CHAPTER 15

Tests of the Null Hypothesis of Cointegration Based on Efficient Tests for a Unit MA Root

Michael Jansson

ABSTRACT

A new family of tests of the null hypothesis of cointegration is proposed. Each member of this family is a plug-in version of a point optimal stationarity test. Appropriately selected tests dominate existing cointegration tests in terms of local asymptotic power.

1. INTRODUCTION

In recent years, several papers have studied the problem of testing the null hypothesis of cointegration against the alternative of no cointegration. A variety of testing procedures have been proposed, but very little is known about the asymptotic power properties of these tests. In an attempt to shed some light on the issue of power, this chapter makes two contributions.

First, a new test of the null hypothesis of cointegration is introduced. Similar to the tests proposed by Park (1990), Shin (1994), Choi and Ahn (1995), and Xiao and Phillips (2002), the test developed in this chapter can be viewed as an extension of an existing test of the null hypothesis of stationarity. Unlike the tests introduced in the cited studies, the test proposed herein is based on a stationarity test (derived in Rothenberg (2000)), which is known to enjoy nearly optimal local asymptotic power properties.

Second, the paper compares the power of the new test to the power of previously proposed tests by numerical evaluation of the local asymptotic power functions. It turns out that a cointegration test based on an optimal stationarity test inherits the good (relative to competing test procedures) local asymptotic power properties of the stationarity tests upon which it is based. In particular, the new test dominates existing tests in terms of local asymptotic power.

Section 2 motivates the testing procedure introduced in this paper. Section 3 presents the model and the assumptions under which the development of formal results will proceed. The new family of tests is introduced in Section 4. Section 5 investigates the asymptotic properties of the tests and two competing test

This paper draws on material in Chapter 2 of the author's Ph.D. dissertation at University of Aarhus, Denmark.

procedures. Finally, Section 6 offers a few concluding remarks, while mathematical derivations appear in three Appendices.

2. MOTIVATION

The leading special case of the testing problem considered in this chapter is the problem of testing the null hypothesis $\theta = 1$ against the alternative hypothesis $\theta < 1$ in the model

$$y_t = \beta' x_t + v_t, \quad t = 1, \dots, T, \tag{2.1}$$

where v_t and x_t are independent zero mean Gaussian time series (of dimensions 1 and k, respectively), $\Delta x_t \sim$ i.i.d. $\mathcal{N}(0, I_k)$ with initial condition $x_0 = 0$, and v_t is generated by the model

$$\Delta v_t = u_t^y - \theta u_{t-1}^y, \quad t = 2, \dots, T, \tag{2.2}$$

where Δ is the difference operator, $u_t^y \sim$ i.i.d. $\mathcal{N}(0, 1)$, and the initial condition is $v_1 = u_1^y$. The parameters $\beta \in \mathbb{R}^k$ and $\theta \in (-1, 1)$ are assumed to be unknown.

In the literature on stationarity testing, the model (2.2) of v_t is often referred to as the moving average model. A convenient feature of the moving average model is that the null hypothesis of stationarity can be formulated as a simple parametric restriction.[1] Indeed, v_t is stationary if and only if the moving average coefficient θ in (2.2) equals unity. (The "if" part is true because $v_t = u_t^y \sim$ i.i.d. $\mathcal{N}(0, 1)$ when $\theta = 1$, whereas the "only if" part follows from the fact v_t is an integrated process with a random walk-type nonstationarity whenever θ differs from unity.) By implication, the time series y_t and x_t are cointegrated (in the sense of Engle and Granger (1987)) if and only if $\theta = 1$.

If β was known, the null hypothesis of cointegration could be tested by applying a stationarity test to the observed series $v_t = y_t - \beta' x_t$. Studying the moving average model (2.2), Rothenberg (2000, Section 4) derived the family of point optimal (PO) tests of the null hypothesis $\theta = 1$.[2] The stationarity test derived in Rothenberg (2000) rejects for large value of

$$P_T(\bar{\lambda}) = \sum_{t=1}^{T} u_t^y(0)^2 - \sum_{t=1}^{T} u_t^y(\bar{\lambda})^2,$$

where $u_t^y(l) = \sum_{i=0}^{t-1}(1 - T^{-1}l)^i \Delta v_{t-i}$ (for $l \in \{0, \bar{\lambda}\}$ and $t \in \{1, \dots, T\}$), $v_0 = 0$, and $\bar{\lambda} > 0$ is some prespecified constant. The test based on $P_T(\bar{\lambda})$ is the PO test of $\theta = 1$ against the point alternative $\theta = 1 - T^{-1}\bar{\lambda}$ in the model (2.2).

[1] An alternative to the moving average model, which also parameterizes stationarity as a point, is the "local-level" unobserved components model. As discussed by Stock (1994), the two models are closely related. In fact, it can be shown that the two models give rise to identical Gaussian power envelopes for tests of the null hypothesis cointegration whenever a constant term is included in the model (Jansson 2005). For this reason, only the moving average model will be studied here.

[2] See also Saikkonen and Luukkonen (1993), who derived the family of PO location invariant tests of $\theta = 1$.

https://doi.org/10.1017/CBO9780511614491.016 Published online by Cambridge University Press

By implication, the test is also the PO test of $\theta = 1$ against $\theta = 1 - T^{-1}\bar{\lambda}$ in the model (2.1)–(2.2) when β is known and $\{x_t\}$ is independent of $\{v_t\}$.[3]

It follows from Rothenberg (2000) that the test based on $P_T(\bar{\lambda})$ is "nearly" optimal (has local asymptotic power function "close" to the Gaussian power envelope) if $\bar{\lambda}$ is chosen appropriately. In particular, such PO stationarity tests have better local asymptotic power properties than the stationarity tests by Park and Choi (1988), Kwiatkowski et al. (1992), Choi and Ahn (1998), and Xiao (2001), respectively.

When β is unknown (as is assumed here), it seems natural to test the null hypothesis of cointegration by using a plug-in approach in which a stationarity test is applied to an estimate of v_t. The cointegration tests proposed by Park (1990), Shin (1994), Choi and Ahn (1995), and Xiao and Phillips (2002) are all of the plug-in variety, being based on the stationarity tests proposed by Park and Choi (1988), Kwiatkowski et al. (1992), Choi and Ahn (1998), and Xiao (2001), respectively. This chapter explores the extent to which the superiority of Rothenberg's stationarity test (Rothenberg 2000) is inherited by a plug-in cointegration test based upon it. Specifically, it is explored whether a plug-in cointegration test based on Rothenberg's stationarity test dominates the tests by Park (1990), Shin (1994), Choi and Ahn (1995), and Xiao and Phillips (2002) in terms of local asymptotic power.

3. THE MODEL AND ASSUMPTIONS

The plug-in cointegration test based on Rothenberg's stationarity test (Rothenberg 2000) will be developed under the assumption that $z_t = (y_t, x_t')'$ is an observed $(k+1)$-vector time series (partitioned into a scalar y_t and a k-vector x_t) generated by

$$z_t = \mu_t^z + z_t^0, \quad t = 1, \ldots, T, \tag{3.1}$$

where μ_t^z is a deterministic component and z_t^0 is a zero mean stochastic component. Partitioning z_t^0 conformably with z_t as $z_t^0 = (y_t^0, x_t^{0\prime})'$, it is assumed that z_t^0 is generated by the potentially cointegrated system

$$y_t^0 = \beta' x_t^0 + v_t, \tag{3.2}$$

$$\Delta x_t^0 = u_t^x, \tag{3.3}$$

where v_t is an error process with initial condition $v_1 = u_1^y$ and generating mechanism

$$\Delta v_t = u_t^y - \theta u_{t-1}^y, \quad t = 2, \ldots, T. \tag{3.4}$$

[3] If $\{x_t\}$ and $\{v_t\}$ are not independent, more powerful tests can often be found. Jansson (2004) has developed PO tests under the assumption that β is known and $(u_t^y, \Delta x_t')'$ is Gaussian white noise. These tests are more powerful than the test based on $P_T(\bar{\lambda})$ whenever the correlation between u_t^y and Δx_t is nonzero, but the source of these power gains is not exploitable when β is unknown (as is assumed in this chapter).

In (3.2) – (3.4), $\beta \in \mathbb{R}^k$ and $\theta \in (-1, 1]$ are unknown parameters and $u_t = (u_t^y, u_t^{x\prime})'$ is a stationary process whose long-run variance covariance matrix

$$\Omega = \lim_{T\to\infty} T^{-1} \sum_{t=1}^{T}\sum_{s=1}^{T} E\left(u_t u_s'\right)$$

is assumed to be positive definite.

For concreteness, the deterministic component μ_t^z is assumed to be a pth order polynomial time trend:

$$\mu_t^z = \alpha_z' d_t, \quad t = 1, \ldots, T, \tag{3.5}$$

where $d_t = (1, \ldots, t^p)'$ and α_z is a $(p+1) \times m$ matrix of unknown parameters. The leading special cases of (3.5) are the constant mean ($p = 0$) and linear trend ($p = 1$) cases corresponding to $d_t = 1$ and $d_t = (1, t)'$, respectively.

In the development of distributional results, it will be assumed that

$$T^{-1/2} \sum_{t=1}^{\lfloor T\cdot \rfloor} u_t \to_d \Omega^{1/2} W(\cdot), \tag{3.6}$$

and

$$T^{-1} \sum_{t=2}^{T} \left(\sum_{s=1}^{t-1} u_s \right) u_t' \to_d \Omega^{1/2} \int_0^1 W(r)\, dW(r)' \Omega^{1/2\prime} + \Gamma', \tag{3.7}$$

where $\lfloor\cdot\rfloor$ denotes the integer part of the argument, $W(\cdot)$ is a Wiener process of dimension m, and

$$\Gamma = \lim_{T\to\infty} T^{-1} \sum_{t=2}^{T}\sum_{s=1}^{t-1} E\left(u_t u_s'\right)$$

is the one-sided long-run covariance matrix of u_t.

Similar to the model of Section 2, the model (3.1)–(3.7) enjoys the property that the null hypothesis of cointegration can be formulated as a simple parametric restriction. Indeed, the problem of testing the null hypothesis of cointegration against the alternative of no cointegration can once again be formulated as the problem of testing

$$H_0 : \theta = 1 \quad \text{versus} \quad H_1 : \theta < 1.$$

The model (3.1)–(3.7) generalizes (2.1) and (2.2) in several respects. The presence of the deterministic component μ_t^z in (3.1) relaxes the zero mean assumption of (2.1) and (2.2). Moreover, the high-level assumptions (3.6) and (3.7) on the latent errors u_t accommodate quite general forms of contemporaneous and serial correlation (and do not require normality). Indeed, the convergence results (3.6) and (3.7) hold (jointly) under a variety of weak dependence conditions on u_t. For instance, the following assumption suffices:

A1. $u_t = \sum_{i=0}^{\infty} C_i \varepsilon_{t-i}$, where $\{\varepsilon_t : t \in \mathbb{Z}\}$ is i.i.d. $(0, I_m)$, $\sum_{i=0}^{\infty} C_i$ has full rank, and $\sum_{i=1}^{\infty} i \|C_i\| < \infty$, where $\|\cdot\|$ is the Euclidean norm.

Under A1, the long-run covariance matrix of u_t is $\Omega = (\sum_{i=0}^{\infty} C_i)(\sum_{i=0}^{\infty} C_i)'$, a positive definite matrix. The assumption that Ω is positive definite is a standard, but important, regularity condition. It implies that x_t^0 is a non-cointegrated integrated process and rules out multicointegration (in the sense of Granger and Lee 1990) under the null hypothesis of cointegration.

4. A FAMILY OF COINTEGRATION TESTS

Conformably with z_t, partition α_z as $\alpha_z = (\alpha_y, \alpha_x)$. Defining $\alpha = \alpha_y - \alpha_x \beta$, the following relation can be obtained by combining (3.1), (3.2), and (3.5):

$$y_t = \alpha' d_t + \beta' x_t + v_t, \quad t = 1, \ldots, T. \tag{4.1}$$

The family of cointegration tests proposed herein is obtained by applying (a suitably modified version of) Rothenberg's stationarity test (Rothenberg 2000) to an estimate of the error term v_t in (4.1).

Suppose (4.1) is estimated by OLS:

$$y_t = \hat{\alpha}' d_t + \hat{\beta}' x_t + \hat{v}_t. \tag{4.2}$$

As it turns out, tests constructed by applying stationarity tests to $\hat{v}_t$ generally have limiting distributions with complicated nuisance parameter dependencies unless x_t satisfies a certain exogeneity condition.[4] In the case of the stationarity tests proposed by Park and Choi (1988), Kwiatkowski et al. (1992), Choi and Ahn (1998), and Xiao (2001), this problem can be circumvented by employing an asymptotically efficient (under H_0) estimation procedure when constructing a plug-in cointegration tests (for details, see Park (1990), Shin (1994), Choi and Ahn (1995), and Xiao and Phillips (2002)). These properties are shared by the PO stationarity test, implying that the plug-in versions of Rothenberg's stationarity tests (Rothenberg 2000) should employ asymptotically efficient (under H_0) estimators of α and β in the construction of estimates of v_t. For concreteness, it is assumed that Park's canonical cointegrating regression (CCR) (Park 1992) estimators of α and β are used. (A brief discussion of alternative estimation strategies is provided at the end of this section.)

To construct the CCR estimators, consistent (under H_0 and local alternatives) estimators of Ω and Γ are needed. Suppose Ω and Γ are estimated by kernel estimators of the form

$$\hat{\Omega} = T^{-1} \sum_{t=1}^{T} \sum_{s=1}^{T} k\left(\frac{|t-s|}{\hat{b}_T}\right) \hat{u}_t \hat{u}_s', \tag{4.3}$$

[4] Specifically, $\lim_{T\to\infty} T^{-1} \sum_{t=1}^{T} \sum_{s=1}^{T} E(u_t^y u_s^{x\prime})$ and $\lim_{T\to\infty} T^{-1} \sum_{t=1}^{T} \sum_{s=1}^{t} E(u_t^y u_s^{x\prime})$ must be zero if these nuisance parameter dependencies are to be avoided. That is, Ω must be block diagonal and Γ must be block upper triangular, where Ω and Γ are the matrices defined in Section 3.

and

$$\hat{\Gamma} = T^{-1} \sum_{t=2}^{T} \sum_{s=1}^{t-1} k\left(\frac{|t-s|}{\hat{b}_T}\right) \hat{u}_t \hat{u}_s', \tag{4.4}$$

where $k(\cdot)$ is a (measurable) kernel function, $\hat{b}_T$ is a sequence of (possibly random) bandwidth parameters, and $\hat{u}_t = (\hat{v}_t, \Delta\hat{x}_t^{0\prime})'$, where $\hat{v}_t$ are the OLS residuals from (4.2) and $\hat{x}_t^0$ are the OLS residuals from

$$x_t = \hat{\alpha}_x' d_t + \hat{x}_t^0. \tag{4.5}$$

The consistency requirement on $\hat{\Omega}$ and $\hat{\Gamma}$ is met under the following assumption on $k(\cdot)$ and $\hat{b}_T$.

A2. (i) $k(0) = 1$, $k(\cdot)$ is continuous at zero and $\bar{k}(0) + \int_0^\infty \bar{k}(r)\,dr < \infty$, where $\bar{k}(r) = \sup_{s \geq r} |k(s)|$ (for all $r \geq 0$).
(ii) $\hat{b}_T = \hat{a}_T b_T$, where $\hat{a}_T$ and b_T are positive, $\hat{a}_T + \hat{a}_T^{-1} = O_p(1)$, and, b_T is nonrandom with $b_T^{-1} + T^{-1/2} b_T = o(1)$.

Assumption A2 (i) is adapted from Jansson (2002) and is discussed there, while A2 (ii) is adapted from Andrews (1991).

Partition $\hat{\Gamma}$ and $\hat{\Omega}$ in conformity with $u_t = (u_t^y, u_t^{x\prime})'$ and let $\hat{\Gamma}_{x\cdot} = (\hat{\gamma}_{xy}, \hat{\Gamma}_{xx})$, $\hat{\omega}_{yy\cdot x} = \hat{\kappa}'\hat{\Omega}\hat{\kappa}$, and $\hat{\gamma}_{yy\cdot x} = \hat{\kappa}'\hat{\Gamma}\hat{\kappa}$, where $\hat{\kappa} = (1, -\hat{\omega}_{xy}'\hat{\Omega}_{xx}^{-1})'$. Let $\tilde{\alpha}$ and $\tilde{\beta}$ be the OLS estimators obtained from the multiple regression

$$y_t^\dagger = \tilde{\alpha}' d_t + \tilde{\beta}' x_t^\dagger + \tilde{v}_t, \tag{4.6}$$

where $y_t^\dagger = y_t - \hat{\omega}_{xy}'\hat{\Omega}_{xx}^{-1}\Delta\hat{x}_t^0 + \hat{\beta}'\hat{\Gamma}_{x\cdot}\hat{\Sigma}^{-1}\hat{u}_t$, $x_t^\dagger = x_t + \hat{\Gamma}_{x\cdot}\hat{\Sigma}^{-1}\hat{u}_t$, $\hat{\Sigma} = T^{-1}\sum_{t=1}^T \hat{u}_t\hat{u}_t'$, and $\hat{\beta}$ is the OLS estimator from (4.2). The estimators $\tilde{\alpha}$ and $\tilde{\beta}$ from (4.6) are Park's CCR estimators (Park 1992) of α and β. Under H_0, these estimators are asymptotically efficient (in the sense of Saikkonen (1991)). In addition, the behavior of suitably normalized partial sums involving the residuals $\tilde{v}_t$ is such that asymptotically pivotal (under H_0) test statistics can be constructed using these residuals.

Let $\tilde{v}_0 = 0$ and define $\tilde{u}_t^y(l) = \sum_{i=0}^{t-1}(1 - T^{-1}l)^i \Delta\tilde{v}_{t-i}$ (for $l \in \{0, \bar{\lambda}\}$ and $t \in \{1, \ldots, T\}$). The proposed test rejects H_0 for large values of

$$Q_T(\bar{\lambda}) = \frac{\sum_{t=1}^T \tilde{u}_t^y(0)^2 - \sum_{t=1}^T \tilde{u}_t^y(\bar{\lambda})^2 - 2\bar{\lambda}\hat{\gamma}_{yy\cdot x}}{\hat{\omega}_{yy\cdot x}}, \tag{4.7}$$

where $\bar{\lambda} > 0$ is a prespecified constant. (Guidance on the choice of $\bar{\lambda}$ will be provided in Section 5.)

In the numerator of $Q_T(\bar{\lambda})$, the term $\sum_{t=1}^T \tilde{u}_t^y(0)^2 - \sum_{t=1}^T \tilde{u}_t^y(\bar{\lambda})^2$ is a plug-in version of the test statistic $P_T(\bar{\lambda})$ of Section 2. The statistic $Q_T(\bar{\lambda})$ is a modified version of $P_T(\bar{\lambda})$ in which two nonparametric corrections are employed in order to produce a test statistic which is asymptotically pivotal under H_0. Specifically, the term $-2\bar{\lambda}\hat{\gamma}_{yy\cdot x}$ corrects $\sum_{t=1}^{T} \tilde{u}_t^y(0)^2 - \sum_{t=1}^{T} \tilde{u}_t^y(\bar{\lambda})^2$ for "serial correlation bias," while the denominator removes scale parameter dependencies from the limiting distribution of $Q_T(\bar{\lambda})$.

Remark. Lemma A.2 in Appendix A summarizes the properties of $\tilde{v}_t$ that are used in the derivation of the distributional result reported in Theorem 5.1 of Section 5. These properties are shared by the "fully modified" (Phillips and Hansen 1990) residual process

$$\check{v}_t = y_t - \hat{\omega}_{xy}'\hat{\Omega}_{xx}^{-1}\Delta\hat{x}_t^0 - \check{\alpha}'d_t - \check{\beta}'x_t,$$

where $\check{\alpha}$ and $\check{\beta}$ are asymptotically efficient estimators of α and β. As a consequence, the test can also be based on $\check{v}_t$. Likewise, the test can be based on the DOLS (Stock and Watson 1993) residuals $\ddot{v}_t$ from the regression

$$y_t = \ddot{\alpha}'d_t + \ddot{\beta}'x_t + \ddot{\gamma}(L)\Delta x_t + \ddot{v}_t,$$

where $\ddot{\gamma}(L)$ is a two-sided lag polynomial.

5. ASYMPTOTIC THEORY

Similar to the existing cointegration tests, the test based on $Q_T(\bar{\lambda})$ has nontrivial power against local alternatives of the form $1-\theta = O(T^{-1})$. This fact motivates the reparameterization $\theta = \theta_T = 1 - T^{-1}\lambda$, where λ is a non-negative constant. Under this reparameterization, the null and alternative hypotheses are $\lambda = 0$ and $\lambda > 0$, respectively. A similar reparameterization was implicitly employed in the definition of $Q_T(\bar{\lambda})$, which is a plug-in version of the optimal test against the alternative $\theta = 1 - T^{-1}\bar{\lambda}$. Theorem 5.1 characterizes the limiting distribution of $Q_T(\bar{\lambda})$ under H_0 and local alternatives.

Theorem 5.1. *Let z_t be generated by (3.1)–(3.5) and suppose A1–A2 hold. Moreover, suppose $\theta = \theta_T = 1 - T^{-1}\lambda$ for some $\lambda \geq 0$. Then*

$$Q_T(\bar{\lambda}) \to_d 2\bar{\lambda}\int_0^1 \tilde{U}_{\bar{\lambda}}^{\lambda}(r)\, d\tilde{U}^{\lambda}(r) - \bar{\lambda}^2\int_0^1 \tilde{U}_{\bar{\lambda}}^{\lambda}(r)^2 dr,$$

where $\tilde{U}_{\bar{\lambda}}^{\lambda}(r) = \int_0^r e^{-\bar{\lambda}(r-s)}\, d\tilde{U}^{\lambda}(s)$,

$$d\tilde{U}^{\lambda}(r) = dU^{\lambda}(r) - \left(\int_0^1 X(s)\, dU^{\lambda}(s)\right)' \times \left(\int_0^1 X(s)X(s)'\, ds\right)^{-1} X(r)\, dr,$$

$U^{\lambda}(r) = U(r) + \lambda\int_0^r U(s)ds$, $X(r) = \left(V(r)', 1, \ldots, r^p\right)'$ *and U and V are independent Wiener processes of dimensions* 1 *and* k, *respectively.*

Table 15.1. *Percentiles of $Q_T(\bar{\lambda})$*

	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$
Constant Mean						
$\bar{\lambda}$	10	12	14	16	18	20
90%	−4.19	−5.78	−7.24	−8.68	−10.14	−11.61
95%	−3.24	−4.82	−6.34	−7.74	−9.17	−10.64
97.5%	−2.33	−3.90	−5.46	−6.83	−8.23	−9.66
99%	−1.09	−2.74	−4.21	−5.62	−7.08	−8.63
Linear trend						
$\bar{\lambda}$	14	16	18	19	21	23
90%	−6.72	−8.25	−9.64	−10.74	−12.15	−13.55
95%	−5.70	−7.24	−8.57	−9.83	−11.20	−12.59
97.5%	−4.73	−6.27	−7.53	−8.84	−10.19	−11.60
99%	−3.50	−4.96	−6.26	−7.67	−8.98	−10.43

To implement the test, the analyst must specify an alternative $\theta = 1 - T^{-1}\bar{\lambda}$ against which good power is desired. The approach recommended here is to choose $\bar{\lambda}$ in such a way that the local asymptotic power against the alternative $\theta = 1 - T^{-1}\bar{\lambda}$ is approximately equal to 50% when the 5% test based on $Q_T(\bar{\lambda})$ is used. In related testing problems, a similar approach has been advocated by Elliott, Rothenberg, and Stock (1996), Stock (1994), and Rothenberg (2000). Table 15.1 tabulates the recommended values of $\bar{\lambda}$ for $k = 1, \ldots, 6$ regressors in the constant mean and linear trend) case and reports selected percentiles of the asymptotic null distributions of the corresponding $Q_T(\bar{\lambda})$ statistics.[5]

The local asymptotic power properties of the new test will be compared to those of the cointegration tests proposed by Xiao and Phillips (2002) and Shin (1994), respectively.[6] The cointegration test proposed by Xiao and Phillips (2002) rejects H_0 for large values of

$$R_T = \hat{\omega}_{yy\cdot x}^{-1/2} \ \max_{1\le t\le T} \left| T^{-1/2} \sum_{s=1}^{t} \tilde{u}_s^y(0) \right|, \tag{5.1}$$

whereas Shin's test (Shin 1994) rejects for large values of

$$S_T = \hat{\omega}_{yy\cdot x}^{-1} T^{-2} \sum_{t=1}^{T-1} \left[\sum_{s=1}^{t} \tilde{u}_s^y(0) \right]^2, \tag{5.2}$$

[5] The percentiles were computed by generating 20,000 draws from the discrete time approximation (based on 2,000 steps) to the limiting random variables.

[6] The local power results of Jansson and Haldrup (2002) indicate that none of the cointegration tests proposed by Park (1990) and Choi and Ahn (1995) are superior to the test by Shin (1994). Therefore, cointegration tests by Park (1990) and Choi and Ahn (1995) are not studied here.

where $\hat{\omega}_{yy\cdot x}$ and $\tilde{u}_t^y(0)$ are defined as in Section 4.[7] It is shown in Appendix B that $R_T \to_d \sup_{0\leq r\leq 1}|\tilde{U}^\lambda(r)|$ and $S_T \to_d \int_0^1 \tilde{U}^\lambda(r)^2\, dr$ under the assumptions of Theorem 5.1, where $\tilde{U}^\lambda(r) = \int_0^r d\tilde{U}^\lambda(s)$.

Figure 15.1(a) plots the local asymptotic power functions of the constant mean $Q_T(10)$, R_T, and S_T tests in the case where x_t is a scalar ($k = 1$).[8] The test based on $Q_T(10)$ dominates existing tests in terms of local asymptotic power whenever λ exceeds 5. Even for alternatives close to H_0, where S_T enjoys certain optimality properties (Harris and Inder 1994), the new test is very competitive in terms of power.

Figure 15.1(b) investigates the optimality properties of Q_T, (10) by plotting its local asymptotic power function against two benchmarks. For any alternative $\lambda > 0$, the level of the quasi-envelope plotted in Figure 15.1(b) is obtained by maximizing (over $\bar{\lambda} > 0$) the power of a cointegration test based on a member of the family $\{Q_T(\bar{\lambda}) : \bar{\lambda} > 0\}$ of test statistics proposed herein. As a consequence, the optimality of the choice $\bar{\lambda} = 10$ can be evaluated by comparing the power of $Q_T(10)$ to the quasi-envelope. The power of $Q_T(10)$ is almost indistinguishable from the quasi-envelope for values of λ between 8 and 16 and is reasonably close to the quasi-envelope for values of λ outside this range. By choosing $\bar{\lambda}$ smaller (greater) than 10, the difference between the power of $Q_T(\bar{\lambda})$ and the quasi-envelope can be decreased for small (large) values of λ at the expense of a greater gap for large (small) values of λ. Therefore, although $Q_T(10)$ fails to attain the quasi-envelope, no other value of $\bar{\lambda}$ delivers a test statistic $Q_T(\bar{\lambda})$ with uniformly better power properties.

The envelope plotted in Figure 15.1(b) is an upper bound on the local asymptotic power of (a class of cointegration tests that contains all) plug-in cointegration tests. That bound, developed in a follow-up paper (Jansson 2005), can be used to investigate the optimality properties of $\{Q_T(\bar{\lambda}) : \bar{\lambda} > 0\}$ within the class of tests that are invariant under transformations of the form $y_t \to y_t + a'd_t + b'x_t$, where $a \in \mathbb{R}^{p+1}$ and $b \in \mathbb{R}^k$. The presence of a visible difference between the quasi-envelope and the power envelope suggests that an even more powerful cointegration test might exist. A confirmation of that conjecture is provided in Jansson (2005), where a cointegration test (not of the plug-in variety) with nearly optimal local asymptotic power properties is developed.

[7] Strictly speaking, R_T and S_T are modifications of the test statistics proposed by Xiao and Phillips (2002) and Shin (1994). Unlike R_T and S_T, tests by Shin (1994) and Xiao and Phillips (2002) are not based on estimation procedure by Park (1992). Under the assumptions of Theorem 5.1, the difference between R_T and Xiao and Phillips's test statistic (Xiao and Phillips 2002) is asymptotically negligible, as is the difference between S_T and test statistic by Shin (1994).

[8] The power functions were obtained by generating 20,000 draws from the discrete time approximation (based on 2,000 steps) to the limiting distributions of the test statistics for selected values of λ.

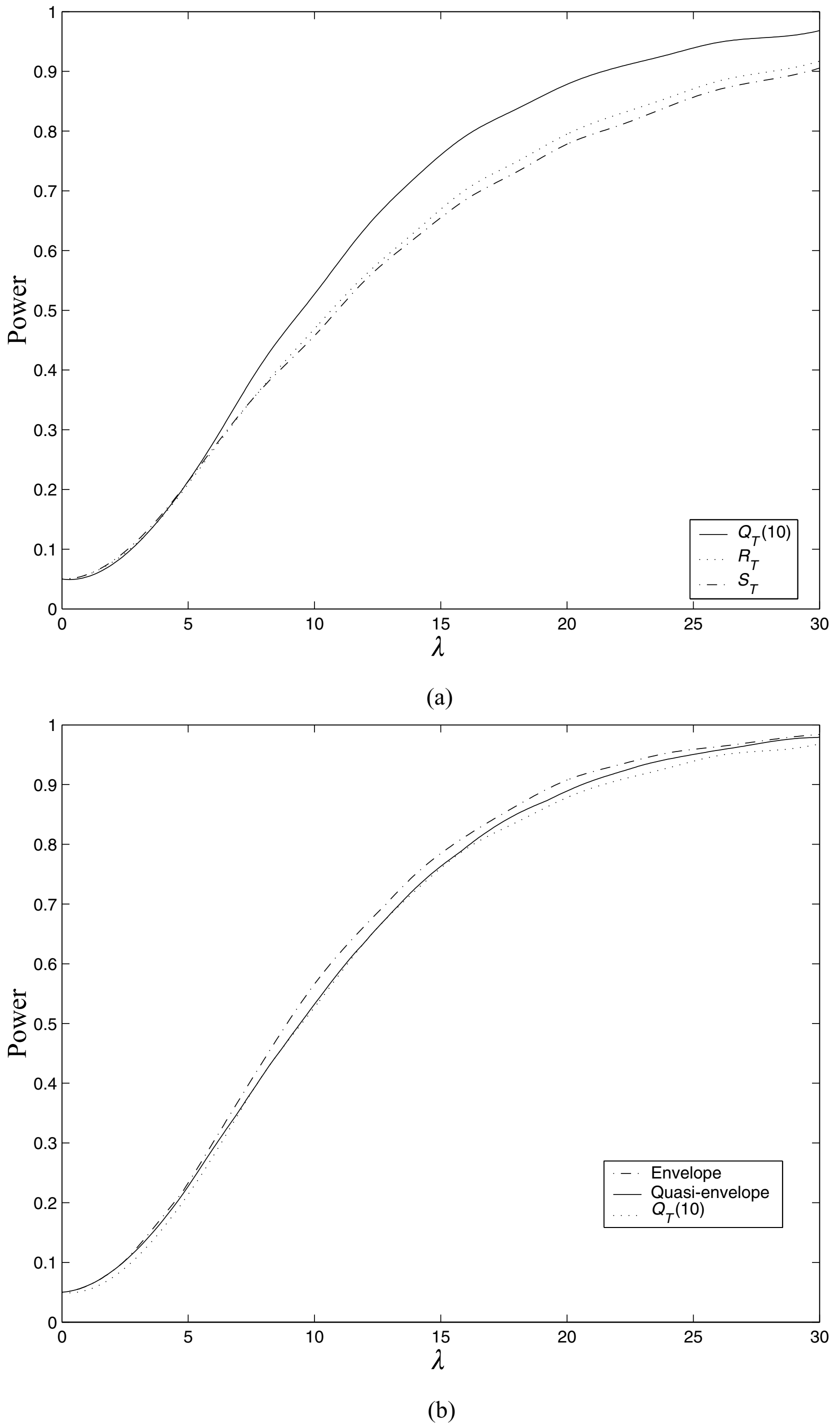


Figure 15.1. Power curves (5% level tests, constant mean, scalar x).

The result for the linear trend case are qualitatively similar to those for the constant mean case as can be seen from Figures 15.2(a) and 15.2(b).

The test statistic $Q_T(\bar{\lambda})$ has been constructed with local alternatives in mind. As the following theorem shows $Q_T(\bar{\lambda})$ can also be used to detect distant alternatives. Indeed, the test which rejects for large values of $Q_T(\bar{\lambda})$ is consistent in the sense that power against any fixed alternative $\theta = \bar{\theta} < 1$ tends to one as T increases without bound.

Theorem 5.2. *Let z_t be generated by (3.1)–(3.5) and suppose A1–A2 hold. Moreover, suppose $\theta < 1$ is fixed. Then $\lim_{T\to\infty} \Pr[Q_T(\bar{\lambda}) > c] = 1$ for any $c \in \mathbb{R}$.*

6. CONCLUSION

A new family of tests of the null hypothesis of cointegration was proposed. Each member of this family is a plug-in version of a PO stationarity test. Similar to the PO stationarity tests upon which they are based, the cointegration tests proposed in this chapter have good power properties. In particular, an appropriately selected version of the new test dominates existing cointegration tests in terms of local asymptotic power.

APPENDIX A: PROOF OF THEOREM 5.1

The proof of Theorem 5.1 utilizes the following two lemmas.

Lemma A.1. *Under the assumptions of Theorem 5.1, $\hat{\Omega} \to_p \Omega$ and $\hat{\Gamma} \to_p \Gamma$.*

Lemma A.2. *Under the assumptions of Theorem 5.1,*

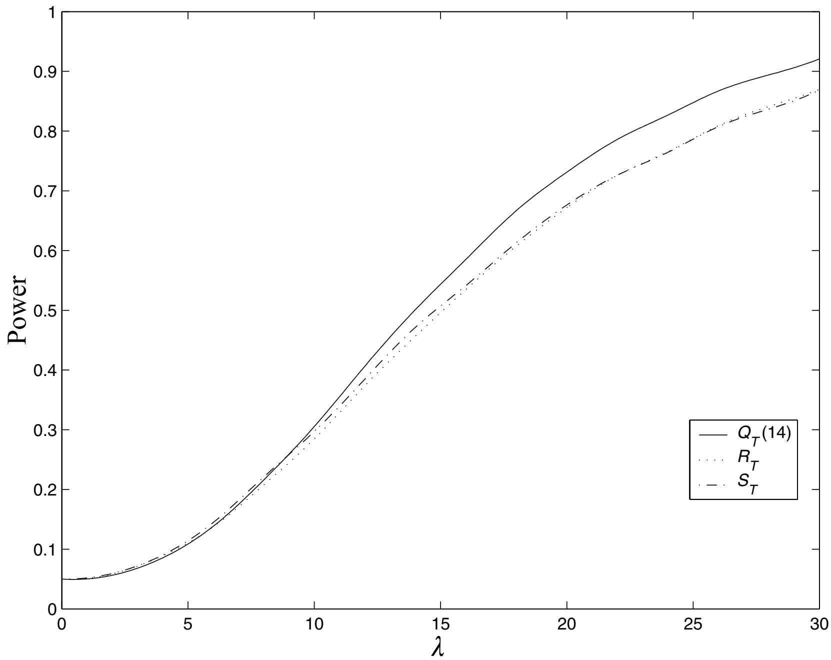
$$T^{-1/2} \sum_{t=1}^{\lfloor T\cdot \rfloor} \tilde{u}_t^y(0) \to_d \omega_{yy\cdot x}^{1/2} \tilde{U}^\lambda(\cdot)$$

and

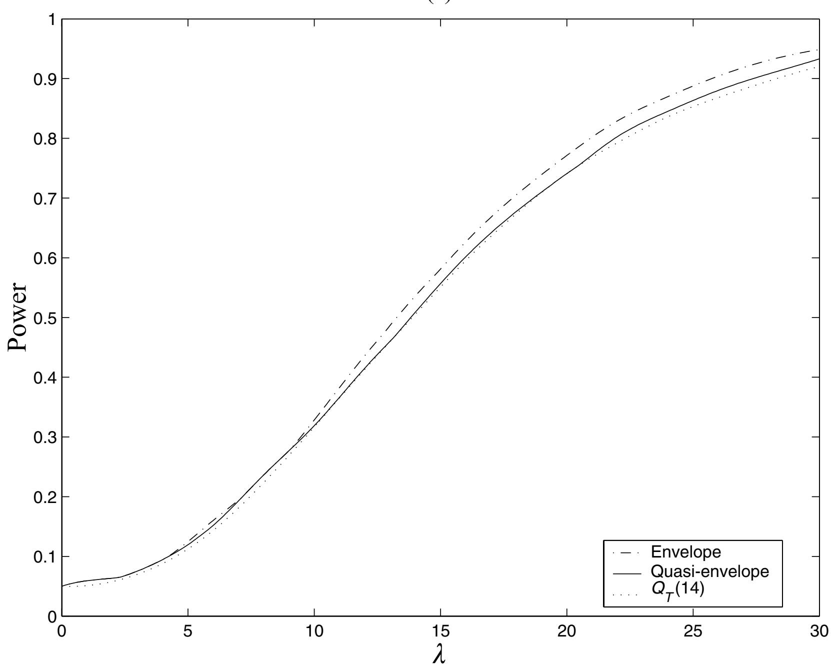
$$T^{-1} \sum_{t=2}^{T} \left[\sum_{s=1}^{t-1} \tilde{u}_s^y(0) \right] \tilde{u}_t^y(0) \to_d \omega_{yy\cdot x} \int_0^1 \tilde{U}^\lambda(r)\, d\tilde{U}^\lambda(r) + \gamma_{yy\cdot x}$$

jointly, where $\omega_{yy\cdot x} = \kappa'\Omega\kappa$, $\gamma_{yy\cdot x} = \kappa'\Gamma\kappa$, and $\kappa = (1, -\omega_{xy}'\Omega_{xx}^{-1})'$.

Under H_0, Lemma A.1 follows from Corollary 4 of Jansson (2002). The extension to local alternatives is straightforward, but tedious, and can be established by proceeding as in the proof of Lemma 5 of Jansson and Haldrup (2002). Lemma A.2 follows from Lemma 6(c)–(f) of Jansson and Haldrup (2002) and the fact that $\tilde{u}_t^y(0) = \tilde{v}_t$.



(a)



(b)

Figure 15.2. Power curves (5% level tests, linear trend, scalar x).

Proof of Theorem 5.1. By Lemma A.1, $\hat{\gamma}_{yy\cdot x} \to_p \gamma_{yy\cdot x}$ and $\hat{\omega}_{yy\cdot x} \to_p \omega_{yy\cdot x}$. Since

$$\begin{aligned}\sum_{t=1}^{T} \tilde{u}_t^y(0)^2 - \sum_{t=1}^{T} \tilde{u}_t^y(\bar{\lambda})^2 &= \sum_{t=1}^{T} \tilde{u}_t^y(0)^2 - \sum_{t=1}^{T} \left[\tilde{u}_t^y(0) + \tilde{u}_t^y(\bar{\lambda}) - \tilde{u}_t^y(0)\right]^2 \\ &= -\sum_{t=1}^{T} \left[\tilde{u}_t^y(0) - \tilde{u}_t^y(\bar{\lambda})\right]^2 + 2\sum_{t=1}^{T} \left[\tilde{u}_t^y(0) - \tilde{u}_t^y(\bar{\lambda})\right] \tilde{u}_t^y(0),\end{aligned}$$

the proof of Theorem 5.1 can therefore be completed by establishing the following convergence results:

$$\sum_{t=1}^{T} \left[\tilde{u}_t^y(0) - \tilde{u}_t^y(\bar{\lambda})\right]^2 \to_d \bar{\lambda}^2 \omega_{yy\cdot x} \int_0^1 \tilde{U}_{\bar{\lambda}}^{\lambda}(r)^2\, dr, \tag{A.1}$$

$$\sum_{t=1}^{T} \left[\tilde{u}_t^y(0) - \tilde{u}_t^y(\bar{\lambda})\right] \tilde{u}_t^y(0) \to_d \bar{\lambda} \left(\omega_{yy\cdot x} \int_0^1 \tilde{U}_{\bar{\lambda}}^{\lambda}(r)\, d\tilde{U}^{\lambda}(r) + \gamma_{yy\cdot x}\right). \tag{A.2}$$

Let $\bar{\theta}_T = 1 - T^{-1}\bar{\lambda}$. Using the relation $\tilde{u}_t^y(\bar{\lambda}) = \tilde{u}_t^y(0) - \bar{\lambda} T^{-1} \sum_{j=1}^{t-1} \bar{\theta}_T^{t-1-j} \tilde{u}_j^y(0)$ and summation by parts,

$$\tilde{u}_t^y(0) - \tilde{u}_t^y(\bar{\lambda}) = T^{-1}\bar{\lambda} \left(\tilde{U}_{t-1}^y - \bar{\lambda}\bar{\theta}_T^{t-2} T^{-1} \sum_{j=1}^{t-2} \bar{\theta}_T^{-j} \tilde{U}_j^y\right), \tag{A.3}$$

where $\tilde{U}_t^y = \sum_{j=1}^{t} \tilde{u}_j^y(0)$. Now, $T^{-1/2}\tilde{U}_{\lfloor T\cdot \rfloor}^y \to_d \omega_{yy\cdot x}^{1/2} \tilde{U}^{\lambda}(\cdot)$ by Lemma A.2. Moreover, $\lim_{T\to\infty} \sup_{0\le r\le 1} |\bar{\theta}_T^{\lfloor Tr \rfloor} - \exp(-\bar{\lambda} r)| = 0$, so

$$T^{1/2}\left[\tilde{u}_{\lfloor T\cdot \rfloor}^y(0) - \tilde{u}_{\lfloor T\cdot \rfloor}^y(\bar{\lambda})\right] \to_d \bar{\lambda}\omega_{yy\cdot x}^{1/2} \tilde{U}_{\bar{\lambda}}^{\lambda}(\cdot) \tag{A.4}$$

by the continuous mapping theorem (CMT), Theorem 13.4 of Billingsley (1999), and the fact that

$$\tilde{U}_{\bar{\lambda}}^{\lambda}(r) = \tilde{U}^{\lambda}(r) - \bar{\lambda} \int_0^r \exp\left[\bar{\lambda}(s-r)\right] \tilde{U}^{\lambda}(s)\, ds, \qquad r \in [0,1].$$

Using (A.4) and applying CMT,

$$\begin{aligned}\sum_{t=1}^{T} \left[\tilde{u}_t^y(0) - \tilde{u}_t^y(\bar{\lambda})\right]^2 &= \int_0^1 \left(T^{1/2}\left[\tilde{u}_{\lfloor Tr \rfloor}^y(0) - \tilde{u}_{\lfloor Tr \rfloor}^y(\bar{\lambda})\right]\right)^2 dr \\ &\quad + \left[\tilde{u}_T^y(0) - \tilde{u}_T^y(\bar{\lambda})\right]^2 \\ &\to_d \int_0^1 \left(\bar{\lambda}\omega_{yy\cdot x}^{1/2} \tilde{U}_{\bar{\lambda}}^{\lambda}(r)\right)^2 dr,\end{aligned}$$

establishing (A.1).

By (A.3),

$$\sum_{t=1}^{T}\left(\tilde{u}_t^y(0)-\tilde{u}_t^y(\bar{\lambda})\right)\tilde{u}_t^y(0)=T^{-1}\bar{\lambda}\sum_{t=1}^{T}\tilde{U}_{t-1}^y\tilde{u}_t^y(0)+T^{-1}\bar{\lambda}\sum_{t=1}^{T}$$
$$\times\left(-\bar{\lambda}\bar{\theta}_T^{t-2}T^{-1}\sum_{j=1}^{t-2}\bar{\theta}_T^{-j}\tilde{U}_j^y\right)\tilde{u}_t^y(0).$$

Now,

$$T^{-1}\sum_{t=1}^{T}\tilde{U}_{t-1}^y\tilde{u}_t^y(0)\rightarrow_d\omega_{yy\cdot x}\int_0^1\tilde{U}^\lambda(r)\,d\tilde{U}^\lambda(r)+\gamma_{yy\cdot x}$$

by Lemma A.2. Moreover,

$$T^{-1}\sum_{t=1}^{T}\left(-\bar{\lambda}\bar{\theta}_T^{t-2}T^{-1}\sum_{j=1}^{t-2}\bar{\theta}_T^{-j}\tilde{U}_j^y\right)\tilde{u}_t^y(0)$$
$$=T^{-1}\sum_{t=1}^{T}\left[\tilde{u}_{t-1}^y(0)-\tilde{u}_{t-1}^y(\bar{\lambda})\right]\tilde{U}_{t-1}^y\rightarrow_d\bar{\lambda}\omega_{yy\cdot x}\int_0^1\tilde{U}_{\bar{\lambda}}^\lambda(r)\tilde{U}^\lambda(r)\,dr,$$

where the equality uses summation by parts, (A.3), and $\tilde{U}_T^y=0$, while the last line uses (A.4), Lemma A.2, and CMT. Combining the preceding displays, the limiting distribution of $\sum_{t=1}^{T}(\tilde{u}_t^y(0)-\tilde{u}_t^y(\bar{\lambda}))\tilde{u}_t^y(0)$ can be represented as

$$\bar{\lambda}\left(\omega_{yy\cdot x}\left[\int_0^1\tilde{U}^\lambda(r)\,d\tilde{U}^\lambda(r)+\bar{\lambda}\int_0^1\tilde{U}_{\bar{\lambda}}^\lambda(r)\tilde{U}^\lambda(r)\,dr\right]+\gamma_{yy\cdot x}\right)$$
$$=\bar{\lambda}\left(\omega_{yy\cdot x}\int_0^1\tilde{U}_{\bar{\lambda}}^\lambda(r)d\tilde{U}^\lambda(r)+\gamma_{yy\cdot x}\right),$$

where the equality follows from integration by parts. Therefore, (A.2) holds and the proof is complete. ■

APPENDIX B: LIMITING DISTRIBUTIONS OF R_T AND S_T

The limiting distribution of R_T is derived as follows:

$$R_T=\hat{\omega}_{yy\cdot x}^{-1/2}\ \max_{1\le t\le T}\left|T^{-1/2}\sum_{s=1}^{t}\tilde{u}_s^y(0)\right|$$
$$=\left(\omega_{yy\cdot x}^{-1/2}+o_p(1)\right)\sup_{0\le r\le 1}\left|T^{-1/2}\sum_{s=1}^{\lfloor Tr\rfloor}\tilde{u}_s^y(0)\right|$$
$$\rightarrow_d\sup_{0\le r\le 1}\left|\tilde{U}^\lambda(r)\right|,$$

where the second equality uses Lemma A.1, while the last line uses Lemma A.2 and CMT.

Similarly,

$$
\begin{aligned}
S_T &= \hat{\omega}_{yy\cdot x}^{-1} T^{-2} \sum_{t=1}^{T-1} \left(\sum_{s=1}^{t} \tilde{u}_s^y(0) \right)^2 \\
&= \left[\omega_{yy\cdot x}^{-1} + o_p(1) \right] \left[\int_0^1 \left(T^{-1/2} \sum_{s=1}^{\lfloor Tr \rfloor} \tilde{u}_s^y(0) \right)^2 dr \right] \\
&\to_d \int_0^1 \tilde{U}^{\lambda}(r)^2 dr,
\end{aligned}
$$

where the second equality uses Lemma A.1, while the last line uses Lemma A.2 and CMT.

APPENDIX C: PROOF OF THEOREM 5.2

Let $\hat{u}_t^y(l) = \sum_{j=0}^{t-1} (1 - T^{-1}l)^j \Delta \hat{v}_{t-j}$ for $l \in \{0, \bar{\lambda}\}$ and $t \in \{1, \ldots, T\}$, where $\hat{v}_0 = 0$ and $\{\hat{v}_t\}$ are the residuals from (4.2). The following lemmas are used in the proof of Theorem 5.2.

Lemma C.1. *Under the assumptions of Theorem 5.2,*

$$
T^{-1/2} \hat{v}_{\lfloor T\cdot \rfloor} \to_d (1-\theta) \omega_{yy\cdot x}^{1/2} \hat{U}(\cdot),
$$

where $\hat{U}(r) = U(r) - (\int_0^1 X(s)U(s))'(\int_0^1 X(s)X(s)'ds)^{-1} X(r)$, while U and X are defined as in Theorem 5.1.

Lemma C.2. *Under the assumptions of Theorem 5.2, $T^{-3/2}\hat{\gamma}_{yy} \to_p 0$, $T^{-1}\hat{\gamma}_{xy} \to_p 0$ and $T^{-1}\hat{\gamma}_{yx} \to_p 0$.*

Lemma C.3. *Under the assumptions of Theorem 5.2,*

$$
T^{-2} \left[\sum_{t=1}^{T} \tilde{u}_t^y(0)^2 - \sum_{t=1}^{T} \tilde{u}_t^y(\bar{\lambda})^2 \right] = T^{-2} \left[\sum_{t=1}^{T} \hat{u}_t^y(0)^2 - \sum_{t=1}^{T} \hat{u}_t^y(\bar{\lambda})^2 \right] + o_p(1).
$$

Lemma C.1 follows from standard spurious regression results. The proof of Lemma C.2 uses $T^{-1}\hat{\sigma}_{yy} = T^{-2}\sum_{t=1}^{T} \hat{v}_t^2 = O_p(1)$, $\hat{\Sigma}_{xx} = T^{-1}\sum_{t=1}^{T} \Delta \hat{x}_t^0 \Delta\, \hat{x}_t^{0\prime} = O_p(1)$, and the fact that $T^{-1/2}\sum_{i=0}^{T-1} |k(i/\hat{b}_T)| \to_p 0$ under A2

(Jansson 2002). For instance,

$$\begin{aligned}
\left|T^{-3/2}\hat{\gamma}_{yy}\right| &= \left|T^{-3/2}\sum_{i=0}^{T-1} k\left(\frac{i}{\hat{b}_T}\right)\left(T^{-1}\sum_{t=1}^{T-i}\hat{v}_{t+i}\hat{v}_t\right)\right| \\
&\leq T^{-1/2}\sum_{i=0}^{T-1}\left|k\left(\frac{i}{\hat{b}_T}\right)\right|\left|\left(T^{-2}\sum_{t=1}^{T-i}\hat{v}_{t+i}\hat{v}_t\right)\right| \\
&\leq T^{-1/2}\sum_{i=0}^{T-1}\left|k\left(\frac{i}{\hat{b}_T}\right)\right|\left(T^{-2}\sum_{t=1}^{T-i}\hat{v}_{t+i}^2\right)^{1/2}\left(T^{-2}\sum_{t=1}^{T-i}\hat{v}_t^2\right)^{1/2} \\
&\leq \left(T^{-2}\sum_{t=1}^{T}\hat{v}_t^2\right)\left(T^{-1/2}\sum_{i=0}^{T-1}\left|k\left(\frac{i}{\hat{b}_T}\right)\right|\right) \to_p 0,
\end{aligned}$$

where the second inequality uses the Cauchy–Schwarz inequality. Finally, the proof of Lemma C.3 uses $T^{-1}\hat{\gamma}_{yx} \to_p 0$ and a considerable amount of tedious algebra. To conserve space, the details are omitted.

Proof of Theorem 5.2. For any T,

$$\begin{aligned}
&Pr\left[Q_T(\bar{\lambda}) > c\right] \\
&\quad = Pr\left[\sum_{t=1}^{T}\tilde{u}_t^y(0)^2 - \sum_{t=1}^{T}\tilde{u}_t^y(\bar{\lambda})^2 - 2\bar{\lambda}\hat{\gamma}_{yy\cdot x} - c\hat{\omega}_{yy\cdot x} > 0\right].
\end{aligned}$$

By Lemmas C.2 and C.3 and using $T^{-1}\hat{\sigma}_{yy} = O_p(1)$,

$$\begin{aligned}
&T^{-2}\left[\sum_{t=1}^{T}\tilde{u}_t^y(0)^2 - \sum_{t=1}^{T}\tilde{u}_t^y(\bar{\lambda})^2 - 2\bar{\lambda}\hat{\gamma}_{yy\cdot x} - c\hat{\omega}_{yy\cdot x}\right] \\
&\quad = T^{-2}\left[\sum_{t=1}^{T}\hat{u}_t^y(0)^2 - \sum_{t=1}^{T}\hat{u}_t^y(\bar{\lambda})^2\right] + o_p(1).
\end{aligned}$$

In view of the portmanteau theorem (for example, Billingsley 1999), the proof of Theorem 5.2 can therefore be completed by showing that $T^{-2}[\sum_{t=1}^{T}\hat{u}_t^y(0)^2 - \sum_{t=1}^{T}\hat{u}_t^y(\bar{\lambda})^2]$ has a limiting distribution with positive support.

Let $\bar{\theta}_T = 1 - T^{-1}\bar{\lambda}$. The relation $\hat{u}_t^y(\bar{\lambda}) = \hat{u}_t^y(0) - \bar{\lambda}T^{-1}\sum_{j=1}^{t-1}\bar{\theta}_T^{t-1-j}\hat{u}_j^y(0)$ can be restated as follows:

$$\hat{u}_{T,\lfloor Tr\rfloor}^y(l) = \hat{u}_{T,\lfloor Tr\rfloor}^y - \bar{\lambda}\bar{\theta}_T^{\lfloor Tr\rfloor-1}\int_0^{\lfloor Tr\rfloor/T}\bar{\theta}_T^{-\lfloor Ts\rfloor}\hat{u}_{T,\lfloor Ts\rfloor}^y ds, \quad 0 \leq r \leq 1.$$

Now, $\lim_{T\to\infty}\sup_{0\leq r\leq 1}|\bar{\theta}_T^{\lfloor Tr\rfloor} - \exp(-\bar{\lambda}r)| = 0$, so it follows from the preceding display, Lemma C.1, and CMT that $T^{-1/2}\hat{u}_{\lfloor T\cdot\rfloor}^y(\bar{\lambda}) \to_d (1-\theta)\omega_{yy\cdot x}^{1/2}\hat{U}_{\bar{\lambda}}(\cdot)$, where

$$\hat{U}_{\bar{\lambda}}(r) = \hat{U}(r) - \bar{\lambda}\int_0^r \exp(-\bar{\lambda}(r-s))\hat{U}(s)ds.$$

Using this result, Lemma C.1, and CMT,

$$T^{-2}\left[\sum_{t=1}^{T}\hat{u}_t^y(0)^2 - \sum_{t=1}^{T}\hat{u}_t^y(\bar{\lambda})^2\right]$$
$$\to_d (1-\theta)^2\omega_{yy\cdot x}\left[\int_0^1 \hat{U}(r)^2dr - \int_0^1 \hat{U}_{\bar{\lambda}}(r)^2dr\right],$$

so it suffices to show that $\Pr[\int_0^1 \hat{U}(r)^2dr - \int_0^1 \hat{U}_{\bar{\lambda}}(r)^2dr > 0] = 1$.

Since $\hat{U}_{\bar{\lambda}}(r) = \hat{U}(r) - \bar{\lambda}\int_0^1 1(s \le r)\exp[-\bar{\lambda}(r-s)]\hat{U}(s)ds$, where $1(\cdot)$ is the indicator function, it follows from straightforward algebra that

$$\int_0^1 \hat{U}(r)^2dr - \int_0^1 \hat{U}_{\bar{\lambda}}(r)^2dr = \int_0^1\int_0^1 K_{\bar{\lambda}}(r,s)\hat{U}(r)\hat{U}(s)drds,$$

where

$$K_{\bar{\lambda}}(r,s) = \frac{\bar{\lambda}}{2}\left(\exp\left[-\bar{\lambda}(2-r-s)\right] + \exp\left[-\bar{\lambda}\,|r-s|\right]\right).$$

The desired result now follows from the fact that the function $K_{\bar{\lambda}}(\cdot,\cdot)$ is positive definite in the sense that $\int_0^1\int_0^1 K_{\bar{\lambda}}(r,s)f(r)f(s)drds > 0$ for any nonzero, continuous function $f(\cdot)$. ■

ACKNOWLEDGMENTS

Helpful comments from Jim Stock, two referees, and participants at the 2001 NSF Symposium on Identification and Inference for Econometric Models are gratefully acknowledged.

References

Andrews, D. W. K. (1991), "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59, 817–858.

Billingsley, P. (1999), *Convergence of Probability Measures*, 2nd edn. New York: Wiley.

Choi, I., and B. C. Ahn (1995), "Testing for Cointegration in a System of Equations," *Econometric Theory*, 11, 952–983.

Choi, I., and B. C. Ahn (1998), "Testing the Null of Stationarity for Multiple Time Series," *Journal of Econometrics*, 88, 41–77.

Elliott, G., T. J. Rothenberg, and J. H. Stock (1996), "Efficient Tests for an Autoregressive Unit Root," *Econometrica*, 64, 813–836.

Engle, R. F., and C. W. J. Granger (1987), "Cointegration and Error Correction: Representation, Estimation, and Testing," *Econometrica*, 55, 251–276.

Granger, C. W. J., and T. Lee (1990), "Multicointegration," in *Advances in Econometrics: Co-integration, Spurious Regressions, and Unit Roots* (ed. by T. B. Fomby and G. F. Rhodes, Jr.), Greenwich, CT: Jai Press, pp. 71–84.

Harris, D., and B. Inder (1994), "A Test of the Null Hypothesis of Cointegration," in *Nonstationary Time Series Analysis and Cointegration* (ed. by C. Hargreaves), Oxford: Oxford University Press, 133–152.

Jansson, M. (2002), "Consistent Covariance Matrix Estimation for Linear Processes," *Econometric Theory*, 18, 1449–1459.

Jansson, M. (2005), "Point Optimal Tests of the Null Hypothesis of Cointegration," *Journal of Econometrics* 124, 187–201.

Jansson, M. (2004), "Stationarity Testing with Covariates," *Econometric Theory*, 20, 56–94.

Jansson, M., and N. Haldrup (2002), "Regression Theory for Nearly Cointegrated Time Series," *Econometric Theory*, 18, 1309–1335.

Kwiatkowski, D., P. C. B. Phillips, P. Schmidt, and Y. Shin (1992), "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, 159–178.

Park, J. Y. (1990), "Testing for Unit Roots and Cointegration by Variable Addition," in *Advances in Econometrics: Co-integration, Spurious Regression, and Unit Roots* (ed. by T. B. Fomby and G. F. Rhodes, Jr.), Greenwich, CT: JAI Press, pp. 107–133.

Park, J. Y. (1992), "Canonical Cointegrating Regressions," *Econometrica*, 60, 119–143.

Park, J. Y., and B. Choi (1988), "A New Approach to Testing for a Unit Root," CAE Working Paper 88–23, Cornell University.

Phillips, P. C. B., and B. E. Hansen (1990), "Statistical Inference in Instrumental Variables Regression with I(1) Variables," *Review of Economic Studies*, 57, 99–125.

Rothenberg, T. J. (2000), "Testing for Unit Roots in AR and MA Models," in *Applications of Differential Geometry to Econometrics* (ed. by P. Marriott and M. Salmon), Cambridge, U.K.: Cambridge University Press, pp. 281–293.

Saikkonen, P. (1991), "Asymptotically Efficient Estimation of Cointegration Regressions," *Econometric Theory*, 7, 1–21.

Saikkonen, P., and R. Luukkonen (1993), "Point Optimal Tests for Testing the Order of Differencing in {ARIMA} Models," *Econometric Theory*, 9, 343–362.

Shin, Y. (1994), "A Residual-Based Test of the Null of Cointegration Against the Alternative of No Cointegration," *Econometric Theory*, 10, 91–115.

Stock, J. H. (1994), "Unit Roots, Structural Breaks and Trends," in *Handbook of Econometrics, Volume IV* (ed. by R. F. Engle and D. L. McFadden), New York: North Holland, pp. 2739–2841.

Stock, J. H., and M. W. Watson (1993), "A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems," *Econometrica*, 61, 783–820.

Xiao, Z. (2001), "Testing the Null Hypothesis of Stationarity Against an Autoregressive Unit Root Alternative," *Journal of Time Series Analysis*, 22, 87–105.

Xiao, Z., and P. C. B. Phillips (2002), "A CUSUM Test for Cointegration Using Regression Residals," *Journal of Econometrics*, 108, 43–61.