Queen's Economics Department Working Paper No. 811

Testing for Cointegration in Linear Quadratic Models

Allan W. Gregory

Department of Economics
Queen's University
94 University Avenue
Kingston, Ontario, Canada
K7L 3N6

5-1991

TESTING FOR COINTEGRATION IN LINEAR QUADRATIC MODELS

Allan W. Gregory
Department of Economics
Queen's University
Kingston, Ontario K7L 3N6

#811

The author would like to thank the department of Agricultural and Resource Economics at the University of California, Berkeley and the department of Economics at the University of British Columbia for the hospitality shown me while undertaking this research. The author would also like to thank Don Andrews, Anindya Banerjee, Søren Johansen, Adrian Pagan, Peter Phillips, Christina Roeder, Bill Schworm, Gregor Smith, Jim Stock and Tony Wirjanto for helpful comments. Financial support of the Social Sciences and Humanities Research Council of Canada is acknowledged.

Abstract

This paper evaluates the finite sample performance of various tests for cointegration by Monte Carlo methods. The evaluation takes place within the linear quadratic model. The results indicate sharp differences in the tests to detect cointegrating relations especially when the cost of adjustment term and the number of regressors are large. Although no single test dominates for all the parameter settings considered, overall the augmented Dickey-Fuller and the Z_α test of Phillips (1987) seem the most reliable in terms of test size and power.

Testing for Cointegration in Linear Quadratic Models

1. Introduction

The theory of cointegration has been a powerful tool in economics to analyze relationships between nonstationary or integrated time series. The setting is very natural since departures from equilibrium relations are permitted in the short run but not in the long run. While many methods are available to test for cointegration (or more correctly non-cointegration), there have been few attempts to compare the finite sample properties of the various tests in an economically meaningful environment[1]. From the perspective of applied work the tests are rather straightforward to calculate so that there is relatively little to recommend one test over another.

The purpose of this paper is to use Monte Carlo methods to evaluate the relative finite sample performance of various tests for cointegration. The evaluation takes place within the class of linear quadratic models that has been extensively employed in the rational expectations literature (see Sargent, 1987). These economies are the only class of optimizing models that give rise to linear decision rules (in the variables) and hence have well-understood properties for the integrated variables. Comparing finite sample test performance in an economic model rather than some (arbitrary) statistical model is then likely to provide the most compelling guidance to the applied researcher. The particular tests of cointegration we compare are

[1] Gonzalo (1989) and Stock and Watson (1991) investigate the finite sample properties of alternative methods of estimating the long run equilibrium relations using statistical models. In this paper paper we are concerned with testing for cointegrating relations rather than evaluating the relative properties of alternative estimators of the cointegrating vector. While the testing for and estimation of cointegrating vectors are of course related, in this study we also consider the size of the tests for cointegration in which case there are no cointegrating relations.

those of Engle and Granger (1987), Hansen (1990), Johansen (1988 and 1990), Park, Ouliaris and Choi (1988), Phillips (1987), Phillips and Ouliaris (1990) and Stock and Watson (1988).

The organization of this paper is as follows. Section 2 develops the linear quadratic model and provides the motivation for the various tests for cointegration. Section 3 describes the tests for cointegration and Section 4 outlines the estimator of the long-run covariance matrix which plays such a prominent role in the various tests. Section 5 discusses the Monte Carlo design and presents the results. Section 6 concludes.

2. Linear Quadratic Models and Cointegration

The linear quadratic model is a popular and tractable dynamic model in which agents minimize a multi-period quadratic cost function (see Sargent, 1987). Agents are assumed to track the long-run target variable y_s^* as given by a static equilibrium theory and choose the actual y_s so as to minimize the weighted sum of the costs of being away from equilibrium $(y_s - y_s^*)$ and the costs of adjustment $(y_s - y_{s-1})$. The problem is:

$$\min_{\{y_s\}} E_t \sum_{s=t}^{\infty} \beta^{s-t} [\delta(y_s - y_s^*)^2 + (y_s - y_{s-1})^2], \tag{2.1}$$

for $s \geq t$, where the expectation is taken with respect to information available to the agent at time t ($\mathbb{F}_t$), $\beta \in (0,1)$ is a discount factor and $\delta > 0$ is a weighting factor (see Kennan, 1979). The static equilibrium relationship is $y_t^* = x_t^T \theta + e_t$, where e_t is a mean zero, independently and identically distributed error with variance σ_e^2, and x_t is a (kx1) vector of forcing variables. We assume that e_t is in $\mathbb{F}_t$ but unknown to the investigating econometrician whose information set is $\mathbb{G}_t \subset \mathbb{F}_t$[2].

The forward solution to (2.1) is:

$$y_t = \lambda\, y_{t-1} + (1 - \lambda)\ (1 - \beta\lambda)\ E_t \sum_{s=t}^{\infty} (\beta\lambda)^{s-t}\, y_s^* \ , \qquad (2.2)$$

where $\lambda < 1$ is the stable root of the Euler equation obtained from the first-order conditions.

The model has been used to explain, for example, the demand for labor by firms (Sargent, 1978, Hansen and Sargent, 1980), the demand for labor and capital by firms (Meese, 1980), the demand and supply of labor (Kennan, 1979, 1988), natural resource extraction (Hansen, Epple, and Roberds, 1985), the demand for money (Cuthbertson and Taylor, 1987, Domowitz and Hakkio, 1990, and Gregory, Smith, and Wirjanto, 1990), and the supply of money (Mercenier and Sekkat, 1988). Hansen and Sargent (1988) have also analyzed and developed software for computable general equilibrium linear quadratic models.

The Wiener-Kolmogorov prediction formula can be used to replace the expectations in (2.2) given the law of motion for the forcing variable (see Sargent, 1987). In this paper we shall be concerned with the case where x_t is a kx1 vector of integrated processes of order 1 denoted I(1):

$$(I - L)\ A(L)x_t \ = \ \varepsilon_t, \qquad (2.3)$$

where $\{\varepsilon_t\}$ is independently and identically distributed with a mean of 0 and variance of Σ and the roots of $A(L) = I-A_1L-\ldots-A_pL^p$ lie outside the unit circle. To simplify the description of the solution of the model we will

---

[2] If we assume e_t is observable by the econometrician then y_t is a deterministic function of the information set. In order to keep the informational issues as simple as possible we assume e_t is serially uncorrelated.

assume that x_t is scalar $(k = 1)$[3]. Given the stochastic process for x_t in (2.3), equation (2.2) can be solved. For instance if $\Delta x_t = \varepsilon_t$ $(k = 1)$ then the error correction model (ECM) can be obtained as:

$$\Delta y_t = (\lambda - 1)(y_{t-1} - \theta x_{t-1}) + (1 - \lambda)\theta\, \Delta x_t + (1 - \beta\lambda)(1 - \lambda)e_t. \tag{2.4}$$

Alternatively with $\Delta x_t = \rho\Delta x_{t-1} + \varepsilon_t$ and $|\rho| < 1$, then:

$$\Delta y_t = (\lambda-1)(y_{t-1}-\theta x_{t-1}) + (1-\lambda)\theta\, \Delta x_t /(1-\rho\lambda\beta) + (1 - \beta\lambda)(1 - \lambda)e_t. \tag{2.5}$$

In general, the solution will depend upon the serial correlation properties of Δx_t. However regardless of the exact nature of (2.3), the following relation always holds:

$$y_t = \theta\, x_t + \eta_t, \qquad t = 1,\ldots,T \tag{2.6}$$

where η_t is a stationary error. Hence y_t and x_t are cointegrated and θ is the cointegrating vector. Tests for cointegration using (2.6) are then applied as a weak test of the model (2.1) without having to know (2.3) with solutions such as (2.4) or (2.5).[4]

The general form for η_t $(k = 1)$ is:

$$\eta_t = [\Psi(L)\ \lambda/(1-\lambda L)]\ \varepsilon_t + [\delta\ \lambda/(1-\lambda L)]\ e_t, \tag{2.7}$$

where $\Psi(L)$ depends upon the nature of x_t in (2.3). For instance $\Delta x_t = \varepsilon_t$, $\Psi(L) = -\theta$ and for $\Delta x_t = \rho\Delta x_{t-1} + \varepsilon_t$, $\Psi(L) = -\theta(1-\rho\beta)/[(1-\rho\beta\lambda)(1-\rho L)]$.

---

[3] However when we develop the tests and perform the Monte Carlo analysis we consider vector x_t.

[4] Single equation estimation of the linear quadratic model with integrated processes have been considered by Dolado, Galbraith and Banerjee (1989) and Gregory, Pagan and Smith (1990) using methods developed by Phillips and Hansen (1990).

We can make some predictions regarding the tests for cointegration in the linear quadratic model by noting the relationship between the relative cost parameter δ and the stable root λ. The stable root $\lambda < 1$ satisfies:

$$\lambda^2\beta+1 = \lambda+\lambda\beta+\lambda\delta, \qquad (2.8)$$

where $\lambda \to 1$ as $\delta \to 0$. That is, as the cost of adjustment gets large (a small δ) the stable root approaches 1 and η_t in (2.6) is nearly integrated. In these circumstances we might expect that tests for cointegration in linear quadratic models (like augmented Dickey-Fuller, see also Phillips and Ouliaris, 1990 and references therein) would encounter difficulties in detecting a cointegrated relation like (2.6) when the stable root (high cost of adjustment) is near unity. Despite the fact that such tests are asymptotically appropriate with serially correlated errors, finite sample evidence in Schwert (1989) for unit root tests suggest a lack of power if η_t is nearly integrated. Unfortunately, applied work has yielded point estimates for the root that have typically been 0.9 or greater (see for example, Meese, 1980; Mendis and Muellbauer, 1982; Nickell, 1984, 1986 and Sargent, 1978).

It is also quite clear that systems approaches like Johansen (1988 and 1990), Phillips and Durlauf (1986), Stock and Watson (1988), and Phillips and Ouliaris (1988 and 1990) may suffer similar problems to the single equation methods when the stable root is near unity. The procedures of Johansen (1988 and 1990) and Stock and Watson (1988) examine the system of equations in vector autoregressive form which for independent Δx_t is:

$$y_t = \lambda\, y_{t-1} + (1-\lambda)\,\theta\, x_{t-1} + (1-\lambda)\,\theta\,\varepsilon_t + (1-\beta\lambda)\,(1-\lambda)\,e_t \qquad (2.9)$$
$$x_t = x_{t-1} + \varepsilon_t\ .$$

Let $z_t = (y_t, x_t)^T$. The tests for cointegration of Johansen (1988 and 1990) and Stock and Watson (1988) for (2.9) consider:

$$z_t = Rz_{t-1} + \nu_t, \tag{2.10}$$

and test whether R = I. Clearly from (2.9) as λ approaches one this restriction is closer to being true and the tests should do poorly. Another multivariate approach due to Phillips and Ouliaris (1988 and 1990) is to examine the long-run covariance matrix of Δz_t (the spectrum of Δz_t at frequency zero) for singularities. With independent Δx_t we have:

$$\Delta z_t = \begin{cases} \Delta y_t = \left(1/1-\lambda L\right)(1-\lambda)(1-\beta\lambda)\ (e_t-e_{t-1}) + \left(1/1-\lambda L\right)\ (1-\lambda)\theta\ \varepsilon_t \\ \Delta x_t = \varepsilon_t\ . \end{cases} \tag{2.11}$$

If $E[e_s\varepsilon_t] = 0$, for all s and t, the long-run covariance matrix of Δz_t denoted by Ω is (with the variance of Δx_t equal to σ_ε^2):

$$\Omega = \begin{bmatrix} \theta^2\ \sigma_\varepsilon^2 & \theta\ \sigma_\varepsilon^2 \\ \theta\ \sigma_\varepsilon^2 & \sigma_\varepsilon^2 \end{bmatrix}, \tag{2.12}$$

which is clearly singular. From (2.11) it is also evident that the closer is λ to one, the more difficult it will be to estimate the long-run relation and hence to detect such singularities as (2.12).

In the next section we describe the tests for cointegration that are used in the Monte Carlo analysis. The descriptions are brief and the interested reader is advised to consult the original sources for further details.

3. Tests for Cointegration

(i) Augmented Dickey Fuller (ADF) Test:

The most widely used cointegration test is the augmented Dickey-Fuller (ADF) t-ratio test (see Said and Dickey, 1984), recommended by Engle and Granger (1987). Its asymptotic properties have been studied by Phillips and Ouliaris (1990). The test is based on the residuals from a cointegrating regression and is constructed to test the null hypothesis of no cointegration by testing the null of a unit root in the residuals against the alternative that the root is less than unity. One first estimates the cointegrating regression (2.6) by ordinary least squares (OLS) and tests the null hypothesis of no cointegration using a scalar unit root test $t(\hat{\alpha})$ on the residuals:

$$\Delta\hat{\eta}_t = \hat{\alpha}\ \hat{\eta}_{t-1} + \sum_{i=1}^{m} \hat{\phi}_i\ \Delta\hat{\eta}_{t-i} + \hat{\nu}_t, \tag{3.1}$$

where the lag length m is chosen sufficiently large in order for $\hat{\nu}_t$ to be serially uncorrelated. The distribution of $t(\hat{\alpha})$ depends upon the number of regressors in (2.6) with asymptotic critical values provided in Engle and Yoo (1987) and Phillips and Ouliaris (1990). In the Monte Carlo work we take $m = 1$ and 6 and denote them as ADF_1 and ADF_6 respectively.

(ii) Phillips's Z_α and Z_t Test:

Closely related to the the ADF tests are those suggested by Phillips (1987) and Phillips and Perron (1988). Equation (2.6) is estimated by OLS, the residuals are obtained and the following test regression is run:

$$\hat{\eta}_t = \hat{\alpha}\ \hat{\eta}_{t-1} + \hat{\zeta}_t\ . \tag{3.2}$$

Again we test the unit root hypothesis on the residuals. The test statistics are:

$$Z_\alpha = T\ (\hat{\alpha}-1)-\tfrac{1}{2}\ [\hat{\omega}^2_\xi-\hat{\sigma}^2_\xi]\ \left(T^{-2}\sum_{t=2}^{T}\hat{\eta}^2_{t-1}\right)^{-1}, \tag{3.3}$$

and

$$Z_t = \left[(\hat{\alpha}-1)\left(\sum_{t=2}^{T}\hat{\eta}^2_{t-1}\right)^{\frac{1}{2}}/\ \hat{\omega}_\xi\right]-\tfrac{1}{2}\ [\hat{\omega}^2_\xi-\hat{\sigma}^2_\xi]\ \left(\hat{\omega}^2_\xi\ T^{-2}\sum_{t=2}^{T}\hat{\eta}^2_{t-1}\right)^{-1/2}, \tag{3.4}$$

where $\hat{\sigma}^2_\xi = T^{-1}\sum_{t=1}^{T}\hat{\zeta}^2_t$ and $\hat{\omega}^2_\xi$ is an estimator of the spectrum of ζ at frequency zero (the long-run variance). In the Monte carlo experiments we estimate the long-run covariance matrix using kernel estimators due to Andrews (1991) and Andrews and Monahan (1990). Exact details of this calculation can be found in Section 4. The critical values for the limiting distribution of (3.3) and (3.4) again depend upon the number of regressors k and can be found in Phillips and Ouliaris (1990).

(iii) Stock and Watson's Minimum Eigenvalue Test:

The next two multivariate tests (Stock and Watson; 1988 and Johansen; 1988 and 1990) are especially useful in determining the number of cointegrating relations in situations where the researcher does not wish to assume that the x's (k>1) themselves are not cointegrated. This generality certainly is an advantage over the other tests where the only possible cointegrating relation is between the y and the x via (2.6).

Let $z_t = [y_t, x_t^T]^T$ and estimate:

$$z_t = \hat{\Pi}\ z_{t-1} + \hat{\nu}_t\ . \tag{3.4}$$

Obtain an estimate of $V = \sum_{i=1}^{\infty} E\left[\nu_t\nu^T_{t-i}\right]$ say using Andrews (1991) and find the k+1 vector of eigenvalues from:

$$\left[T^{-2}\sum z_t z^T_{t-1} - T^{-1}\hat{V}^T\right]\ \left[T^{-2}\sum z_{t-1}z^T_{t-1}\right]^{-1}. \tag{3.5}$$

Let $\hat{\Lambda}_{min}$ be the minimum real of that vector. Under the null of no cointegration the estimated minimum eigenvalue should be insignificantly different from one. The test statistic suggested by Stock and Watson (1988) is:

$$SW = T(\hat{\Lambda}_{min}-1). \qquad (3.6)$$

The critical values for SW depend upon the dimension of z and are in Stock and Watson (1988).

(iv) Johansen's Likelihood Ratio Test:

A closely related test to (3.6), derived by Johansen (1988 and 1990) is a likelihood ratio test for cointegration. This is not to say that the other tests presented in this paper could not be viewed as likelihood ratio tests. However, since the Johansen test is so firmly ground in likelihood theory, we have reserved this label for these tests. Suppose the data generating process for z_t may be written as a m^{th} order vector autoregression:

$$z_t = \Pi_1 z_{t-1} + \Pi_2 z_{t-2} + \ldots + \Pi_m z_{t-m} + \psi_t \ , \qquad (3.7)$$

where ψ_t is independent mean zero with a constant covariance. We may rewrite (3.7) as:

$$\Delta z_t = \Gamma_1 \Delta z_{t-1} + \Gamma_2 \Delta z_{t-2} + \ldots + \Gamma_{m-1} \Delta z_{t-m+1} + \Gamma_m \ z_{t-m} + \psi_t, \qquad (3.8)$$

where $\Gamma_i = -I + \Pi_1 + \ldots + \Pi_i$, $i = 1,\ldots m$ and $\Gamma_m = I-\Pi_1-\ldots-\Pi_m$. The intuition behind the test is simply to test the rank of Γ_m. If Γ_m has rank of zero then the null hypothesis of no cointegration cannot be rejected. The test of Johansen (1988) is a likelihood ratio obtained as:

$$LR =-T \sum_{i=1}^{k+1} \ln \ (1-\hat{\Lambda}_i), \qquad (3.9)$$

where $\hat{\Lambda}$ are the eigenvalues from solving (called squared partial canonical correlations or reduced rank regression):

$$|\Lambda \ S_{mm} - S_{m0} \ S_{00}^{-1} \ S_{0m}| = 0, \qquad (3.10)$$

where we define $z_{0t} = \Delta z_t$, $z_{1t} = [\Delta z_{t-1},\ldots,\Delta z_{t-m+1}]$, $z_{mm} = z_{t-m}$ and the following moment relations:

$$M_{ij} = T^{-1}\sum_{t=1}^{T} z_{it}\, z_{jt}^{T} \qquad i,j = 0,1,\ldots m \tag{3.11}$$

$$S_{ij} = M_{ij} - M_{i1}M_{11}^{-1}M_{1j} \qquad i,j = 0,m\ .$$

The critical values depend on the number of regressors and are in Johansen (1988). In the Monte Carlo analysis we set m = 2 which is sufficient to ensure that ψ_t is serially uncorrelated in all experiments. Notice also that this test is a multivariate unit root test. As mentioned above Johansen's tests may used to test for other possible cointegrating relations in the vector x . We investigate the finite sample performance of this test which is, strictly speaking, not directly comparable to the other tests in this study which only allow for one cointegrating relation between y and the vector x. However Johansen (1990) and Johansen and Juselius (1990) suggest a test which is a special case of (3.11) designed for testing for one cointegrating vector in a system of equations. It involves only the maximum eigenvalue in the vector $\hat{\Lambda}$ in (3.10):

$$LR_1 = -T\ \ln(1-\hat{\Lambda}_{max})\ . \tag{3.12}$$

Compare (3.12) with the Stock-Watson test (3.6). In order to be appropriate for the critical values supplied in Johansen and Juselius (1990, Table A2) a constant has been added to (3.8) so that z_{1t} would also have a one in it.

(v) Park, Ouliaris and Choi's Spurious Regressors Test:

Park, Ouliaris and Choi (1988) have developed a variable addition test in which additional regressors (powers of time trends) are added to a (potentially) cointegrating regression. If the variables do indeed define a cointegrating relation then additional variables should have no explanatory power. On the other hand, if the regression is spurious (no cointegration), results from Phillips (1986) indicate that F tests on additional trend terms should diverge. Park, Ouliaris and Choi (1988) obtain a limiting distribution by dividing the usual F test by the sample size. Unlike all the other tests of cointegration discussed in this paper, the Park, Ouliaris and Choi (1988) test can be formulated with a null hypothesis of no cointegration or a null of cointegration (see also Hansen, 1991). However since we wish to make direct comparisons with the other tests we choose the formulation with the null hypothesis of no cointegration.

The unrestricted least squares regression is:

$$y_t = \sum_{i=0}^{q} \hat{\alpha}_i t^i + x_t^T \hat{\theta} + \hat{\eta}_t, \tag{3.13}$$

and denote the residual sum of squares as RSS_q. The Park, Ouliaris and Choi test statistic is:

$$J(0,q) = (RSS_0 - RSS_q)/RSS_q, \tag{3.14}$$

where RSS_0 is the (restricted) residual sum of squares from regressing y_t on x_t and a constant. Critical values are given in Table 1 of Park, Ouliaris and Choi (1988) and again depend upon the number of regressors. For the Monte Carlo analysis q = 3 and is denoted as J.

(vi) Hansen's Cochrane-Orcutt Technique:

All of the tests for cointegration share the feature that the limiting distribution of the test statistic depends on the number of regressors in the cointegrating relation. The tests proposed by Hansen (1990) are based upon the Cochrane-Orcutt technique and yield limiting distributions which are invariant to the number of regressors. Consider equation (2.6) with the following relation for the error structure:

$$\begin{aligned} y_t &= x_t^T \theta + \eta_t \\ \eta_t &= \rho \eta_{t-1} + \xi_t . \end{aligned} \tag{3.15}$$

Estimate sequentially θ and ρ by OLS. Quasi-difference the data using the estimated $\hat{\rho}$:

$$y_t^* = y_t - \hat{\rho} y_{t-1} \qquad (3.16)$$
$$x_t^* = x_t - \hat{\rho} x_{t-1},$$

and estimate:

$$y_t^* = x_t^{*T} \tilde{\theta} + \tilde{\eta}_t \ . \qquad (3.17)$$

We may iterate this procedure or follow some finite sample modifications suggested in Hansen (1990). In the Monte Carlo work we iterated five times which resulted in the convergence of $\hat{\rho}$. Tests for unit roots can now be applied to the residual $\tilde{\eta}_t$ from the transformed estimated equation (3.17) (instead of $\hat{\eta}_t$ from (2.6)). In the Monte Carlo work we consider ADF_1, Z_α, and Z_t which we denote by $HADF_1$, HZ_α and HZ_t respectively.

The advantage of these tests is that the limiting distribution does not depend on the number of regressors and the distributions for the $HADF_1$ and HZ_t and HZ_α are the the Dickey Fuller t-test and coefficient distribution for the unit root hypothesis respectively. The intuition behind this invariance is that the limiting distribution of $\tilde{\theta}$ in (3.17) converges to a constant and not a random variable as in the since under H_0, $\hat{\rho} \rightarrow 1$ and thus y_t^* and x_t^* are asymptotically first differences. Hansen (1990) has observed that the standard residual based tests suffer a considerable loss of power as the number of regressors increases and that tests based upon $\tilde{\eta}_t$ may be more powerful.[5] We check this statement in the context of the linear quadratic model.

---

[5] In private correspondence Professor Johansen has suggested that the Cochrane-Orcutt procedure may in fact have little local power in certain directions.

(vi) Phillips and Ouliaris's Trace and Variance Ratio Test:

Phillips and Ouliaris (1990) suggest examining the long-run covariance matrix of Δz_t, denoted as Ω, for singularities. The tests they propose do not estimate Ω from Δz_t (which would in fact lead to an inconsistent test, see Phillips and Ouliaris, 1990, Theorem 5.3) but instead are based on the residuals from the following first-order vector autoregression:

$$z_t = \hat{\Pi} z_{t-1} + \hat{\nu}_t. \tag{3.18}$$

Let $\hat{\Omega}$ be the estimated long-run covariance matrix (the estimated spectrum of ν_t at frequency zero)[6]. If the variables are cointegrated then there should be singularities in Ω (Phillips and Ouliaris, 1988). The trace test is:

$$P_z = T \operatorname{tr}\left[\hat{\Omega} M_{zz}^{-1}\right] \quad , \; M_{zz} = T^{-1}\Sigma\, z_t z_t^{T} \; , \tag{3.19}$$

and the variance ratio test is:

$$P_u = T\left[\hat{\Omega}_{yy} - \hat{\Omega}_{xy}^{T}\, \hat{\Omega}_{xx}^{-1}\, \hat{\Omega}_{xy}\right] / \; T^{-1}\Sigma\hat{\eta}_t^2, \tag{3.20}$$

where $\hat{\eta}_t$ is the residual from the cointegrating relation (2.6). The critical values for both (3.19) and (3.20) are in Phillips and Ouliaris (1990) and depend upon the dimensionality of z_t.

---

[6] There are some other tests for singularities in the long-run covariance matrix of $(1-L)z_t$ suggested in Phillips and Ouliaris (1988). A feature of these tests is that the limiting distributions are unknown. We have done some preliminary Monte Carlo work using the following test procedure: (i) obtain the smallest eigenvalue of the estimated long-run covariance matrix (ii) use the empirical distribution function to construct critical values for the smallest eigenvalue when the null hypothesis of no cointegration is true and (iii) test the null of no cointegration with data generated from the linear quadratic model. Results from this test proved to be less powerful than the others considered here and consequently this test was not pursued any further.

The idea motivating (3.19) is that under the null of no cointegration there is no singularity in Ω and hence the ratio should stabilize asymptotically (the numerator is an estimate of $t\Omega$ and the denominator is the corresponding sample moment). P_u in (3.20) tests whether the conditional variance of y given x is significantly different from zero.

4. Estimating Long Run Covariance Matrices

The approach used throughout the Monte Carlo work for estimating long-run covariance matrices is due to Andrews (1991) with some important modifications in Andrews and Monahan (1990). Let V_t be a nx1 vector whose long-run covariance matrix is given by Ω. Prewhiten V_t by a finite vector autoregression. Obtain the residuals from this and use an automatic bandwidth for a kernel estimator of the heteroskedastic-autocorrelation consistent (HAC) variance covariance matrix. Recolor to obtain the estimate of the long-run covariance matrix.

First the prewhitening:

$$V_t = \sum_{r=1}^{b} \hat{A}_r V_{t-r} + V_t^* \qquad t = b+1 \ldots T , \tag{4.1}$$

where $\hat{A}_r$ are nxn parameter estimates and V_t^* are the corresponding residuals (in the Monte Carlo work b = 1). The HAC estimator is in terms of V_t^*:

$$\hat{\Omega}^*(S_T) = \sum_{j=-T+1}^{T-1} k\left(j/S_T\right) \hat{\Gamma}^*(j), \qquad \hat{\Gamma}^*(j) = \begin{cases} T^{-1}\sum_{t=j+1}^{T} V_t^* V_{t-j}^{*T} & \text{for } j\geq 0 \\ T^{-1}\sum_{t=-j+1}^{T} V_{t+j}^* {}^* V_t^T & \text{for } j<0 \end{cases} \tag{4.2}$$

where S_T is the data dependent (automatic) bandwidth and $k(\cdot)$ is the real-valued quadratic spectral kernel:

$$k(x) = 25/(12\pi^2x^2)\left\{\frac{\sin(6\pi x/5)}{6\pi x/5}-\cos(6\pi x/5)\right\}. \tag{4.3}$$

For the quadratic spectral kernel $S_T = 1.3221\ (\hat{\alpha}^*\ T)^{1/5}$ where $\hat{\alpha}^*$ is obtained by regressing V_t^* on V_{t-1}^* with associated coefficient matrix A (nxn) and innovation covariance matrix Σ and then calculating:

$$\hat{\alpha}^* = \left\{\frac{2\ \text{vec}\ \hat{g}^T\ W_T\ \text{vec}\ \hat{g}}{\text{tr}\ W_T\ (I + K_{nn})\ \hat{f}\otimes\hat{f}}\right\}, \tag{4.4}$$

where

$$\hat{f} = 1/2\pi\ (I-\hat{A})^{-1}\hat{\Sigma}\ (I-\hat{A}^T)^{-1} \tag{4.5}$$

$$\hat{g} = 1/2\pi\ (I-\hat{A})^{-3}\left[\hat{A}\hat{\Sigma}+\hat{A}^2\hat{\Sigma}\hat{A}^T+\hat{A}^2\hat{\Sigma}-6\hat{A}\hat{\Sigma}\hat{A}^T+\hat{\Sigma}(\hat{A}^T)^2+\hat{A}\hat{\Sigma}(\hat{A}^T)^2+\hat{\Sigma}\hat{A}^T\right](I-\hat{A}^T)^{-3}.$$

W_T is a n^2xn^2 diagonal weight matrix with 2's for diagonal elements that correspond to diagonal elements of Ω and 1's for diagonal elements that correspond to non-diagonal elements of Ω, vec is the vectorization operator, $\otimes$ is the Kronecker cross-product and K_{nn} is an n^2xn^2 commutation matrix that transforms vec(A) into vec(A^T). Finally the estimate of the long-run covariance matrix is obtained by recoloring:

$$\hat{\Omega} = \hat{D}\ \hat{\Omega}^*(S_T)\ \hat{D}^T \quad \text{and} \quad \hat{D} = \left[I_n-\sum_{r=1}^{b}\hat{A}_r\right]^{-1}. \tag{4.6}$$

To calculate $\hat{V}$ for the Stock-Watson test (3.6) we do not prewhiten but follow Andrews (1991) directly. That is:

$$\hat{V} = = \sum_{j=1}^{T-1} k\left(j/S_T\right)\ \hat{\Gamma}(j) \qquad \hat{\Gamma}(j) = T^{-1}\sum_{t=j+1}^{T}\hat{\nu}_t\ \hat{\nu}_{t-j}^T. \tag{4.7}$$

5. Monte Carlo Design and Results

In order to illustrate the issues in as clear a manner as possible we outline the basic experiment whose results appear in Table 1. The design is similar to Gregory, Pagan and Smith (1990) and West (1986). The computer package used in the analysis is GAUSS386 and the programs are available from the author upon request.

For each experiment we do 2500 replications with observation set $T = 50$, 100 and 200. We first consider the effects from adding regressors $k = 1, \ldots, 4$ with the corresponding vector θ equaling 1. The stable root $\lambda = .8$ (implying a $\delta = .06$), $\beta = .97$, e_t and ε_t are normally and independently distributed with mean zero, $COV[e_t, \varepsilon_s^T] = 0$ for all t and s, $VAR[(1-\lambda)(1-\beta\lambda)\ e_t] = 1$ and $VAR[\varepsilon_t] = I_k$. Thus we start with the situation in which Δx_t is exogenous. The variance for e_t has been scaled up in order to avoid singularities which would be caused by λ approaching unity in the cointegrating relation (see equation (2.11)). As Park, Ouliaris and Choi have documented many of the tests become highly unstable when covariance matrices have diagonal elements of highly unequal size. Moreover we think the case with a (relatively) equivalent contribution from e_t in the cointegrating relation (2.6) is more realistic.

Since we intend to make power comparisons of the various tests to detect cointegrating relations it is essential to adjust for differences in test size. For the basic experiment we calculate the size of the test for each T under the following data generating process:

$$y_t = y_{t-1} + u_t \qquad u_t = \lambda u_{t-1} + \zeta_t \tag{5.1}$$
$$x_t = x_{t-1} + \varepsilon_t,$$

where $\xi_t = (\zeta_t, \varepsilon_t^T) \sim NID(0, I_{k+1})$. Clearly y_t and x_t are not cointegrated. Notice that we calculate the size of the tests for cointegration with a serial correlation in u_t that roughly corresponds to the amount of serial correlation in the cointegrating relation (2.6). Table 1 reports the rejection frequencies at the five percent level of significance[7]. SIZE refers to the rejection frequency at the five percent level when (5.1) is true. Power is determined (using data generated from the linear quadratic model) first (labelled POWER) on comparisons using the size-adjusted tests (critical values are calculated so that each tests has the same rejection frequency when the null hypothesis is true) and then against the asymptotic critical values. These latter values appear in parentheses.

Before detailing the individual test performance it is worthwhile to provide some overall assessment. First the tests produce roughly similar results for $k = 1$ and small sample size $T = 50$. Discrepancies occur as the number of regressors and sample size increase. All tests share the feature that power falls as k increases, although some more sharply than others. Except for the LR and LR_1 tests of Johansen (1988 and 1990), tests sizes are either close to the asymptotic values or undersized by $T = 200$ for all k. This result is comforting in practice as comparisons are almost always against the asymptotic critical values.

The ADF_1 size (power) falls as k increases with a bias towards the null hypothesis of no cointegration. Despite this size distortion the power properties are good. The ADF_6 is basically the same as ADF_1 with slightly better size and sightly worse power (both size-adjusted and against asymptotic values). The Z_α and Z_t follow a similar pattern to the ADF tests: the bias is towards the null but there is still good power (size adjusted and asymptotic) which does not diminish very much as k increases. These tests are the best in this regard.

---

[7]Test results for the one percent and ten percent were also calculated. For the most part these results are qualitatively similar to those at the five percent level and are available upon request.

The LR and LR_1 tests have a tendency to overreject when the null is true and the rejection frequency rises in k. The overrejection occurs even at T = 200 for k = 3 or 4. The size results are somewhat better for the maximum eigenvalue test, LR_1 than the trace test LR which uses all the eigenvalues. Despite this overrejection the size adjusted power is quite good producing similar results to the ADF's, Z_α and Z_t tests. As would be expected given the size results, the power against asymptotic critical values is the highest for the larger k.

The SW test is undersized and has slightly worse power than ADF_1, Z_α and the Z_t especially against the asymptotic critical values for large k. The J test is undersized for all experiments but appears to approach asymptotic size in a monotonic fashion (unlike many of the other tests). The test has the smallest power for k = 1 but the power is remarkably constant as k increases.

Hansen (1990) has suggested that one potential advantage of his test is that power may be better as the number of regressors in the cointegrating regression gets large since the limiting distribution of the tests based upon the Cochrane-Orcutt transformation are invariant to k. For the linear quadratic model the power does indeed stay relatively constant for all the $HADF_1$, HZ_α and HZ_t but these same tests are not as powerful as their untransformed (regressor dependent) counterparts. Also for large k there is some tendency to overreject for $HADF_1$ and HZ_t with T = 50.

Both P_z and P_u are undersized and have comparable size-adjusted power with Z_α and Z_t for k = 1. However both size-adjusted and nominal power drops rapidly as k increases.

This completes the basic experiment and now we consider the performance of the tests under some different parameter settings. To keep the number of experiments down we restrict attention to the situation of $k = 1$ and only the changes that appear in the column heading are made. That is, all other parameters are set as in the basic experiment in Table 1 column 1. Size again is determined according to (5.1) with the appropriate changes noted in the columns of the tables.

In Table 2 rejection frequencies are reported for experiments that have differing amounts of serial correlation. The first two columns have $\lambda = .9$ and $\lambda = .7$ respectively. As expected, test performance deteriorates for $\lambda = .9$ (particularly the Hansen tests and the P_u and P_z up to $T = 100$) but improves for the smaller cost of adjustment ($\lambda = .7$). At sample size $T = 100$ with $\lambda = .9$ we see over a 50 percent fall in power for almost all of the tests. On the other hand, with $\lambda = .7$ and $T = 200$ all of the tests reject the null hypothesis of no cointegration at least 87 percent of the time. Positive serial correlation in Δx_t (column 3 in Table 2, $\rho = .8$) tends to raise the test size and lower power (both size adjusted and nominal). The latter observation is especially true for the ADF, Z_α and Z_t tests. Negative serial correlation ($\rho = -.5$) produces similar results to Table 1 column 1 for all the tests except SW. The SW tests with negative serial correlation reject the null too frequently (15 percent at $T = 200$) and there is a substantial loss of power (size adjusted) at $T = 100$.

In Table 3 we no longer assume that the regressor x_t is strictly exogenous. We know that endogenous regressors creates additional nuisance parameters for statistical inference on the cointegrating vector (see Banerjee, Dolado, Hendry and Smith, 1986; Phillips, 1991; Phillips and Hansen 1990; Saikkonen, 1990 and Stock and Watson, 1991). Four experiments are considered with different settings for the covariances (the variances are unchanged).

Denote the covariance between e and ε as $\sigma_{e\varepsilon}$. The four experiments are $\sigma_{e\varepsilon}$ = .2, -.2, .8 and -.8; the size of the tests are calculated using these values for $\sigma_{\zeta\varepsilon}$ in (5.1). We see that with increased positive covariance the ADF, Z_α, Z_t, P_z and P_u have higher test sizes and some loss of power (size adjusted). Interestingly all of the Hansen tests have the highest power with virtually unchanged test sizes for $\sigma_{e\varepsilon}$ = .8. The J test also has increased in power but to a lesser extent than the Hansen tests. Negative covariance has little effect on the tests both in terms of size and power; the exception is the P_u whose power is less than the size for $\sigma_{e\varepsilon}$ = -.8.

In Table 4 alternative values for the cointegrating vector are considered.[8] The first two columns (with k = 1) double (θ = 2) and halve (θ = .5) the coefficient compared to the base case. The only test to be affected appreciably by changing the magnitude of the cointegrating vector is the P_u test in which power falls (rises) as θ rises (falls).

The last two experiments in Table 4 investigate the loss in power from faulty inclusion. For each case the true cointegrating relation has only one variable but additional I(1) variables (k = 2 and k = 3) are included in the test regression (θ_1 = 1 with all other coefficients set to zero). As expected power falls for both of these experiments compared to Table 1 column 1. The power loss is generally comparable to the corresponding values in Table 1 for k = 2 ($\theta_1 = \theta_2 = 1$) and k = 3 ($\theta_1 = \theta_2 = \theta_3 = 1$); the principal exceptions are the three Hansen tests. Overall, if we compare these results

---

[8] There are no corresponding size results since we used those from the base experiment in Table 1 column 1. However with different values for θ the variance changes and this could effect size.

to those in Table 2, we see that the power loss from additional regressors is much less than the loss due to a high stable root in linear quadratic models.

6. Conclusion

The purpose of this paper has been to evaluate the finite sample performance of various tests for cointegration under the class of linear quadratic models. The results indicate sharp differences in the tests to detect cointegrating relations especially when the cost of adjustment and the number of regressors are large. For all experiments considered we found that no one test dominates in terms of test size and power and hence it is difficult to give clear advice. Nevertheless our overall impression is that at least for the linear quadratic model the ADF tests and the Z_α test appear to be the most reliable in terms of test size and power.

In addition we find that for sample sizes of around one hundred observations (a typical macro data set), with several regressors and high costs of adjustment the tests lack power. Unfortunately economic data seems to be characterized by slow speeds of adjustment (high stable roots). This suggests a flaw in the strategy of pretesting for cointegration and then estimating a linear quadratic model **only if** the null hypothesis of no cointegration is rejected.

Lastly a word of caution. Econometricians are quite familiar with test conflict: we calculate two (or more) tests and find one test rejects the null hypothesis (at some significance level) while another does not. For example in testing linear hypotheses we have the inequality of the Wald, likelihood ratio and Lagrange multiplier tests which on occasion gives rise to test conflict. The evidence presented in this paper suggests the instances of test conflict are likely to be more numerous than is usually encountered.

For this reason we think it is of considerable practical importance to calculate and report several tests for cointegration in applied studies. Of course, tests for cointegration are just one part of the model evaluation process that should be conducted in any specification analysis.

Table 1: Comparisons of Tests for Cointegration: λ = .8 (5% level)

	k = 1		k = 2		k = 3		k = 4	
	SIZE	POWER	SIZE	POWER	SIZE	POWER	SIZE	POWER
ADF_1								
T=50	.04	.38 (.30)	.03	.31 (.17)	.02	.19 (.10)	.01	.20 (.07)
T=100	.03	.89 (.81)	.02	.78 (.54)	.01	.66 (.33)	.01	.53 (.21)
T=200	.03	1.0 (1.0)	.01	1.0 (1.0)	.01	1.0 (.95)	.01	.99 (.83)
ADF_6								
T=50	.06	.14 (.17)	.03	.13 (.09)	.02	.10 (.05)	.01	.09 (.03)
T=100	.05	.49 (.49)	.04	.29 (.26)	.03	.23 (.13)	.02	.18 (.07)
T=200	.05	.96 (.96)	.04	.85 (.83)	.03	.72 (.60)	.02	.58 (.39)
Z_α								
T=50	.02	.58 (.33)	.02	.37 (.19)	.01	.24 (.08)	.00	.21 (.04)
T=100	.01	.98 (.87)	.01	.89 (.63)	.01	.76 (.40)	.00	.60 (.25)
T=200	.01	1.0 (1.0)	.01	1.0 (1.0)	.01	1.0 (.97)	.01	1.0 (.90)
Z_t								
T=50	.07	.41 (.49)	.04	.37 (.31)	.04	.25 (.20)	.03	.24 (.16)
T=100	.04	.94 (.91)	.03	.84 (.73)	.02	.73 (.53)	.02	.61 (.41)
T=200	.04	1.0 (1.0)	.02	1.0 (1.0)	.02	1.0 (.99)	.01	.99 (.94)
LR								
T=50	.12	.21 (.38)	.19	.11 (.34)	.35	.09 (.47)	.55	.07 (.59)
T=100	.08	.80 (.86)	.12	.59 (.77)	.18	.50 (.77)	.26	.40 (.79)
T=200	.07	1.0 (1.0)	.08	1.0 (1.0)	.12	.97 (1.0)	.16	.98 (1.0)
LR_1								
T=50	.12	.12 (.23)	.16	.06 (.17)	.24	.04 (.21)	.35	.05 (.37)
T=100	.08	.58 (.71)	.09	.28 (.44)	.13	.18 (.33)	.17	.35 (.64)
T=200	.08	1.0 (1.0)	.06	.96 (.97)	.09	.84 (.92)	.12	.99 (1.0)
SW								
T=50	.00	.54 (.13)	.00	.24 (.04)	.00	.17 (.00)	.00	.14 (.00)
T=100	.01	.95 (.70)	.01	.60 (.30)	.01	.40 (.14)	.01	.26 (.06)
T=200	.02	1.0 (1.0)	.01	.98 (.95)	.01	.86 (.89)	.01	.71 (.52)
J								
T=50	.01	.16 (.04)	.01	.27 (.06)	.01	.19 (.05)	.01	.20 (.04)
T=100	.02	.30 (.24)	.02	.32 (.13)	.02	.22 (.10)	.02	.19 (.07)
T=200	.03	.48 (.31)	.04	.44 (.32)	.03	.31 (.21)	.03	.31 (.18)

See notes at the bottom of Table 1.

Table 1: Continued

	k = 1		k = 2		k = 3		k = 4	
	SIZE	POWER	SIZE	POWER	SIZE	POWER	SIZE	POWER
$HADF_1$								
T=50	.08	.31 (.40)	.11	.26 (.41)	.14	.21 (.43)	.18	.18 (.45)
T=100	.05	.42 (.42)	.06	.28 (.31)	.08	.23 (.27)	.08	.21 (.27)
T=200	.05	.47 (.47)	.05	.24 (.25)	.06	.28 (.31)	.06	.14 (.15)
HZ_α								
T=50	.02	.52 (.42)	.04	.43 (.36)	.04	.40 (.37)	.06	.31 (.35)
T=100	.01	.51 (.40)	.02	.35 (.27)	.02	.30 (.21)	.01	.28 (.20)
T=200	.02	.52 (.44)	.03	.26 (.21)	.02	.17 (.13)	.02	.13 (.10)
HZ_t								
T=50	.10	.34 (.47)	.13	.27 (.45)	.17	.21 (.48)	.23	.15 (.49)
T=100	.06	.42 (.44)	.06	.30 (.32)	.07	.25 (.29)	.08	.22 (.29)
T=200	.05	.47 (.47)	.05	.25 (.25)	.06	.16 (.17)	.06	.14 (.15)
P_z								
T=50	.02	.22 (.05)	.01	.09 (.01)	.00	.06 (.00)	.00	.03 (.00)
T=100	.01	.86 (.33)	.01	.43 (.11)	.00	.26 (.02)	.00	.15 (.00)
T=200	.01	1.0 (1.0)	.01	.98 (.93)	.01	.83 (.43)	.00	.61 (.18)
P_u								
T=50	.02	.35 (.07)	.01	.20 (.01)	.00	.15 (.00)	.00	.13 (.00)
T=100	.01	.90 (.53)	.01	.46 (.10)	.01	.30 (.02)	.00	.14 (.00)
T=200	.02	1.0 (1.0)	.02	.98 (.82)	.01	.78 (.30)	.01	.40 (.06)

Notes:
SIZE refers to the number of rejections when the null hypothesis of no cointegration is true using the asymptotic critical values. POWER refers to the number of rejections for the size-corrected tests. Beside these in parentheses are rejection frequencies when asymptotic critical values are used. ADF_1 and ADF_6 are augmented Dickey-Fuller tests (Engle and Granger, 1987) with 1 and 6 lags respectively; Z_α and Z_t are the tests of Phillips (1987), LR is the Johansen (1988) likelihood ratio (trace) test, LR_1 is the likelihood ratio test based on the maximal eigenvalue (Johansen, 1990), SW is Stock and Watson (1988) minimum eigenvalue test, J is the cubic trend regression test of Park, Ouliaris and Choi (1988), $HADF_1$, HZ_α and HZ_t use the Hansen (1990) Cochrane-Orcutt procedures, P_z and P_u are the trace and variance ratio tests in Phillips and Ouliaris (1990). For the Z_α, Z_t, HZ_α, HZ_t, P_z and P_u, the long-run covariance estimators are due to Andrews (1990). They are obtained from a prewhitened quadratic spectral kernel with a vector autoregression of order 1 for the prewhitening (Andrews and Monahan, 1990) and an automatic bandwidth estimator which is also a vector autoregression of order 1 (Andrews, 1990). For each experiment there are 2500 replications.

Table 2: Serial Correlation (5% level)

	k = 1							
	$\lambda = .9$		$\lambda = .7$		$\rho = .8$		$\rho = -.5$	
	SIZE	POWER	SIZE	POWER	SIZE	POWER	SIZE	POWER
ADF_1								
T=50	.03	.21 (.13)	.03	.67 (.54)	.06	.06 (.07)	.03	.37 (.25)
T=100	.03	.41 (.30)	.04	.99 (.98)	.06	.19 (.22)	.04	.81 (.74)
T=200	.03	.91 (.84)	.03	1.0 (1.0)	.05	.75 (.76)	.03	1.0 (1.0)
ADF_6								
T=50	.04	.14 (.12)	.05	.22 (.24)	.07	.09 (.13)	.06	.17 (.19)
T=100	.05	.24 (.22)	.05	.65 (.66)	.06	.27 (.29)	.06	.44 (.48)
T=200	.05	.71 (.64)	.06	.99 (.99)	.05	.78 (.77)	.05	.97 (.96)
Z_α								
T=50	.03	.25 (.12)	.01	.88 (.65)	.06	.04 (.05)	.04	.47 (.43)
T=100	.01	.68 (.32)	.02	1.0 (.99)	.04	.24 (.19)	.03	.97 (.90)
T=200	.01	1.0 (.86)	.02	1.0 (1.0)	.04	.84 (.76)	.02	1.0 (1.0)
Z_t								
T=50	.05	.20 (.20)	.04	.82 (.78)	.10	.05 (.10)	.10	.32 (.60)
T=100	.04	.41 (.40)	.04	1.0 (1.0)	.07	.20 (.25)	.08	.89 (.95)
T=200	.04	.96 (.89)	.04	1.0 (1.0)	.06	.76 (.80)	.06	1.0 (1.0)
LR								
T=50	.16	.06 (.21)	.10	.41 (.56)	.19	.34 (.66)	.12	.16 (.31)
T=100	.11	.31 (.47)	.07	.96 (.98)	.11	.84 (.99)	.09	.60 (.77)
T=200	.07	.91 (.94)	.08	1.0 (1.0)	.08	1.0 (1.0)	.06	1.0 (1.0)
LR_1								
T=50	.14	.04 (.13)	.11	.21 (.35)	.17	.19 (.43)	.12	.09 (.19)
T=100	.12	.16 (.29)	.08	.87 (.92)	.11	.84 (.99)	.09	.46 (.59)
T=200	.08	.72 (.83)	.07	1.0 (1.0)	.10	1.0 (1.0)	.08	.99 (1.0)
SW								
T=50	.01	.24 (.04)	.00	.78 (.35)	.00	.32 (.01)	.16	.05 (.37)
T=100	.01	.57 (.20)	.01	1.0 (.98)	.00	.59 (.07)	.17	.18 (.90)
T=200	.01	.95 (.73)	.02	1.0 (1.0)	.00	.87 (.52)	.15	.97 (1.0)
J								
T=50	.00	.49 (.06)	.02	.51 (.22)	.01	.17 (.04)	.01	.53 (.14)
T=100	.00	.41 (.10)	.03	.67 (.46)	.02	.21 (.10)	.02	.63 (.28)
T=200	.03	.48 (.20)	.04	.87 (.80)	.04	.42 (.32)	.03	.75 (.58)

See notes at the bottom of Table 1.

Table 2: Continued

	k = 1							
	λ = .9		λ = .7		ρ = .8		ρ = -.5	
	SIZE	POWER	SIZE	POWER	SIZE	POWER	SIZE	POWER
$HADF_1$								
T=50	.10	.11 (.19)	.08	.60 (.70)	.15	.11(.35)	.08	.24 (.34)
T=100	.07	.10 (.13)	.06	.80 (.81)	.10	.25 (.41)	.06	.34 (.38)
T=200	.05	.09 (.10)	.07	.90 (.91)	.06	.43 (.45)	.05	.42 (.41)
HZ_α								
T=50	.02	.26 (.17)	.02	.80 (.71)	.04	.30 (.29)	.03	.44 (.37)
T=100	.02	.15 (.10)	.02	.87 (.80)	.03	.44 (.38)	.02	.45 (.37)
T=200	.02	.12 (.07)	.03	.93 (.91)	.03	.46 (.42)	.02	.45 (.39)
HZ_t								
T=50	.13	.07 (.23)	.09	.68 (.77)	.19	.06 (.40)	.11	.31 (.42)
T=100	.08	.10 (.13)	.06	.82 (.83)	.10	.26 (.45)	.07	.37 (.40)
T=200	.05	.09 (.10)	.06	.91 (.92)	.07	.43 (.46)	.06	.40 (.41)
P_z								
T=50	.04	.05 (.02)	.01	.60 (.15)	.04	.08 (.04)	.02	.24 (.08)
T=100	.02	.12 (.29)	.01	.99 (.89)	.02	.37 (.16)	.02	.81 (.62)
T=200	.01	.89 (.60)	.02	1.0 (1.0)	.02	.92 (.73)	.02	1.0 (1.0)
P_u								
T=50	.02	.12 (.02)	.01	.58 (.14)	.02	.08 (.01)	.01	.42 (.06)
T=100	.02	.30 (.09)	.01	.99 (.83)	.02	.42 (.12)	.01	.89 (.51)
T=200	.02	.84 (.45)	.02	1.0 (1.0)	.02	.95 (.75)	.02	1.0 (1.0)

Notes: See notes at the bottom of Table 1.

Table 3: Endogenous Regressors (5% level)

	$k = 1$							
	$\sigma_{e\varepsilon} = .2$		$\sigma_{e\varepsilon} = -.2$		$\sigma_{e\varepsilon} = .8$		$\sigma_{e\varepsilon} = -.8$	
	SIZE	POWER	SIZE	POWER	SIZE	POWER	SIZE	POWER
ADF_1								
T=50	.03	.39 (.31)	.08	.26 (.34)	.06	.27 (.33)	.05	.25 (.26)
T=100	.03	.88 (.81)	.07	.30 (.33)	.08	.69 (.83)	.07	.68 (.77)
T=200	.03	1.0 (1.0)	.04	1.0 (1.0)	.08	1.0 (1.0)	.09	1.0 (1.0)
ADF_6								
T=50	.06	.15 (.18)	.05	.17 (.18)	.06	.15 (.20)	.06	.13 (.15)
T=100	.05	.49 (.86)	.05	.47 (.47)	.06	.43 (.51)	.06	.39 (.45)
T=200	.05	.95 (.96)	.05	.95 (.95)	.07	.92 (.97)	.07	.93 (.96)
Z_α								
T=50	.01	.64 (.34)	.02	.58 (.33)	.05	.38 (.37)	.04	.33 (.30)
T=100	.01	.99 (.86)	.01	.97 (.86)	.06	.84 (.89)	.05	.84 (.84)
T=200	.01	1.0 (1.0)	.02	1.0 (1.0)	06	1.0 (1.0)	.06	1.0 (1.0)
Z_t								
T=50	.04	.53 (.49)	.05	.47 (.46)	.09	.37 (.50)	.08	.31 (.43)
T=100	.03	.95 (.91)	.03	.93 (.91)	.09	.82 (.93)	.09	.78 (.90)
T=200	.04	1.0 (1.0)	.04	1.0 (1.0)	.09	1.0 (1.0)	.09	1.0 (1.0)
LR								
T=50	.11	.16 (.30)	.12	.28 (.48)	.11	.11 (.21)	.12	.87 (.95)
T=100	.09	.68 (.78)	.08	.91 (.94)	.08	.49 (.60)	.08	1.0 (1.0)
T=200	.08	1.0 (1.0)	.08	1.0 (1.0)	.07	.98 (.99)	.07	1.0 (1.0)
LR_1								
T=50	.11	.09 (.20)	.11	.17 (.29)	.12	.06 (.15)	.11	.68 (.84)
T=100	.09	.46 (.60)	.09	.72 (.83)	.08	.28 (.40)	.09	1.0 (1.0)
T=200	.08	.99 (.99)	.08	1.0 (1.0)	.08	.92 (.96)	.08	1.0 (1.0)
SW								
T=50	.00	.59 (.14)	.00	.52 (.12)	.02	.34 (.15)	.01	.31 (.13)
T=100	.01	.94 (.71)	.01	.93 (.70)	.03	.86 (.75)	.03	.77 (.66)
T=200	.01	1.0 (1.0)	.01	1.0 (1.0)	.06	1.0 (1.0)	.05	1.0 (1.0)
J								
T=50	.01	.52 (.17)	.01	.37 (.08)	.02	.62 (.28)	.01	.08 (.02)
T=100	.02	.57 (.32)	.02	.39 (.18)	.03	.66 (.52)	.03	.10 (.05)
T=200	.03	.78 (.61)	.04	.60 (.45)	.04	.85 (1.0)	.05	.24 (.23)

See notes at the bottom of Table 1.

Table 3: Continued

	k = 1							
	$\sigma_{e\varepsilon}$ = .2		$\sigma_{e\varepsilon}$ = -.2		$\sigma_{e\varepsilon}$ = .8		$\sigma_{e\varepsilon}$ = -.8	
	SIZE	POWER	SIZE	POWER	SIZE	POWER	SIZE	POWER
$HADF_1$								
T=50	.08	.40 (.51)	.08	.26 (.34)	.09	.57 (.78)	.08	.20 (.26)
T=100	.07	.54 (.58)	.07	.30 (.33)	.06	.99 (.99)	.06	.20 (.21)
T=200	.05	.66 (.67)	.07	.32 (.34)	.05	1.0 (1.0)	.05	.18 (.18)
HZ_α								
T=50	.01	.66 (.52)	.02	.39 (.33)	.04	.90 (.82)	.04	.22 (.21)
T=100	.02	.65 (.57)	.02	.36 (.30)	.02	1.0 (1.0)	.02	.18 (.16)
T=200	.02	.72 (.64)	.02	.34 (.31)	.02	1.0 (1.0)	.02	.16 (.18)
HZ_t								
T=50	.09	.44 (.57)	.10	.30 (.38)	.10	.60 (.88)	.10	.20 (.28)
T=100	.07	.57 (.60)	.07	.31 (.34)	.06	.99 (1.0)	.06	.21 (.22)
T=200	.05	.67 (.67)	.06	.33 (.35)	.06	1.0 (1.0)	.05	.18 (.18)
P_z								
T=50	.02	.25 (.06)	.02	.20 (.05)	.03	.12 (.07)	.04	.09 (.05)
T=100	.02	.85 (.53)	.01	.86 (.51)	.03	.74 (.56)	.03	.64 (.47)
T=200	.02	1.0 (1.0)	.01	1.0 (1.0)	.03	1.0 (1.0)	.03	1.0 (1.0)
P_u								
T=50	.01	.41 (.08)	.01	.16 (.02)	.03	.01 (.00)	.02	.00 (.00)
T=100	.02	.89 (.58)	.01	.58 (.17)	.03	.23 (.07)	.03	.00 (.00)
T=200	.02	1.0 (1.0)	.02	.99 (.91)	.04	.99 (.97)	.03	.00 (.00)

Notes: See notes at the bottom of Table 1.

Table 4: Cointegrating Vectors (5% level)

	k = 1		k = 2	k = 3
	$\theta = 2.0$	$\theta = .5$	$\theta_1=1,\ \theta_2=0$	$\theta_1=1,\ \theta_2=\theta_3=0$
	POWER	POWER	POWER	POWER
ADF_1				
T=50	.37 (.28)	.39 (.30)	.31 (.16)	.26 (.12)
T=100	.86 (.79)	.89 (.82)	.77 (.59)	.67 (.40)
T=200	1.0 (1.0)	1.0 (1.0)	1.0 (.99)	1.0 (.96)
ADF_6				
T=50	.13 (.17)	.15 (.19)	.11 (.10)	.12 (.05)
T=100	.43 (.45)	.48 (.50)	.33 (.28)	.30 (.14)
T=200	.95 (.95)	.95 (.96)	.86 (.83)	.75 (.65)
Z_α				
T=50	.57 (.30)	.63 (.35)	.40 (.18)	.32 (.11)
T=100	.98 (.84)	.98 (.87)	.90 (.65)	.77 (.46)
T=200	1.0 (1.0)	1.0 (1.0)	1.0 (1.0)	1.0 (.98)
Z_t				
T=50	.44 (.44)	.49 (.48)	.38 (.32)	.31 (.25)
T=100	.92 (.89)	.94 (.92)	.83 (.75)	.75 (.60)
T=200	1.0 (1.0)	1.0 (1.0)	1.0 (1.0)	1.0 (.99)
LR				
T=50	.52 (.73)	.12 (.25)	.07 (.26)	.04 (.34)
T=100	.99 (1.0)	.54 (.67)	.39 (.58)	.18 (.48)
T=200	1.0 (1.0)	1.0 (1.0)	.97 (.99)	.83 (.92)
LR_1				
T=50	.32 (.49)	.05 (.16)	.06 (.17)	.05 (.22)
T=100	.95 (.98)	.35 (.47)	.29 (.44)	.17 (.33)
T=200	1.0 (1.0)	.93 (.96)	.97 (.98)	.80 (.90)
SW				
T=50	.52 (.13)	.55 (.13)	.25 (.03)	.15 (.01)
T=100	.92 (.68)	.96 (.71)	.61 (.31)	.46 (.15)
T=200	1.0 (1.0)	1.0 (1.0)	.99 (.96)	.92 (.81)
J				
T=50	.22 (.04)	.63 (.20)	.40 (.10)	.40 (.10)
T=100	.25 (.11)	.66 (.42)	.41 (.21)	.39 (.17)
T=200	.45 (.30)	.85 (.72)	.63 (.42)	.56 (.35)

See notes at the bottom of Table 1.

Table 4: Continued

	k = 1		k = 2	k = 3
	θ = 2.0	θ = .5	θ_1=1, θ_2=0	θ_1=1, θ_2=θ_3=0
	POWER	POWER	POWER	POWER
$HADF_1$				
T=50	.21 (.27)	.47 (.63)	.34 (.51)	.36 (.62)
T=100	.23 (.24)	.80 (.83)	.45 (.52)	.51 (.59)
T=200	.20 (.22)	.94 (.95)	.47 (.48)	.53 (.55)
HZ_α				
T=50	.30 (.24)	.81 (.67)	.63 (.50)	.65 (.59)
T=100	.23 (.20)	.89 (.83)	.58 (.48)	.67 (.55)
T=200	.20 (.18)	.97 (.94)	.54 (.46)	.61 (.52)
HZ_t				
T=50	.23 (.31)	.51 (.73)	.32 (.59)	.29 (.69)
T=100	.23 (.24)	.81 (.85)	.47 (.53)	.51 (.62)
T=200	.21 (.22)	.95 (.95)	.49 (.49)	.55 (.56)
P_z				
T=50	.22 (.05)	.26 (.06)	.09 (.01)	.07 (.00)
T=100	.81 (.48)	.84 (.53)	.47 (.14)	.28 (.04)
T=200	1.0 (1.0)	1.0 (1.0)	.98 (.85)	.83 (.48)
P_u				
T=50	.02 (.00)	.59 (.14)	.32 (.02)	.42 (.00)
T=100	.09 (.01)	.97 (.76)	.68 (.18)	.62 (.11)
T=200	.59 (.19)	1.0 (1.0)	1.0 (.92)	.96 (.79)

See notes at the bottom of Table 1.

References

Andrews, D. K., 1991, Heteroskedasticity and autocorrelation consistent covariance matrix estimation, **Econometrica**, forthcoming.

Andrews, D. K. and J. C. Monahan, 1990, An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator, Cowles Foundation Working Paper No. 942.

Banerjee, A., J. Dolado, D. F. Hendry and G. W. Smith, 1986, Exploring equilibrium relationships in econometrics through static models: some monte carlo evidence, **Oxford Bulletin of Economics and Statistics** 48, 253-277.

Dolado, J., J. W. Galbraith and A. Banerjee, 1989, Estimating Euler equation with integrated series, Applied Economics Discussion Paper series 81, University of Oxford, Oxford.

Domowitz, I. and C. S. Hakkio, 1990, Interpreting an error correction model: partial adjustment, forward-looking behavior, and dynamic international money demand, **Journal of Applied Econometrics** 5, 29-46.

Cuthbertson, K. and M. P. Taylor, 1987, The demand for money: a dynamic rational expectations model, **Economic Journal** 88, 661-692.

Engle, R. F. and C. W. J. Granger, 1987, Cointegration and error correction representation, estimation, and testing, **Econometrica** 55, 251-276.

Engle, R. F. and B. S. Yoo, 1987, Forecasting and testing in cointegrated systems, **Journal of Econometrics** 35, 135-159.

Gonzalo, J. 1989, Comparison of five alternative methods of estimating long run equilibrium relations, University of California, San Diego.

Gregory, A. W., A. R. Pagan, and G. W. Smith, 1990, Estimating linear quadratic models with integrated processes, forthcoming in **Essays in Honour of Rex Bergstrom**, edited by P.C.B. Phillips, Blackwell Publishers.

Gregory, A. W., G. W. Smith and T. Wirjanto, 1990, Synthesis of money-demand and indicator models, forthcoming in **The Bank of Canada Monetary Seminar May 1990, Proceedings**, edited by D. Longworth, F. Caramazza and B. Montador, Bank of Canada.

Hansen, B., 1990, A powerful, simple test for cointegration using Cochrane-Orcutt, Working paper No. 260, University of Rochester, Rochester.

Hansen, B., 1991, Tests for parameter instability in regressions with I(1) processes, unpublished paper, University of Rochester, January.

Hansen, L. P., D. Epple, and W. Roberds, 1985, Linear-quadratic duopoly models of resource depletion, in: T. S. Sargent, ed., **Energy, Foresight and Strategy** , Resources for the Future: Washington D. C. .

Hansen, L.P. and T.J. Sargent, 1980, Formulating and estimating dynamic linear rational expectations models, **Journal of Economic Dynamics and Control**, 2, 7-46.

Hansen, L. P. and T. J. Sargent, 1981, A note on Wiener-Kolmogorov prediction formulas for rational expectations model, **Economics Letters** 8, 255-260.

Hansen, L. P. and T. J. Sargent, 1988, Recursive Models of Dynamic Economies, mimeo, University of Chicago.

Johansen, S., 1988, Statistical analysis of cointegrating vectors, **Journal of Economic, Dynamics and Control**, 12, 231-254.

Johansen, S., 1990, Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models, forthcoming in **Econometrica**.

Johansen, S. and K. Juselius, 1990, Maximum likelihood estimation and inference on cointegration —with applications to the demand for money, **Oxford Bulletin of Economics and Statistics**, 52, 169-210.

Kennan, J., 1979, The Estimation of partial adjustment models with rational expectations, **Econometrica** 47, 1441-1456.

Kennan, J., 1988, An econometric analysis of fluctuations in aggregate labour supply and demand, **Econometrica** 56, 317-333.

Mendis, L. and J. Muellbauer, 1982, Employment functions, returns to scale and expectations, mimeo, Nuffield College, Oxford.

Mercenier, J. and K. Sekkat, 1988, Money stock targeting and money supply - an intertemporal optimization approach (with application to Canada), **Journal of Applied Econometrics** 3, 215-228.

Messe, R., 1980, Dynamic demand schedules for labor and capital under rational expectations, **Journal of Econometrics** 14, 141-158.

Nickel, S., 1984, An investigation of the determinants of manufacturing employment in the United Kingdom, **Review of Economics Studies** 51, 529-557.

Nickel, S., 1986, Dynamic models of labor demand, in: O. Ashenfelter and R. Layard eds., **Handbook of Labor Economics, Vol I**, North Holland, Amsterdam.

Park, J., S. Ouliaris and B. Choi, 1988, Spurious regressions and tests for cointegration, mimeo, Cornell University.

Phillips, P.C.B., 1986, Understanding spurious regressions in econometrics, **Journal of Econometrics**,33, 311-340.

Phillips, P.C.B., 1987, Time series regressions with unit roots, **Econometrica** 55, 271-302.

Phillips, P.C.B., 1991, Optimal inferences in cointegrated systems, **Econometrica** 59, 283-326.

Phillips, P.C.B. and S. N. Durlauf, 1986, Multiple time series with integrated variables, **Review of Economic Studies** 53, 473-396.

Phillips, P.C.B. and B. Hansen, 1990, Statistical inference in instrumental variables regression with I(1) processes, **Review of Economic Studies**, 57, 99-125.

Phillips, P.C.B. and S. Ouliaris, 1988, Testing for cointegration using principal components method, **Journal of Economics Dynamics and Control** 12, 205-130.

Phillips, P.C.B. and S. Ouliaris, 1990, Asymptotic properties of residual based tests for cointegration, **Econometrica** 58, 165-193.

Phillips, P.C.B. and P. Perron, 1988, Testing for a unit root in time series regression, **Biometrika** 75, 335-346.

Said, S. E. and D. A. Dickey , 1984, Testing for unit roots in autoregressive moving average models of unknown orders, **Biometrika** 71, 599-607.

Saikkonen, P., 1990, Estimation of cointegrated systems by an autoregressive approximation, mimeo, University of Helsinki.

Sargent, T. J., 1978, Estimation of dynamic labor demand schedules under rational expectations, **Journal of Political Economy** 86, 1009-1045.

Sargent, T. J., 1987, **Macro-economic theory**, Second Edition , Academic Press, Orlando.

Schwert, W. G., 1989, Test for unit roots: a Monte Carlo investigation, **Journal of Business and Economics Statistics** 7, 147-159.

Stock, J. H. and M. W. Watson, 1988, Testing for common trends, **Journal of the American Statistical Association** 83, 1097-1107.

Stock, J. H. and M. W. Watson, 1991, A simple estimator of cointegrating vectors in higher order integrated systems, Working Paper No. 91-3, Federal Reserve Bank of Chicago, Chicago, Illinois.Research.

West, K. D., 1986, Full-versus limited information of a rational expectations model. Some numerical comparisons, **Journal of Econometrics** 33, 367-385.