Point Optimal Tests of the Null Hypothesis of Cointegration

Michael Jansson*
Department of Economics, UC Berkeley

April 30, 2002

Abstract. This paper obtains an asymptotic Gaussian power envelope for tests of the null hypothesis of cointegration and proposes a feasible point optimal cointegration test. The local asymptotic power function of the point optimal test is close to the asymptotic Gaussian power envelope and the test is found to perform well in a Monte Carlo experiment.

Keywords: Cointegration, local asymptotic power, point optimal test, power envelope.

JEL Classification: C22.

1. Introduction

The concept of cointegration (Engle and Granger (1987)) has attracted considerable attention in the literature and answers to a variety of questions concerning inference in cointegrated systems have been provided. In particular, an asymptotic optimality theory for the estimation of cointegrating vectors under normality has been developed by Phillips (1991a) and several asymptotically efficient estimation procedures have been proposed (see e.g. Hansen (1992b), Johansen (1988, 1991), Park (1992), Phillips (1991b), Phillips and Hansen (1990), Saikkonen (1991, 1992) and Stock and Watson (1993)). Moreover, an asymptotic Gaussian power envelope for tests of the unit root assumption underlying these cointegration methods has been obtained by Elliott, Rothenberg, and Stock (1996). In contrast, although numerous papers have considered the problem of testing the null hypothesis of cointegration against the alternative of no cointegration (examples include Choi and Ahn (1995), Hansen (1992c), Harris

*Address correspondence to: Michael Jansson, Department of Economics, 549 Evans Hall #3880, UC Berkeley, Berkeley, CA 94720-3880, USA; e-mail:mjansson@econ.berkeley.edu.
This paper draws on material in Chapter 2 of the author's Ph.D. dissertation at University of Aarhus, Denmark. Helpful suggestions from Peter Boswijk, Bent J. Christensen, Graham Elliott, Niels Haldrup, Svend Hylleberg, Nick Kiefer and Jim Powell are gratefully acknowledged. The paper has benefited from the comments of seminar participants at University of Aarhus, University of California, Berkeley, University of British Columbia, Cambridge University, Tilburg University, the 1999 Econometric Society European Meeting and the 1999 CNMLE conference. A MATLAB program that implements the tests proposed in this paper is available from the author upon request.

and Inder (1994), Harris (1997), Jansson (2001a), Leybourne and McCabe (1993), McCabe, Leybourne, and Shin (1997), Park (1990), Quintos and Phillips (1993), Saikkonen and Luukkonen (1996), Shin (1994), Tanaka (1993), Xiao (1999) and Xiao and Phillips (2002)), no asymptotic optimality theory for that testing problem has been developed.

This paper obtains an asymptotic Gaussian power envelope for tests of the null hypothesis of cointegration and proposes a feasible point optimal cointegration test. By construction, the point optimal test attains the asymptotic Gaussian power envelope at a prespecified alternative. Against other alternatives, the local asymptotic power of the point optimal test is close to the asymptotic Gaussian power envelope. Moreover, the point optimal test is found to perform well in a Monte Carlo experiment.

In terms of the methodology employed, the present paper is related to Elliott, Rothenberg, and Stock's (1996) and Saikkonen and Luukkonen's (1993) studies of unit root testing in autoregressive and moving average models, respectively. As do Elliott, Rothenberg, and Stock (1996) and Saikkonen and Luukkonen (1993), this paper develops asymptotic optimality results by obtaining limiting distributions of optimal test statistics derived under a normality assumption. By accommodating stochastically trending and (possibly) endogenous regressors, the results obtained here generalize the fixed-regressor results of Saikkonen and Luukkonen (1993).

Section 2 presents the model. In Section 3, the asymptotic Gaussian power envelope is derived. Section 4 constructs a feasible point optimal test and Section 5 reports local asymptotic power results and investigates the finite sample performance of the test proposed in this paper. Finally, Section 6 offers a few concluding remarks, while all mathematical derivations are collected in an Appendix.

2. THE MODEL AND ASSUMPTIONS

Let z_t be an observed m-vector time series generated by

$$z_t = \mu_t^z + z_t^0, \qquad 1 \leq t \leq T, \tag{1}$$

where μ_t^z is a deterministic component and z_t^0 is a zero-mean stochastic component. For concreteness, the deterministic component is assumed to be a pth order polynomial time trend:

$$\mu_t^z = \alpha_z' d_t, \tag{2}$$

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

Partition z_t^0 into a scalar y_t^0 and a k-vector x_t^0 $(k = m - 1)$ as $z_t^0 = (y_t^0, x_t^{0\prime})'$ and suppose z_t^0 is generated by the potentially cointegrated system

$$\begin{aligned} y_t^0 &= \beta' x_t^0 + v_t, \\ \Delta x_t^0 &= u_t^x, \end{aligned} \tag{3}$$

where v_t is an error process with generating mechanism

$$\Delta v_t = u_t^y - \theta u_{t-1}^y, \qquad 1 \leq t \leq T. \tag{4}$$

In $(3) - (4)$, $\beta \in \mathbb{R}^k$ and $\theta \in (-1, 1]$ are unknown parameters and $u_t = (u_t^y, u_t^{x\prime})'$ satisfies the following assumption.

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 $\sum_{i=0}^{\infty} i \|C_i\| < \infty$, where $\|\cdot\|$ denotes the Euclidean norm.

The system is initialized at $t = 0$ with $v_0 = u_0^y = 0$ and $x_0^0 = 0$. The moving average specification (4) implies that y_t^0 and x_t^0 are cointegrated if and only if $\theta = 1$. Under A1, x_t^0 is a non-cointegrated integrated process and the cointegration between y_t^0 and x_t^0 is regular (in the sense of Park (1992)) when $\theta = 1$.

Conformably with z_t^0, partition z_t and α_z as $z_t = (y_t, x_t')'$ and $\alpha_z = (\alpha_y, \alpha_x)$. Defining $\alpha = \alpha_y - \alpha_x \beta$, the model can be written in triangular form as

$$\begin{aligned} y_t &= \alpha' d_t + \beta' x_t + v_t, \qquad \Delta v_t = u_t^y - \theta u_{t-1}^y, \\ \Delta x_t &= \alpha_x' \Delta d_t + u_t^x. \end{aligned}$$

This paper considers the problem of testing

$$H_0 : \theta = 1 \qquad \text{vs.} \qquad H_1 : \theta < 1, \tag{$\mathcal{P}$}$$

treating α and β as unknown nuisance parameters. Section 3 derives an asymptotic power envelope under the following strengthening of A1.

A1*. $u_t \sim$ *i.i.d.* $\mathcal{N}(0, \Sigma)$, where Σ is positive definite.

As a by-product of the analysis, a point optimal test statistic is obtained. Section 4 relaxes A1* and constructs a feasible point optimal test which is applicable whenever A1 holds.

3. AN ASYMPTOTIC GAUSSIAN POWER ENVELOPE

Under A1*, the model is fully parametric and the principle of invariance (e.g. Lehmann (1994, Chapter 6)) can be used to develop optimality theory for the testing problem $(\mathcal{P})$. In particular, it is possible to construct an upper bound on the local asymptotic power of a class of cointegration tests. The objective is to find an upper bound which is attainable for any value of α, β, α_x and Σ without knowledge of these nuisance parameters and covers a class of tests which is as large as possible.

For now, suppose α_x and Σ are known so that the only unknown nuisance parameters are α and β. Define the matrices $Y = (y_1, \ldots, y_T)'$, $X = (x_1, \ldots, x_T)'$ and $D = (d_1, \ldots, d_T)'$. The testing problem $(\mathcal{P})$ is invariant under transformations of the form

$$(Y, X) \rightarrow (Y + D \cdot a + X \cdot b, X), \qquad (\mathcal{G})$$

where $a \in \mathbb{R}^{p+1}$ and $b \in \mathbb{R}^k$. It therefore seems natural to restrict attention to tests that are invariant under $(\mathcal{G})$. All previously proposed tests (of which the author is aware) are invariant under this group of transformations, so the class of tests that are invariant under $(\mathcal{G})$ is quite large. A maximal invariant under $(\mathcal{G})$ is $\left((R_\perp' Y)', vec(X)'\right)'$, where $R_\perp$ is a matrix whose columns form an orthonormal basis for the orthogonal complement of the column space of $R = (X, D)$.

Partition Σ in conformity with u_t as

$$\Sigma = \begin{pmatrix} \sigma_{yy} & \sigma_{xy}' \\ \sigma_{xy} & \Sigma_{xx} \end{pmatrix}$$

and for any θ^*, define $\Psi_{\theta^*} = \Psi_{\theta^*}^{1/2} \Psi_{\theta^*}^{1/2\prime}$, where

$$\Psi_{\theta^*}^{1/2} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1-\theta^* & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 1-\theta^* & \cdots & 1-\theta^* & 1 \end{pmatrix}$$

is a $T \times T$ matrix with ones on the diagonal and $1 - \theta^*$ below the diagonal. Under A1*,

$$vec(X) \sim \mathcal{N}\left((\alpha_x' \otimes I_T) vec(D), \Sigma_{xx} \otimes \Psi_0\right),$$

while

$$R_\perp' Y \sim \mathcal{N}\left(\theta\left(R_\perp' \Psi_0^{-1/2} X\right)\Sigma_{xx}^{-1}\sigma_{xy}, \sigma_{yy.x}\left(R_\perp' \Psi_\theta R_\perp\right)\right)$$

conditional on X, where $\sigma_{yy.x} = \sigma_{yy} - \sigma_{xy}'\Sigma_{xx}^{-1}\sigma_{xy}$. Apart from an additive term that does not depend on θ, minus 2 times the log density of the maximal invariant can therefore be written as

$$\mathcal{L}_T(\theta) = \log\left|R'\Psi_\theta^{-1}R\right| + \sigma_{yy.x}^{-1} Y_\theta' \left(\Psi_\theta^{-1} - \Psi_\theta^{-1}R\left(R'\Psi_\theta^{-1}R\right)^{-1}R'\Psi_\theta^{-1}\right)Y_\theta, \qquad (5)$$

where $Y_{\theta^*} = Y - \theta^* \cdot \Psi_0^{-1/2} X \Sigma_{xx}^{-1}\sigma_{xy}$ and the derivation of (5) makes use of the relations

$$R_\perp\left(R_\perp'\Psi_{\theta^*}R_\perp\right)^{-1}R_\perp' = \Psi_{\theta^*}^{-1} - \Psi_{\theta^*}^{-1}R\left(R'\Psi_{\theta^*}^{-1}R\right)^{-1}R'\Psi_{\theta^*}^{-1}$$

and

$$\left|R_\perp'\Psi_{\theta^*}R_\perp\right| = \left|R'\Psi_{\theta^*}^{-1}R\right| \cdot \left|\Psi_{\theta^*}\right| \cdot \left|R'R\right|^{-1} = \left|R'\Psi_{\theta^*}^{-1}R\right| \cdot \left|R'R\right|^{-1}.$$

The expression (5) differs from its fixed-regressor counterpart (e.g. King (1980), King and Hillier (1985)) in two respects. In fixed-regressor settings, the term corresponding to $\log\left|R'\Psi_\theta^{-1}R\right|$ is non-random and can be omitted. Moreover, the definition of Y_θ reflects the fact that correlations between Y and X must be taken into account when X is random. As in the fixed-regressor case, the term

$$Y_\theta'\left(\Psi_\theta^{-1} - \Psi_\theta^{-1}R\left(R'\Psi_\theta^{-1}R\right)^{-1}R'\Psi_\theta^{-1}\right)Y_\theta$$

in (5) can be interpreted as the weighted sum of squared residuals from a GLS regression (of Y_θ on R using the covariance matrix Ψ_θ).

It follows from the Neyman-Pearson Lemma that the test which rejects when $\mathcal{L}_T(1) - \mathcal{L}_T\left(\bar\theta\right)$ is large is the most powerful invariant test of $\theta = 1$ vs. $\theta = \bar\theta < 1$. An asymptotic analogue of that optimality result can be obtained by deriving the limiting distribution of $\mathcal{L}_T(1) - \mathcal{L}_T\left(\bar\theta\right)$ under a local-to-unity reparameterization of θ and $\bar\theta$ in which $\lambda = T(1-\theta) \geq 0$ and $\bar\lambda = T\left(1-\bar\theta\right) > 0$ are held constant as T increases without bound. A formal statement is provided in Theorem 1, the proof of which represents the limiting distribution of $\mathcal{L}_T(1) - \mathcal{L}_T\left(\bar\theta\right)$ in terms of the random functional

$$
\begin{aligned}
&\varphi\left(\lambda;\bar{\lambda}\right) = \\
&2\bar{\lambda}\int_0^1 V_{\bar{\lambda}}^{\lambda} dV^{\lambda} - \bar{\lambda}^2 \int_0^1 \left(V_{\bar{\lambda}}^{\lambda}\right)^2 \\
&+\left(\int_0^1 Q_{\bar{\lambda}} dV_{\bar{\lambda}}^{\lambda}\right)' \left(\int_0^1 Q_{\bar{\lambda}} Q_{\bar{\lambda}}'\right)^{-1} \left(\int_0^1 Q_{\bar{\lambda}} dV_{\bar{\lambda}}^{\lambda}\right) - \log\left|\int_0^1 Q_{\bar{\lambda}} Q_{\bar{\lambda}}'\right| \\
&-\left(\int_0^1 Q dV^{\lambda}\right)' \left(\int_0^1 Q Q'\right)^{-1} \left(\int_0^1 Q dV^{\lambda}\right) + \log\left|\int_0^1 Q Q'\right|,
\end{aligned}
$$

where

$$
\begin{aligned}
V_{\bar{\lambda}}^{\lambda}(s) &= \int_0^s e^{-\bar{\lambda}(s-t)} dV^{\lambda}(t), \qquad V^{\lambda}(s) = V(s) + \lambda \int_0^s V(t)\,dt, \\
Q_{\bar{\lambda}}(s) &= \int_0^s e^{-\bar{\lambda}(s-t)} dQ(t), \qquad Q(s) = \left(1, \ldots, s^p, W(s)'\right)',
\end{aligned}
$$

and V and W are independent Wiener processes of dimensions 1 and k, respectively.

Theorem 1. *Let z_t be generated by* (1) − (4) *and suppose A1* holds. An upper bound on the local asymptotic power of any invariant test of $\theta = 1$ against $\theta = \theta_T = 1 - T^{-1}\lambda$ is given by $\pi^{\delta}(\lambda) = \Pr\left[\varphi(\lambda;\lambda) > c^{\delta}(\lambda)\right]$, where δ is the asymptotic level of the test, $c^{\delta}(\lambda)$ satisfies $\Pr\left[\varphi(0;\lambda) > c^{\delta}(\lambda)\right] = \delta$ and invariance is with respect to transformations of the form $(\mathcal{G})$.*

Under the reparameterization employed in Theorem 1, the null and alternative hypotheses are $\lambda = 0$ and $\lambda > 0$, respectively. The upper bound provided by the Gaussian power envelope is sharp in the sense that $\pi^{\delta}(\lambda)$ can be attained for any given λ by the test which rejects for large values of the corresponding point optimal test statistic, viz. $\mathcal{L}_T(1) - \mathcal{L}_T(1 - T^{-1}\lambda)$. Indeed, previous research on special cases of the testing problem considered here (e.g. Saikkonen and Luukkonen (1993), Rothenberg (2000)) suggests that a test based on $\mathcal{L}_T(1) - \mathcal{L}_T\left(\bar{\theta}\right)$ should have a local asymptotic power function close to $\pi^{\delta}(\cdot)$ if $\bar{\theta}$ is chosen appropriately. The function $\pi^{\delta}(\cdot)$ therefore constitutes a useful benchmark against which the local power function of any invariant test of the null hypothesis of cointegration can be compared.

The power envelope depends on p, the order of the deterministic component, and k, the dimension of x_t. On the other hand, although the form of the point optimal test statistic depends on Σ, the power envelope does not depend on the covariance matrix of the underlying errors u_t. In particular, the power envelope does not depend on the extent to which the regressors are endogenous in the sense that Δx_t is correlated with the latent error u_t^y. Jansson (2002) derives the asymptotic Gaussian power envelope in a model isomorphic to the present model under the assumption that β is known. That power envelope depends on $\sigma_{yy}^{-1}\sigma_{xy}'\Sigma_{xx}^{-1}\sigma_{xy}$, the squared coefficient of multiple correlation between u_t^y and u_t^x, and substantial power gains over conventional (univariate) tests are available when the correlation is non-zero. In contrast, it follows from Theorem 1 that it is impossible to exploit any correlations between Δx_t and u_t^y when testing the null hypothesis that y_t and x_t are cointegrated with an unknown cointegrating vector β.

4. Feasible Point Optimal Tests

This section constructs a test statistic which can be computed without knowledge of any nuisance parameters and has a limiting distribution of the form $\varphi\left(\lambda;\bar{\lambda}\right)$ under A1. Let $\hat{\Sigma}, \hat{\Gamma}$ and $\hat{\Omega}$ denote estimators of

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

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

and

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

respectively. Let $\hat{\Gamma}_{x\cdot} = \left(\hat{\gamma}_{xy}, \hat{\Gamma}_{xx}\right)$, $\hat{\omega}_{yy.x} = \hat{\kappa}'\hat{\Omega}\hat{\kappa}$ and $\hat{\gamma}_{yy.x} = \hat{\kappa}'\hat{\Gamma}\hat{\kappa}$, where $\hat{\kappa} = \left(1, -\hat{\omega}_{xy}'\hat{\Omega}_{xx}^{-1}\right)'$ and the matrices $\hat{\Gamma}$ and $\hat{\Omega}$ have been partitioned in conformity with u_t. Moreover, let $X_D = M_D X$ and $Y_D = M_D Y$, where $M_D = I - D\left(D'D\right)^{-1}D'$. Define $R^+ = \left(D, X - \hat{U}\hat{\Sigma}^{-1}\hat{\Gamma}_{x\cdot}'\right)$, where $\hat{U} = \left(Y_D - X_D\hat{\beta}, \Psi_0^{-1/2}X_D\right)$ and $\hat{\beta} = \left(X_D'X_D\right)^{-1}X_D'Y_D$. Finally, let $Y_{\theta^*}^+ = Y - \theta^*\cdot\Delta X\hat{\Omega}_{xx}^{-1}\hat{\omega}_{xy} - \hat{U}\hat{\Sigma}^{-1}\hat{\Gamma}_{x\cdot}'\hat{\beta}$ for any θ^*.

The proposed test statistic is

$$P_T\left(\bar{\lambda}\right) = \mathcal{L}_T^+(1) - \mathcal{L}_T^+\left(1 - T^{-1}\bar{\lambda}\right) - 2\bar{\lambda}\hat{\omega}_{yy.x}^{-1}\hat{\gamma}_{yy.x}, \tag{6}$$

where $\bar{\lambda} > 0$ is a prespecified constant and

$$\mathcal{L}_T^+(\theta^*) = \log\left|R^{+\prime}\Psi_{\theta^*}^{-1}R^+\right| + \hat{\omega}_{yy.x}^{-1}Y_{\theta^*}^{+\prime}\left(\Psi_{\theta^*}^{-1} - \Psi_{\theta^*}^{-1}R^+\left(R^{+\prime}\Psi_{\theta^*}^{-1}R^+\right)^{-1}R^{+\prime}\Psi_{\theta^*}^{-1}\right)Y_{\theta^*}^+.$$

Under the assumptions of Theorem 1, $P_T\left(\bar{\lambda}\right)$ is asymptotically equivalent to $\mathcal{L}_T(1) - \mathcal{L}_T\left(\bar{\theta}\right)$. As Theorem 2 shows, $P_T\left(\bar{\lambda}\right)$ has a limiting distribution of the form $\varphi\left(\lambda;\bar{\lambda}\right)$ even when u_t exhibits serial correlation of the form permitted under A1. This robustness property, not shared by $\mathcal{L}_T(1) - \mathcal{L}_T\left(\bar{\theta}\right)$, is achieved by employing two serial correlation corrections. The first correction, employed in the construction of $Y_{\theta^*}^+$ and R^+, is similar to the correction proposed by Park (1992) in the context of estimation of a cointegrating regression. Indeed, the purpose of this correction is to remove "serial correlation bias" from the limiting distribution of the estimator of β appearing in $\mathcal{L}_T^+(\theta^*)$. The second correction term, $-2\bar{\lambda}\hat{\omega}_{yy.x}^{-1}\hat{\gamma}_{yy.x}$, in (6) resembles the correction term in Phillips's (1987) Z_α test for an autoregressive unit root and accommodates serial correlation in $u_t^y - \omega_{xy}'\Omega_{xx}^{-1}u_t^x$.

Theorem 2. *Let z_t be generated by $(1)-(4)$, suppose A1 holds and suppose $\theta = \theta_T = 1 - T^{-1}\lambda$ for some $\lambda \geq 0$. If $\left(\hat{\Sigma},\hat{\Gamma},\hat{\Omega}\right) \to_p (\Sigma,\Gamma,\Omega)$, then $P_T\left(\bar{\lambda}\right) \to_d \varphi\left(\lambda;\bar{\lambda}\right)$.*

A consistent estimator of Σ is $\hat{\Sigma} = T^{-1}\hat{U}'\hat{U}$, while Γ and Ω can be estimated consistently by means of conventional (possibly prewhitened) kernel estimators (e.g. Andrews (1991), Andrews and Monahan (1992)). Primitive sufficient conditions (on the kernel, the bandwidth sequence and the prewhitening procedure) for consistency can be found in previous work by the author (Jansson (2001b), Jansson and Haldrup (2001) and Jansson (2002)). Suffice it to say that the consistency requirement of Theorem 2 is met by a variety of estimators $\hat{\Sigma},\hat{\Gamma}$ and $\hat{\Omega}$.

To implement the feasible point optimal test, a value of $\bar{\lambda}$ must be specified. Following Elliott, Rothenberg, and Stock (1996), the approach advocated here is to choose $\bar{\lambda}$ in such a way that the asymptotic local power against the alternative $\theta = 1 - T^{-1}\bar{\lambda}$ is approximately equal to 50% when the 5% test based on $P_T\left(\bar{\lambda}\right)$ is used. That is, the recommendation is to use the test which is (asymptotically) 0.50-optimal, level 0.05 in the sense of Davies (1969). Table 1a tabulates the recommended values of $\bar{\lambda}$ for $k = 1,\ldots,6$ regressors in the constant mean $(p = 0)$ case and reports selected percentiles of the asymptotic null distributions of the corresponding $P_T\left(\bar{\lambda}\right)$ statistics. These percentiles were computed by generating $20,000$ draws from the discrete time approximation (based on $2,000$ steps) to the limiting random variables. Table 1b contains similar information about the linear trend $(p = 1)$ version of $P_T\left(\bar{\lambda}\right)$.

TABLES 1a-1b ABOUT HERE

5. Power Properties

FIGURE 1 ABOUT HERE

5.1. Local Asymptotic Power.

Figure 1 plots the 5% asymptotic Gaussian power envelope $\pi^{0.05}(\cdot)$ along with the local asymptotic power functions of two feasible cointegration tests in the constant mean case with scalar x_t. The power envelope and the power functions were computed by generating $20,000$ draws from the discrete time approximation (based on $2,000$ steps) to $\varphi(\lambda;\lambda)$, $\varphi(\lambda;9)$ and the limiting distribution of S_T for various values of λ. Results for vector-valued x_t are qualitatively similar and have been omitted to conserve space. The two feasible tests, denoted $P_T(9)$ and S_T, are the tests proposed in Section 4 and the test due to Shin (1994), respectively. The latter test appears to be the most widely used cointegration test in applications. Moreover, S_T is known to enjoy local optimality properties under the assumptions of Theorem 1 (Harris and Inder (1994), Saikkonen and Luukkonen (1996)). In addition, previous research (Jansson and Haldrup (2001), Jansson (2001a)) indicates that none of the tests due to Choi and Ahn (1995), Hansen (1992c), Park (1990), Xiao (1999) and Xiao and Phillips (2002) dominate Shin's (1994) test in terms of local asymptotic power. For these reasons, it seems natural to use the performance of S_T as a benchmark when evaluating new tests such as $P_T(9)$.

As might be expected, the local asymptotic power of S_T is close to the envelope for small values of λ (smaller than 5, say). For larger values of λ, on the other hand, the local asymptotic power of the locally optimal test is well below the envelope. In contrast, the local asymptotic power of $P_T(9)$ is close to the envelope for all values of λ. In particular, the local asymptotic power properties of $P_T(9)$ are similar to those of S_T for small values of λ (when the latter is optimal) and $P_T(9)$ dominates S_T in terms of local asymptotic power for larger values of λ. As is apparent from Figure 2, the situation is similar in the linear trend case although the magnitude of the differences is smaller.

FIGURE 2 ABOUT HERE

5.2. Finite Sample Evidence.

To gauge the extent to which the predictions from the asymptotic power results of Section 5.1 can be expected to be borne out in sample sizes encountered in practice, a small Monte Carlo experiment is conducted. Samples of size $T = 200$ are generated according to the bivariate version of $(1) - (4)$ under A1*. The parameters α, β and α_x are normalized to zero and the variances σ_{yy} and σ_{xx} are normalized to one. The covariance σ_{xy} is $0, 0.2$ or 0.5, while the parameter of interest, θ, takes on values in the set $\{1, 0.975, 0.95, 0.925, 0.90\}$.

The matrix Σ is estimated by $\hat{\Sigma} = T^{-1}\hat{U}'\hat{U}$. Based on the recommendations of Andrews (1991), the parameters Γ and Ω are estimated by means of a kernel estimator using the QS kernel along with a plug-in bandwidth. The plug-in bandwidth is $1.3221 \cdot T^{1/5} \max\left(\min\left(\hat{a}^{1/5}, 5\right), 0.05\right)$, where $\hat{a}$ is computed from Andrews's (1991) equation (6.4) and the censoring involved in the construction of the bandwidth guarantees consistency of the test. Table 2 reports observed rejection rates (based on $5,000$ replications) of 5% level tests using the constant mean and linear trend versions of the test statistics $P_T\left(\bar{\lambda}\right)$ and S_T. To facilitate comparisons, the $P_T\left(\bar{\lambda}\right)$ and S_T test are implemented using the same estimation strategy. That is, both tests use the same estimators $\hat{\Sigma}$ and $\hat{\Gamma}$ and both tests are based on a correction in the spirit of Park (1992). The version of S_T constructed in this fashion is described in Choi and Ahn (1995), where it is denoted $SBDH_{II}$.

TABLE 2 ABOUT HERE

Rejection rates are reasonably close to 5% under the null with $P_T\left(\bar{\lambda}\right)$ exhibiting slightly larger size distortions than S_T. The power of the point optimal test is comparable to the power of the locally optimal test for values of θ close to unity. Against more distant alternatives, the point optimal test dominates the locally optimal test. These power results are consistent with the asymptotic results of Section 5.1. In particular, the simulation evidence is favorable to the feasible point optimal test and suggests that nontrivial power gains can be achieved by employing the test proposed in this paper.

6. Conclusion

An asymptotic Gaussian power envelope for tests of the null hypothesis of cointegration was obtained and a feasible point optimal cointegration test was proposed. The local asymptotic power function of the point optimal test is close to the asymptotic Gaussian power envelope and the test was found to perform well in a simple Monte Carlo experiment.

7. APPENDIX: PROOFS

Throughout the Appendix, $\lfloor\cdot\rfloor$ denotes the integer part of the argument and all functions are understood to be CADLAG functions defined on the unit interval (equipped with the Skorohod topology). Lemma 3, adapted from Jansson (2002), will be used to derive limiting distributions of sample moments of GLS transformed data $\Psi_{1-T^{-1}\bar{\lambda}}^{-1/2}Y_{1-T^{-1}\bar{\lambda}}^{+}$ and $\Psi_{1-T^{-1}\bar{\lambda}}^{-1/2}R^{+}$ from limiting distributions of sample moments of $Y_{1-T^{-1}\bar{\lambda}}^{+}$ and R^{+}. Indeed, the following relation, in which $\left\{F_{Tt}\left(\bar{\lambda}\right):1\leq t\leq T\right\}$ is expressed in terms of $\{F_{Tt}:1\leq t\leq T\}$, motivates the transformations considered in Lemma 3:

$$\left(F_{T1}\left(\bar{\lambda}\right),F_{T2}\left(\bar{\lambda}\right),\ldots,F_{TT}\left(\bar{\lambda}\right)\right)'=\Psi_{1-T^{-1}\bar{\lambda}}^{-1/2}\left(F_{T1},F_{T2},\ldots,F_{TT}\right)'.$$

Lemma 3. *Let $\{F_{Tt}:0\leq t\leq T,T\geq 1\}$ and $\{g_{Tt}:1\leq t\leq T,T\geq 1\}$ be triangular arrays of (vector) random variables with $F_{T0}=0$ for all T. Let $\bar{\lambda}>0$ be given and define $F_{Tt}\left(\bar{\lambda}\right)=\Delta F_{Tt}+\left(1-T^{-1}\bar{\lambda}\right)F_{T,t-1}\left(\bar{\lambda}\right)$ and $g_{Tt}\left(\bar{\lambda}\right)=\Delta g_{Tt}+\left(1-T^{-1}\bar{\lambda}\right)g_{T,t-1}\left(\bar{\lambda}\right)$ with initial conditions $F_{T0}\left(\bar{\lambda}\right)=F_{T0}$ and $g_{T1}\left(\bar{\lambda}\right)=g_{T1}$. If*

$$\begin{pmatrix} F_{T,\lfloor T\cdot\rfloor} \\ T^{-1}\sum_{t=1}^{\lfloor T\cdot\rfloor}g_{Tt} \\ T^{-1}\sum_{t=1}^{\lfloor T\cdot\rfloor}F_{Tt}g_{Tt}' \\ T^{-2}\sum_{t=2}^{\lfloor T\cdot\rfloor}\left(\sum_{i=1}^{t-1}g_{Ti}\right)g_{Tt}' \end{pmatrix}\rightarrow_d\begin{pmatrix} F\left(\cdot\right) \\ G\left(\cdot\right) \\ \int_0^{\cdot}F\left(s\right)dG\left(s\right)'+\Gamma_{FG}\left(\cdot\right) \\ \int_0^{\cdot}G\left(s\right)dG\left(s\right)'+\Gamma_{GG}\left(\cdot\right) \end{pmatrix},\qquad(7)$$

where F and G are continuous semimartingales and Γ_{FG} and Γ_{GG} are continuous, then

$$\begin{pmatrix} F_{T,\lfloor T\cdot\rfloor}\left(\bar{\lambda}\right) \\ g_{T,\lfloor T\cdot\rfloor}-g_{T,\lfloor T\cdot\rfloor}\left(\bar{\lambda}\right) \\ T^{-1}\sum_{t=1}^{\lfloor T\cdot\rfloor}F_{Tt}\left(\bar{\lambda}\right)g_{Tt}\left(\bar{\lambda}\right)' \\ T^{-1}\sum_{t=1}^{\lfloor T\cdot\rfloor}\left(g_{Tt}-g_{Tt}\left(\bar{\lambda}\right)\right)g_{Tt}' \end{pmatrix}\rightarrow_d\begin{pmatrix} F_{\bar{\lambda}}\left(\cdot\right) \\ \bar{\lambda}G_{\bar{\lambda}}\left(\cdot\right) \\ \int_0^{\cdot}F_{\bar{\lambda}}\left(s\right)dG_{\bar{\lambda}}\left(s\right)'+\Gamma_{FG}\left(\cdot\right) \\ \bar{\lambda}\left(\int_0^{\cdot}G_{\bar{\lambda}}\left(s\right)dG\left(s\right)'+\Gamma_{GG}\left(\cdot\right)\right) \end{pmatrix},\qquad(8)$$

where $F_{\bar{\lambda}}\left(s\right)=\bar{\lambda}\int_0^s\exp\left(-\bar{\lambda}\left(s-t\right)\right)dF\left(t\right)$ and $G_{\bar{\lambda}}\left(s\right)=\bar{\lambda}\int_0^s\exp\left(-\bar{\lambda}\left(s-t\right)\right)dG\left(t\right)$. Joint convergence in (7) and (8) also applies.

Proof of Theorems 1 and 2. Under the assumptions of Theorem 1, $\Gamma=0$ and $\Sigma=\Omega$. Theorem 1 therefore follows from the Neyman-Pearson Lemma, Theorem 2 and the fact that the distribution of $\varphi\left(\lambda,\lambda\right)$ is continuous.

Next, to prove Theorem 2, let $\bar{\theta}=\bar{\theta}_T=1-T^{-1}\bar{\lambda}$ and for $\theta^*\in\left\{1,\bar{\theta}\right\}$, let $\left(y_1^+\left(\theta^*\right),\ldots,y_T^+\left(\theta^*\right)\right)=Y_{\theta^*}^+$. Let $x_t^+=x_t-\hat{\Gamma}_{x\cdot}\hat{\Sigma}^{-1}\hat{u}_t$, where $\left(\hat{u}_1,\ldots,\hat{u}_T\right)'=\hat{U}$. Now, $y_t^+\left(\theta^*\right)$ can be written as

$$y_t^+\left(\theta^*\right)=\alpha\left(\theta^*\right)'d_t+\beta'x_t^++v_t^+\left(\theta^*\right),$$

where $\alpha(\theta^*)$ satisfies $\alpha(\theta^*)' d_t = (\alpha + \alpha_x' \Omega_{xx}^{-1} \omega_{xy} (1-\theta))' d_t - \theta^* \hat{\omega}_{xy}' \hat{\Omega}_{xx}^{-1} \alpha_x' \Delta d_t$ and $v_t^+(\theta^*) = v_t^{++} + \hat{v}_t^{++}(\theta^*)$, where $v_t^{++} = (1-\theta)\sum_{s=1}^t u_s^{y.x} + \theta u_t^{y.x}$, $u_t^{y.x} = u_t^y - \omega_{xy}' \Omega_{xx}^{-1} u_t^x$ and $\hat{v}_t^{++}(\theta^*) = -\left(\theta^* \hat{\Omega}_{xx}^{-1} \hat{\omega}_{xy} - \theta \Omega_{xx}^{-1} \omega_{xy}\right)' u_t^x - \left(\hat{\beta} - \beta - (1-\theta)\Omega_{xx}^{-1}\omega_{xy}\right)\hat{u}_t$. Moreover, $x_t^+ = x_t^{++} + \hat{x}_t^{++}$, where $x_t^{++} = x_t - \Gamma_{x\cdot}\Sigma^{-1} u_t$ and $\hat{x}_t^{++} = \left(\Gamma_{x\cdot}\Sigma^{-1} - \hat{\Gamma}_{x\cdot}\hat{\Sigma}^{-1}\right) u_t + \hat{\Gamma}_{x\cdot}\hat{\Sigma}^{-1}(u_t - \hat{u}_t)$. Similarly, $r_t^+ = \left(d_t', x_t^{+\prime}\right)' = r_t^{++} + \hat{r}_t^{++}$, where $r_t^{++} = \left(d_t', x_t^{++\prime}\right)$ and $\hat{r}_t^{++} = \left(0, \hat{x}_t^{++\prime}\right)$.

By proceeding as in the proof of Jansson and Haldrup (2001, Lemma 6), standard weak convergence results for linear processes (e.g. Phillips and Solo (1992), Phillips (1988), Hansen (1992a)) can be used to show that the following hold jointly:

$$T^{1/2}\Upsilon_T r_{\lfloor T\cdot\rfloor}^{+} = T^{1/2}\Upsilon_T r_{\lfloor T\cdot\rfloor}^{++} + o_p(1) \to_d Q(\cdot) \tag{9}$$

$$T^{-1/2}\sum_{t=1}^{\lfloor T\cdot\rfloor} v_t^+(\theta^*) = T^{-1/2}\sum_{t=1}^{\lfloor T\cdot\rfloor} v_t^{++} + o_p(1) \to_d \omega_{yy.x}^{1/2} V^\lambda(\cdot), \tag{10}$$

$$\begin{aligned}\Upsilon_T \sum_{t=1}^{\lfloor T\cdot\rfloor} r_t^+ v_t^+(\theta^*) &= \Upsilon_T \sum_{t=1}^{\lfloor T\cdot\rfloor} r_t^{++} v_t^{++} + o_p(1) \\ &\to_d \omega_{yy.x}^{1/2}\int_0^{\cdot} Q(s)\, dV^\lambda(s),\end{aligned} \tag{11}$$

$$\begin{aligned}T^{-1}\sum_{t=1}^{\lfloor T\cdot\rfloor}\left(\sum_{s=1}^{t-1} v_s^+(\theta^*)\right) v_t^+(\theta^*) &= T^{-1}\sum_{t=1}^{\lfloor T\cdot\rfloor}\left(\sum_{s=1}^{t-1} v_s^{++}\right) v_t^{++} + o_p(1) \\ &\to_d \omega_{yy.x}\int_0^{\cdot} V^\lambda(s)\, dV^\lambda(s) + \gamma_{yy.x}\int_0^{\cdot} ds,\end{aligned} \tag{12}$$

for $\theta^* \in \left\{1, \bar{\theta}\right\}$, where

$$\Upsilon_T = \begin{pmatrix} \operatorname{diag}\left(T^{-1/2}, \ldots, T^{-(p+1/2)}\right) & 0' \\ -T^{-1}\Omega_{xx}^{-1/2}\alpha_x' & T^{-1}\Omega_{xx}^{-1/2} \end{pmatrix}.$$

Since $\mathcal{L}_T^+(\theta^*)$ is invariant under transformations of the form $Y_{\theta^*}^+ \to Y_{\theta^*}^+ + D \cdot a + X^+ \cdot b$, $P_T(\bar{\lambda})$ can be written as $P_T^1(\bar{\lambda}) + P_T^2(\bar{\lambda}) + P_T^3(\bar{\lambda})$, where

$$\begin{aligned}
P_T^1(\bar{\lambda}) &= \log\left|R^{+\prime}R^+\right| - \log\left|R^{+\prime}\Psi_{\bar{\theta}}^{-1}R^+\right|, \\
P_T^2(\bar{\lambda}) &= \hat{\omega}_{yy.x}^{-1}\left(V_1^{+\prime}V_1^+ - V_{\bar{\theta}}^{+\prime}\Psi_{\bar{\theta}}^{-1}V_{\bar{\theta}}^+ - 2\bar{\lambda}\hat{\gamma}_{yy.x}\right), \\
P_T^3(\bar{\lambda}) &= \hat{\omega}_{yy.x}^{-1}V_{\bar{\theta}}^{+\prime}\Psi_{\bar{\theta}}^{-1}R^+\left(R^{+\prime}\Psi_{\bar{\theta}}^{-1}R^+\right)^{-1}R^{+\prime}\Psi_{\bar{\theta}}^{-1}V_{\bar{\theta}}^+ \\
&\quad -\hat{\omega}_{yy.x}^{-1}V_1^{+\prime}R^+\left(R^{+\prime}R^+\right)^{-1}R^{+\prime}V_1^+,
\end{aligned}$$

and $V_{\theta^*}^+ = \left(v_1^+(\theta^*), \ldots, v_T^+(\theta^*)\right)'$ for $\theta^* \in \left\{1, \bar{\theta}\right\}$.

Using (9), Lemma 3, the continuous mapping theorem (CMT) and standard manipulations,

$$P_T^1(\bar{\lambda}) = \log\left|\Upsilon_T R^{+\prime}R^+\Upsilon_T'\right| - \log\left|\Upsilon_T R^{+\prime}\Psi_{\bar{\theta}}^{-1}R^+\Upsilon_T'\right|$$

$$\to_d \log\left|\int_0^1 QQ'\right| - \log\left|\int_0^1 Q_{\bar{\lambda}}Q_{\bar{\lambda}}'\right|.$$

Similarly, using (10), (12), $\left(\hat{\omega}_{yy.x}, \hat{\gamma}_{yy.x}\right) \to_p \left(\omega_{yy.x}, \gamma_{yy.x}\right)$, Lemma 3 and CMT,

$$\begin{aligned}
P_T^2\left(\bar{\lambda}\right) &= \hat{\omega}_{yy.x}^{-1}\left(V_1^{+\prime}V_1^{+} - V_{\bar{\theta}}^{+\prime}\Psi_{\bar{\theta}}^{-1}V_{\bar{\theta}}^{+} - 2\bar{\lambda}\hat{\gamma}_{yy.x}\right) \\
&= 2\hat{\omega}_{yy.x}^{-1}\left(\left(V_1^{+} - \Psi_{\bar{\theta}}^{-1/2}V_{\bar{\theta}}^{+}\right)' V_1^{+} - \bar{\lambda}\hat{\gamma}_{yy.x}\right) \\
&\quad - \hat{\omega}_{yy.x}^{-1}\left(V_1^{+} - \Psi_{\bar{\theta}}^{-1/2}V_{\bar{\theta}}^{+}\right)'\left(V_1^{+} - \Psi_{\bar{\theta}}^{-1/2}V_{\bar{\theta}}^{+}\right) \\
&= 2\omega_{yy.x}^{-1}\left(\left(V^{++} - \Psi_{\bar{\theta}}^{-1/2}V^{++}\right)' V^{++} - \bar{\lambda}\gamma_{yy.x}\right) \\
&\quad - \omega_{yy.x}^{-1}\left(V^{++} - \Psi_{\bar{\theta}}^{-1/2}V^{++}\right)'\left(V^{++} - \Psi_{\bar{\theta}}^{-1/2}V^{++}\right) + o_p\left(1\right) \\
&\rightarrow_d 2\bar{\lambda}\int_0^1 V_{\bar{\lambda}}^{\lambda} dV^{\lambda} - \bar{\lambda}^2\int_0^1 \left(V_{\bar{\lambda}}^{\lambda}\right)^2,
\end{aligned}$$

where $V^{++} = \left(v_1^{++}, \ldots, v_T^{++}\right)'$.

Finally, using (9), (11), $\hat{\omega}_{yy.x} \rightarrow_p \omega_{yy.x}$, Lemma 3 and CMT,

$$\begin{aligned}
P_T^3\left(\bar{\lambda}\right) &= \hat{\omega}_{yy.x}^{-1}\left(\Upsilon_T R^{+\prime}\Psi_{\bar{\theta}}^{-1}V_{\bar{\theta}}^{+}\right)'\left(\Upsilon_T R^{+\prime}\Psi_{\bar{\theta}}^{-1}R^{+}\Upsilon_T'\right)^{-1}\left(\Upsilon_T R^{+\prime}\Psi_{\bar{\theta}}^{-1}V_{\bar{\theta}}^{+}\right) \\
&\quad - \hat{\omega}_{yy.x}^{-1}\left(\Upsilon_T R^{+\prime}V_1^{+}\right)'\left(\Upsilon_T R^{+\prime}R^{+}\Upsilon_T'\right)^{-1}\left(\Upsilon_T R^{+\prime}V_1^{+}\right) \\
&\rightarrow_d \left(\int_0^1 Q_{\bar{\lambda}} dV_{\bar{\lambda}}^{\lambda}\right)'\left(\int_0^1 Q_{\bar{\lambda}}Q_{\bar{\lambda}}'\right)^{-1}\left(\int_0^1 Q_{\bar{\lambda}} dV_{\bar{\lambda}}^{\lambda}\right) \\
&\quad - \left(\int_0^1 Q dV^{\lambda}\right)'\left(\int_0^1 QQ'\right)^{-1}\left(\int_0^1 Q dV^{\lambda}\right).
\end{aligned}$$

The convergence results in the preceding displays hold jointly. Combining these results, Theorem 2 follows. ■

REFERENCES

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

ANDREWS, D. W. K., AND J. C. MONAHAN (1992): "An Improved Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator," *Econometrica*, 60, 953–966.

CHOI, I., AND B. C. AHN (1995): "Testing for Cointegration in a System of Equations," *Econometric Theory*, 11, 952–983.

DAVIES, R. B. (1969): "Beta-optimal Tests and an Application to the Summary Evaluation of Experiments," *Journal of the Royal Statistical Society, Series B*, 31, 524–538.

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.

HANSEN, B. E. (1992a): "Convergence to Stochastic Integrals for Dependent Heterogeneous Processes," *Econometric Theory*, 8, 489–500.

——— (1992b): "Efficient Estimation and Testing of Cointegrating Vectors in the Presence of Deterministic Trends," *Journal of Econometrics*, 53, 87–121.

——— (1992c): "Tests for Parameter Instability in Regressions with I(1) Processes," *Journal of Business and Economic Statistics*, 10, 321–335.

HARRIS, D. (1997): "Principal Components Analysis of Cointegrated Time Series," *Econometric Theory*, 13, 529–557.

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, pp. 133-152.

JANSSON, M. (2001a): "Another Test of the Null Hypothesis of Cointegration," Manuscript, UC Berkeley.

——— (2001b): "Consistent Covariance Matrix Estimation for Linear Processes," Econometric Theory, forthcoming.

——— (2002): "Stationarity Testing with Covariates," Manuscript, UC Berkeley.

Jansson, M., and N. Haldrup (2001): "Regression Theory for Nearly Cointegrated Time Series," Econometric Theory, forthcoming.

Johansen, S. (1988): "Statistical Analysis of Cointegration Vectors," *Journal of Economic Dynamics and Control*, 12, 231–254.

——— (1991): "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models," *Econometrica*, 59, 1551–1580.

King, M. L. (1980): "Robust Tests for Spherical Symmetry and their Application to Least Squares Regression," *Annals of Statistics*, 8, 1265–1271.

King, M. L., and G. H. Hillier (1985): "Locally Best Invariant Tests of the Error Covariance Matrix of the Linear Regression Model," *Journal of the Royal Statistical Society, Series B*, 47, 98–102.

Lehmann, E. L. (1994): *Testing Statistical Hypotheses*. Second Edition. New York: Chapman and Hall.

Leybourne, S. J., and B. P. M. McCabe (1993): "A Simple Test for Cointegration," *Oxford Bulletin of Economics and Statistics*, 55, 97–103.

McCabe, B. P. M., S. J. Leybourne, and Y. Shin (1997): "A Parametric Approach to Testing the Null of Cointegration," *Journal of Time Series Analysis*, 18, 395–413.

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.

——— (1992): "Canonical Cointegrating Regressions," *Econometrica*, 60, 119–143.

Phillips, P. C. B. (1987): "Time Series Regression with a Unit Root," *Econometrica*, 55, 277–301.

——— (1988): "Weak Convergence of Sample Covariance Matrices to Stochastic Integrals Via Martingale Approximations," *Econometric Theory*, 4, 528–533.

——— (1991a): "Optimal Inference in Cointegrated Systems," *Econometrica*, 59, 283–306.

——— (1991b): "Spectral Regression for Cointegrated Time Series," in *Nonparametric and Semiparametric Methods in Economics and Statistics*, ed. by W. A. Barnett, J. Powell, and G. E. Tauchen. Cambridge, U.K.: Cambridge University Press, pp. 413-435.

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.

PHILLIPS, P. C. B., AND V. SOLO (1992): "Asymptotics for Linear Processes," *Annals of Statistics*, 20, 971–1001.

QUINTOS, C. E., AND P. C. B. PHILLIPS (1993): "Parameter Constancy in Cointegrating Regressions," *Empirical Economics*, 18, 675–706.

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, Cambridge, U.K.: Cambridge University Press, pp. 281-293.

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

——— (1992): "Estimation and Testing of Cointegrated Systems by an Autoregressive Approximation," *Econometric Theory*, 8, 1–27.

SAIKKONEN, P., AND R. LUUKKONEN (1993): "Point Optimal Tests for Testing the Order of Differencing in ARIMA Models," *Econometric Theory*, 9, 343–362.

——— (1996): "Testing the Order of Differencing in Time Series Regression," *Journal of Time Series Analysis*, 88, 596–601.

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., AND M. W. WATSON (1993): "A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems," *Econometrica*, 61, 783–820.

TANAKA, K. (1993): "An Alternative Approach to the Asymptotic Theory of Spurious Regression, Cointegration, and Near Cointegration," *Econometric Theory*, 9, 36–61.

XIAO, Z. (1999): "A Residual Based Test for the Null Hypothesis of Cointegration," *Economics Letters*, 64, 133–141.

XIAO, Z., AND P. C. B. PHILLIPS (2002): "A CUSUM Test for Cointegration Using Regression Residals," *Journal of Econometrics*, 108, 43–61.

8. Tables

TABLE 1a

Percentiles of $P_T\left(\bar{\lambda}\right)$
Constant Mean: $d_t = 1$

	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$
$\bar{\lambda}$	9	10.5	12.5	14	16	17.5
90%	0.71	0.81	0.80	0.83	0.87	0.89
95%	1.70	1.82	1.82	1.87	1.88	1.91
97.5%	2.71	2.77	2.81	2.91	2.87	2.97
99%	3.93	4.20	4.03	4.27	4.34	4.40

TABLE 1b

Percentiles of $P_T\left(\bar{\lambda}\right)$
Linear Trend: $d_t = (1, t)'$

	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$
$\bar{\lambda}$	13.5	15.5	16.5	18	20	21.5
90%	0.84	0.82	0.94	0.98	1.01	1.09
95%	1.88	1.95	2.01	2.03	2.14	2.27
97.5%	2.87	3.04	3.12	3.03	3.28	3.29
99%	4.09	4.52	4.39	4.45	4.76	4.67

TABLE 2

MONTE CARLO REJECTION RATES
5% LEVEL TESTS, $T = 200$

		Constant Mean		Linear Trend	
σ_{xy}	θ	$P_T\,(9)$	S_T	$P_T\,(13.5)$	S_T
0	1.000	5.5	5.5	4.8	5.3
	0.975	22.8	22.0	11.4	12.6
	0.950	53.8	44.8	31.6	30.3
	0.925	75.2	61.8	56.7	50.4
	0.900	86.8	72.1	74.8	66.1
0.2	1.000	4.7	5.1	5.0	5.4
	0.975	22.4	21.3	10.9	12.4
	0.950	54.8	45.3	30.0	30.0
	0.925	75.9	62.3	53.1	48.3
	0.900	87.0	73.4	72.1	64.1
0.5	1.000	3.5	4.8	3.2	5.1
	0.975	18.6	20.3	7.7	11.2
	0.950	49.3	44.3	24.9	28.8
	0.925	71.7	60.7	48.3	48.0
	0.900	84.2	71.7	67.7	64.1

9. Figures

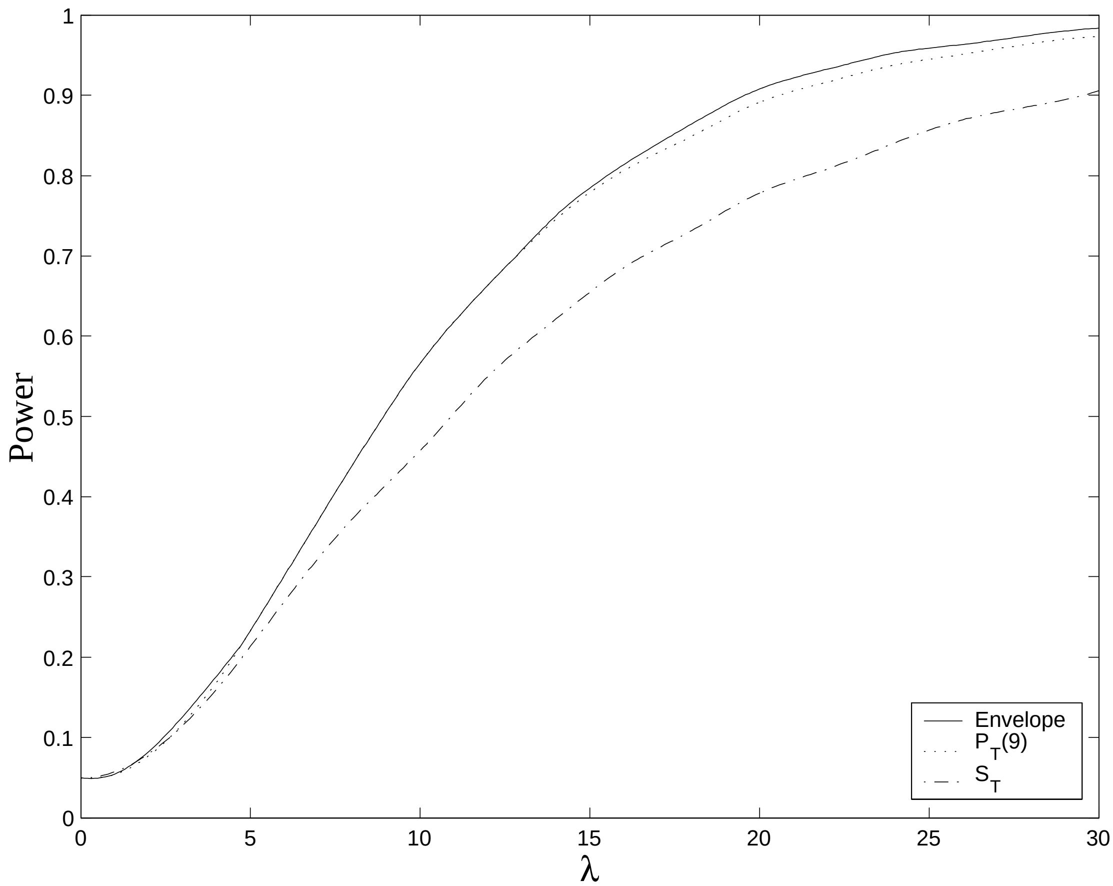

Figure 1: Power Curves

5% Level Tests, Constant Mean, Scalar x.

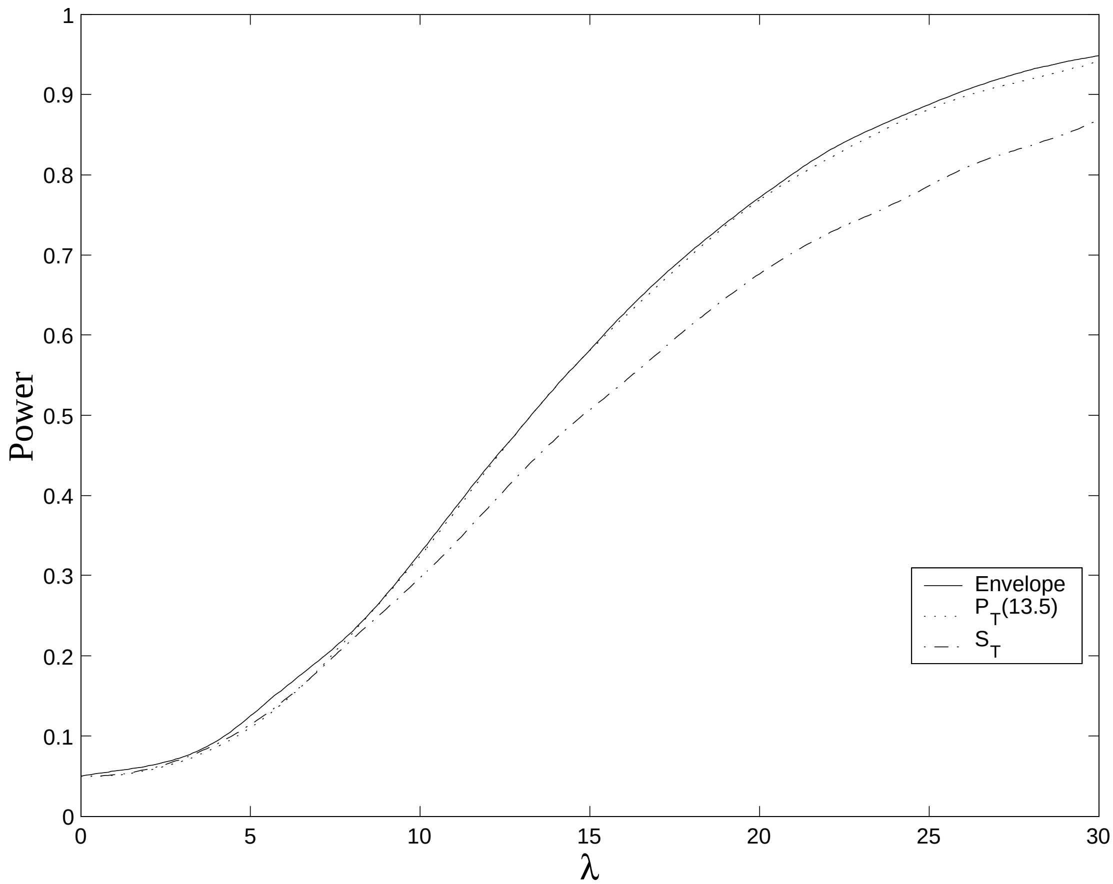

Figure 2: Power Curves

5% Level Tests, Linear Trend, Scalar x.