7], we can also show that $\tilde s^2(\ell)$ is the consistent estimator of the long run variance of $\varepsilon_t, \omega_{1\cdot2}$. Therefore,

$$C \equiv T^{-2}\Sigma\tilde S_t^2/\tilde s^2(\ell) \to \omega_{1\cdot2}\int Q^2/\omega_{1\cdot2} = \int Q^2.$$

Similarly, we can show for (11) and (12)

$$T^{-1/2}S_{\mu[Tr]} \to Q_{\mu1\cdot2} \equiv B_{1\cdot2} - rB_{1\cdot2}(1) - \left(\int_0^r \bar{B}_2'\right)\left(\int_0^1 \bar{B}_2\bar{B}_2'\right)^{-1}\left(\int_0^1 \bar{B}_2\, dB_{1\cdot2}\right)$$

and

$$T^{-1/2}S_{\tau[Tr]} \to Q_{\tau1\cdot2} \equiv B_{1\cdot2} + (2r - 3r^2)B_{1\cdot2}(1) + (-6r + 6r^2)$$

$$\times \int_0^1 B_{1\cdot2} - \left(\int_0^r B_2^{\star\prime}\right)\left(\int_0^1 B_2^{\star} B_2^{\star\prime}\right)^{-1}\left(\int_0^1 B_2^{\star}\, dB_{1\cdot2}\right).$$

Therefore, combining these results with the consistency of $\tilde{s}_{\mu}^2(\ell)$ and $\tilde{s}_{\tau}^2(\ell)$, we can show that C_{μ} and C_{τ} have the same limiting distributions as those of CI_{μ} and CI_{τ}. ∎

Proof of Theorem 3. To save space we consider the standard case only. The proofs are basically the same with more calculus for the demeaned and detrended cases. See also Park and Phillips [19] and Phillips and Ouliaris [27]. Under the alternative of no cointegration (that is, $\sigma_u^2 > 0$), $X_t = \sum_{j=1}^t u_j + \varepsilon_t$. Then, X_t is $I(1)$ so that $T^{-1/2}X_{[Tr]} = T^{-1/2}\sum_{j=1}^{[Tr]} u_j + o_p(1) \to B_u(r) = \sigma_u W_1$. Now, let $\tilde{\beta}$ and $\tilde{\pi}_j$ be the OLS estimates obtained from (10) under the alternative hypothesis. (Other notations are defined similarly.) Then, we can show

$$(\tilde{\beta} - \beta) \to \left(\int B_2 B_2'\right)^{-1}\int B_2 B_u, \quad \text{and} \quad \sum_{i=-K}^{K} (\tilde{\pi}_j - \pi_j) = O_P(K^{1/2}).$$

Let $\tilde{X}_t$ be the residuals obtained from (10) under the alternative hypothesis of no cointegration. Then,

$$T^{-3/2}\tilde{S}_{[Ta]} = T^{-3/2}\sum_{j=1}^{[Ta]} X_j - T^{3/2}\sum_{j=1}^{[Ta]} Z_j'(\tilde{\beta} - \beta) - T^{-3/2}\sum_{j=1}^{[Ta]}\sum_{i=-K}^{K} \Delta Z_{j-i}'(\tilde{\pi}_i - \pi_i).$$

The first component converges weakly to $\int_0^a B_u$, and the second to $\left(\int_0^a B_2'\right)\left(\int B_2 B_2'\right)^{-1}\times$ $\int B_2 B_u$, because B_u is independent of B_2 by construction. So, we need to show that the last term converges in probability to zero, uniformly in r. Indeed,

$$\sup_{a\le1}\left|T^{-3/2}\sum_{j=1}^{[Ta]}\sum_{i=-K}^{K} \Delta Z_{j-i}'(\tilde{\pi}_i - \pi_i)\right| = \sup_{a\le1}\left|T^{-3/2}\sum_{i=-K}^{K}\sum_{j=1}^{[Ta]} \Delta Z_{j-i}'(\tilde{\pi}_i - \pi_i)\right|$$

$$= \sup_{a\le1}\left|T^{-3/2}\sum_{i=-K}^{K} (Z_{[Ta]-i} - Z_{1-i})'(\tilde{\pi}_i - \pi_i)\right|$$

$$\le 2\cdot T^{-1/2}\sup_{t\le T}|Z_t'|T^{-1}\left|\sum_{i=-K}^{K} (\tilde{\pi}_i - \pi_i)\right|$$

$$= O_p(1)O_p\left(\frac{K^{1/2}}{T}\right) = o_p(1).$$

Therefore,

$$T^{-3/2}\tilde{S}_{[Ta]} \to \int_0^a B_u - \left(\int_0^a B_2'\right)\left(\int B_2 B_2'\right)^{-1}\int B_2 B_u = \sigma_u\int_0^a Q_{PO}$$

and

$$T^{-4}\Sigma\tilde{S}_t^2 = T^{-1}\Sigma(T^{-3/2}\tilde{S}_t)^2 \to \sigma_u^2 \int \left(\int_0^a Q_{PO}\right)^2.$$

From KPSS [11], we obtain the result that $(\ell T)^{-1}\tilde{s}^2(\ell) \to L\sigma_u^2 \int Q_{PO}^2$. See also the appendix of [23]. Note that $w(s,\ell) = k(s/\ell)$. For example, for the Bartlett window, $k(s) = 1 - |s|$ and $L = 1$. Therefore, combining the above results we obtain the result

$$(\ell/T)C \equiv T^{-4}\Sigma\tilde{S}_t^2/(\ell T)^{-1}\tilde{s}^2(\ell) \to \sigma_u^2 \int \left(\int_0^a Q_{PO}\right)^2 \bigg/ L\sigma_u^2 \int Q_{PO}^2$$

$$= \int \left(y \int_0^a Q_{PO}\right)^2 \bigg/ L \int Q_{PO}^2. \qquad \blacksquare$$