AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
http://ageconsearch.umn.edu
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

MONASH 14/92

MONASH UNIVERSITY

AUSTRALIA

GIANNINI FOUNDATION OF AGRICULTURAL ECONOMICS LIBRARY
WITHDRAWN
JAN - 6 1993

A TEST OF THE NULL HYPOTHESIS OF COINTEGRATION

David Harris

and

Brett Inder

Working Paper No. 14/92

December 1992

DEPARTMENT OF ECONOMETRICS

ISSN 1032-3813

ISBN 0 7326 0364 1

A TEST OF THE NULL HYPOTHESIS OF COINTEGRATION

David Harris and Brett Inder

Working Paper No. 14/92

December 1992

DEPARTMENT OF ECONOMETRICS, FACULTY OF ECONOMICS COMMERCE & MANAGEMENT

MONASH UNIVERSITY, CLAYTON, VICTORIA 3168, AUSTRALIA.

A Test of the Null Hypothesis of Cointegration

David Harris and Brett Inder

Department of Econometrics

Monash University

Clayton Vic 3168

The cointegration tests of Engle and Granger (1987) test the null hypothesis of no cointegration. We extend the unit root testing framework of Kwiatkowski et al (1992) to testing the null hypothesis of cointegration. A test is developed which is asymptotically equivalent to the locally best invariant (LBI) test and is applicable to a wide range of non-stationary data generating processes. The asymptotic distribution of our test statistic is found to be free of nuisance parameters, and is dependent only on the number of regressors in the cointegrating regression. We tabulate asymptotic critical values for the test based on this distribution, and report on a small power comparison with the Dickey-Fuller test.

November 1992

1 Introduction

In the literature on cointegrated time series, hypothesis tests developed for testing for cointegration have a null hypothesis of no cointegration. Such tests can be found in Engle and Granger (1987), Phillips and Ouliaris (1990) and Johansen (1988). However, there would seem to be some merit in constructing a test of the null hypothesis of cointegration. In fact, Engle is quoted in Phillips and Ouliaris (1990) as writing "a null hypothesis of cointegration would be far more useful in empirical research than the natural null of non-cointegration". Phillips and Ouliaris proceed to suggest two tests of this hypothesis, but then show that these tests are inconsistent. They leave the problem unsolved.

The merit in testing the null hypothesis of cointegration can be seen if we were building a model where the variables were believed, *a priori*, to be cointegrated. The classic example of aggregate consumption and income could be one such case. If we were to use the standard approach of Engle and Granger (1987) based on augmented Dickey-Fuller tests of the null hypothesis of no cointegration, then we are implicitly saying that we believe the variables are not cointegrated unless the data can convincingly demonstrate otherwise. Instead, we suggest a test of the null hypothesis of cointegration, so that we will believe the variables to be cointegrated unless the data can strongly convince us otherwise. Such an approach may prove to be useful.

Not all applications of cointegration tests may involve some *a priori* beliefs on the presence or absence of cointegration. In such cases, one procedure could make use of tests of both null hypotheses. If we accept no cointegration and reject cointegration, then this is strong evidence for no cointegration. Similarly, if we reject no cointegration and accept cointegration then there is strong evidence for cointegration. There is also the possibility of two inconclusive results. If both null hypotheses are rejected, then a type I error may have occurred in one of the tests.[1] If both null hypotheses are accepted, then at least one of the tests is the victim of a lack of power against the particular data generating process.

This paper can be considered a generalisation of Kwiatkowski et al (1992) (referred to as KPSS), who proposed a unit root test with the null hypothesis of stationarity. We extend their model to include explanatory variables, so that we can estimate a cointegrating regression and test the null hypothesis of stationarity of the residuals. On the issue of model specification, we find that the KPSS model is not a completely general data generating process, and we suggest a more general model. However, in the context of the hypothesis testing problem, this difference is found to be unimportant. Locally best invariant (LBI) tests are constructed for quite restrictive models, but we also construct an asymptotically equivalent test which is applicable to very general data generating processes.

---

1 Another possibility is that the potential error correction term is fractionally integrated but this is not explored in this paper.

In the process of deriving a test that is asymptotically equivalent to the LBI test, we find that the asymptotic distribution of the test is a function of independent standard Brownian motions, and is not dependent on any nuisance parameters. The only variable in the distribution is the number of explanatory variables in the cointegrating regression. Hence we can construct tables of percentage points of the distribution under different numbers of regressors for use as critical values.

In section 2 we discuss two models of non-stationary variables, and in section 3 we derive tests of the null hypothesis of cointegration for these models, which turn out to be identical. To make the tests of section 3 more generally applicable, an asymptotically equivalent test is constructed in section 4, and consistency is proved in section 5. Methods for finding critical values for the test are discussed in section 6, and section 7 provides a brief summary of the procedure for our test. Section 8 presents a simple Monte Carlo comparison of the new test with the Dickey-Fuller test for cointegration. An application of our test to the Fisher effect in Australia, following Inder and Silvapulle (1992), is given for illustrative purposes in section 9.

2 The Models

In this section we suggest two models from which tests for cointegration can be derived. Both are basically extensions of the idea of the cointegrating regression as introduced by Engle and Granger (1987) to allow for the possibility of non-stationarity in the error term.

2.1 Model 1: Independent Random Walk

Model 1 is a simple extension of the model used by KPSS for testing for unit roots, and is specified as

$$y_t = x_t'\beta_0 + \mu_t + u_t \tag{2.1}$$

$$x_t = x_{t-1} + v_t \tag{2.2}$$

$$\mu_t = \mu_{t-1} + w_t, \tag{2.3}$$

where y_t is the dependent variable, x_t is a vector of k non-stationary explanatory variables and μ_t is a random walk in the error term of the cointegrating regression (2.1). If we define $\xi_t = [u_t, v_t', w_t]'$ to be a $k+2$ dimensional process, then we assume that ξ_t is serially independent and $\xi_t \sim N(0, \Sigma)$ where

$$\Sigma = \begin{bmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \Sigma_{22} & 0 \\ 0 & 0 & \sigma_3^2 \end{bmatrix}.$$

With this set of assumptions on ξ_t, we can regard x_t as a set of exogenous variables. Also the random walk component of the error term is independent of the white noise component. If model 1 is the true data generating process then y_t and x_t are not cointegrated, due to the presence of the random walk μ_t in equation (2.1). However, if we could impose the restriction $\sigma_3^2 = 0$ then μ_t would collapse to μ_0 for all t. That is, equation (2.1) would be a cointegrating regression with a constant and a stationary error term. Hence, testing the null hypothesis $\sigma_3^2 = 0$ against the alternative $\sigma_3^2 > 0$ will test the null hypothesis of cointegration against the alternative of no cointegration.

Repeated back substitution for μ_{t-1} in equation (2.3) leads to the following representation for (2.1):

$$y_t = \mu_0 + x_t'\beta_0 + \sum_{k=1}^{t} w_k + u_t. \tag{2.4}$$

If we define $z_t = \left(\sum_{k=1}^{t} w_k\right) + u_t$, then $E(z_t) = 0$ and

$$E(z_i z_j) = \min(i,j)\sigma_3^2 + \sigma_1^2 \qquad i = j$$
$$= \min(i,j)\sigma_3^2 \qquad i \neq j.$$

Now equation (2.1) can be expressed as a linear regression in matrix form,

$$y = X\beta + z; \quad z \sim N(0, \sigma_1^2 Q_1(\lambda)), \tag{2.5}$$

where $\lambda = \sigma_3^2/\sigma_1^2$, $Q_1(\lambda) = A\lambda + I$, and A is the matrix with ij^{th} element $\min(i,j)$. The null hypothesis of $\sigma_3^2 = 0$ is equivalent to testing $\lambda = 0$ in equation (2.5) against $\lambda > 0$. If $\lambda = 0$ then $z_t = u_t$ and is stationary, whereas if $\lambda > 0$ then z_t is non-stationary. The advantage of formulating the model in this way is that this is the exact hypothesis testing problem considered by King and Hillier (1985) for constructing LBI tests. This is considered further in section 3.

One possible objection to the specification of model 1 arises from the independence of u_t and w_t. This places a restriction on the data generating process for z_t when $\sigma_3^2 > 0$, which can be seen by writing down the first difference of z_t as $\Delta z_t = w_t + u_t - u_{t-1}$. This imposes a negative first order autocorrelation on the moving average process generating Δz_t, since $E(\Delta z_t \Delta z_{t-1}) = -\sigma_1^2$. This restriction is expressed in terms of the spectrum of the process by Watson (1986). In this sense, model 1 is not a completely general data generating process. In the next section, we suggest a model which does not impose this restriction, although there is no difference between the models under the null hypothesis.

2.2 Model 2: Perfectly Correlated Random Walk

In model 1, the random walk and stationary components of the error term in equation (2.1) were driven by independent processes. Following the idea of Snyder (1985), we let the processes driving the two components be perfectly correlated. As we shall see, this allows for an unrestricted data generating process. Model 2 is

$$y_t = x_t'\beta_0 + \mu_t + u_t \quad (2.6)$$

$$x_t = x_{t-1} + v_t \quad (2.7)$$

$$\mu_t = \mu_{t-1} + \theta u_t. \quad (2.8)$$

If $\zeta_t = [u_t, v_t']'$ then we assume that ζ_t is a serially independent process with distribution $N(0, \Sigma)$ where

$$\Sigma = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \Sigma_{22} \end{bmatrix}.$$

Under these assumptions, x_t is a set of exogenous explanatory variables. At first glance, imposing perfect correlation between the processes driving the random walk and stationary components of the error of (2.6) may seem restrictive, but as discussed below, this is not the case. The coefficient θ is a scale parameter which reflects the relative size of the effects of a shock on the random walk and stationary components. We do not restrict θ to be positive, so that the direction of the effect of a shock may be different for the random walk and stationary components. Clearly, if $\theta = 0$ then the random walk μ_t collapses to a constant μ_0 and it follows that y_t and x_t are cointegrated. Thus we are interested in the two sided testing problem of $\theta = 0$ against $\theta \neq 0$.

Back substituting in equation (2.8) leads to an equivalent representation for model 2 as given in equation (2.4) for model 1:

$$y_t = \mu_0 + x_t'\beta_0 + \theta \sum_{k=1}^{t} u_k + u_t. \quad (2.9)$$

We can then define the compound error term $z_t = \theta\left(\sum_{k=1}^{t} u_k\right) + u_t$, and we note that $E(z_t) = 0$ and

$$\begin{aligned} E(z_i z_j) &= \min(i,j)\theta^2\sigma_1^2 + 2\theta\sigma_1^2 + \sigma_1^2 \qquad i = j \\ &= \min(i,j)\theta^2\sigma_1^2 + \theta\sigma_1^2 \qquad i \neq j. \end{aligned}$$

Equation (2.6) can now be expressed in matrix form as follows

$$y = X\beta + z; \qquad z \sim N(0, \sigma_1^2 Q_2(\theta)), \quad (2.10)$$

where $Q_2(\theta) = A\theta^2 + (J + I)\theta + I$ and J is a $T \times T$ matrix with ij^{th} element equal to 1. Thus we have expressed the testing problem in the same form as for model 1 in equation (2.5), and the theory of locally best unbiased invariant (LBUI) tests in King and Hillier (1985) is applicable.

We showed in section 2.1 that the data generating process for Δz_t was constrained to have a negative first order autocorrelation. For model 2, we find $\Delta z_t = (1+\theta)u_t - u_{t-1}$ and $E(\Delta z_t \Delta z_{t-1}) = -(1+\theta)\sigma_1^2$. Thus the data generating process for Δz_t is positively autocorrelated if $\theta < -1$ and negatively correlated for $\theta > -1$. Watson (1986) showed that this model is completely general (in terms of the Wold representation), in contrast to model 1. More discussion on this point can be found in Snyder (1985) and Watson (1986).

3 Tests for Cointegration

3.1 LBI Test for Model 1

We want to test $\lambda = 0$ against $\lambda > 0$ in equation (2.5) with $Q_1(\lambda) = A\lambda + I$. Since $Q_1(0) = I$, we can directly apply the results of King and Hillier (1985) for the form of the LBI test. The form of the LBI test in the general case is given by

$$s_1 = \frac{\hat{z}' A_0 \hat{z}}{\hat{z}'\hat{z}} < c_0,$$

where $A_0 = -\partial Q_1(\lambda)/\partial\lambda\,|_{\lambda=0}$ and $\hat{z}$ are the residuals from the estimation of equation (2.5) by OLS. Evaluating the partial derivative gives the following LBI test

$$s_1 = \frac{\hat{z}' A \hat{z}}{\hat{z}'\hat{z}} > c_0, \tag{3.1}$$

which KPSS show is equivalent to

$$s_1 = \frac{\sum_{t=1}^{T} K_t^2}{\hat{\sigma}_z^2} \tag{3.2}$$

up to a power of T, where $K_t = \left(\sum_{i=1}^{t} \hat{z}_i\right)$ and $\hat{\sigma}_z^2 = T^{-1}\hat{z}'\hat{z}$ is the estimator of the error variance. The required normalisation by a power of T is explored further in section 4.

3.2 LBUI Test for Model 2

We consider the test for cointegration in model 2 in more detail. The equation of interest is (2.10) with $Q_2(\theta) = A\theta^2 + (J+I)\theta + I$, and since $Q_2(0) = I$ it is again straightforward to follow King and Hillier (1985) to derive locally best tests. The maximal invariant for this hypothesis testing problem is defined to be

$$w = \frac{P\hat{z}}{\hat{z}'\hat{z}}$$

where $M = I - X(X'X)^{-1}X'$, $\hat{z} = My$, $P'P = M$ and $PP' = I_m$, $m = T - k$. The probability density function of the maximal invariant is given by

$$f(w \mid \theta) = \frac{1}{2}\Gamma\left(\frac{m}{2}\right)\pi^{-\frac{m}{2}} \mid PQ_2(\theta)P' \mid^{-\frac{1}{2}} (w'(PQ_2(\theta)P')^{-1}w)^{-\frac{m}{2}} \tag{3.3}$$

and the log of the likelihood ratio test for $\theta = 0$ against $\theta = \delta$ is

$$\log r(\theta) \mid_{\theta=0} = \log f(w \mid \theta + \delta) \mid_{\theta=0} - \log f(w \mid \theta) \mid_{\theta=0} > c_1 \tag{3.4}$$

for some value of δ. At this stage, we are following the steps to construct a one sided test, and direction of the alternative depends on the sign of δ. In what follows, we will show that the critical region for (3.4) depends on δ^2 (and not δ as in a standard LBI test), and hence that the critical region is unaffected by the choice of alternative hypothesis. Expanding $\log f(w \mid \theta + \delta)$ in a Taylor series about $\delta = 0$ gives

$$\log f(w \mid \theta + \delta) = \log f(w \mid \theta) + \delta \frac{\partial \log f(w \mid \theta)}{\partial \theta} + \frac{\delta^2}{2} \frac{\partial^2 \log f(w \mid \theta)}{\partial \theta^2} + O(\delta^3).$$

Substituting this expansion into equation (3.4) gives the following test:

$$\log r(\theta) \mid_{\theta=0} = \delta \frac{\partial \log f(w \mid \theta)}{\partial \theta} \mid_{\theta=0} + \frac{\delta^2}{2} \frac{\partial^2 \log f(w \mid \theta)}{\partial \theta^2} \mid_{\theta=0} + O(\delta^3) > c_1. \tag{3.5}$$

Evaluating the first derivative gives

$$\frac{\partial \log f(w \mid \theta)}{\partial \theta} \mid_{\theta=0} = \text{constant} + \frac{m}{2} \frac{\hat{z}'(J+I)\hat{z}}{\hat{z}'\hat{z}} = \text{constant} + \frac{m}{2},$$

where the constant term is constant with respect to $\hat{z}$. Clearly the first derivative is constant with respect to $\hat{z}$ and can be transferred to the right hand side of (3.5) and combined with c_1. In this way, the dependence of the test on δ disappears, and we are left only with δ^2. Evaluation of the second derivative gives

$$\frac{\partial^2 \log f(w \mid \theta)}{\partial \theta^2} \mid_{\theta=0} = \text{constant} + \frac{m}{2} \frac{\hat{z}'A\hat{z}}{\hat{z}'\hat{z}}.$$

Now dividing (3.5) (without the first term) by δ^2, and letting δ approach zero to maximise power local to the null hypothesis, we obtain the following test:

$$s_2 = \frac{\hat{z}'A\hat{z}}{\hat{z}'\hat{z}} > c_2. \tag{3.6}$$

Thus the test statistic is the same for models 1 and 2.

4 Relaxing Some Assumptions

The tests developed in section 3 are finite sample tests, in the sense that the optimal local power properties hold for finite samples. However, the assumptions of the models from which the tests are derived are quite restrictive. In particular, imposing exogeneity on the explanatory variables and serial independence on the error process of the cointegrating regression is quite unrealistic in practice. In this section, we relax these assumptions and derive a transformed test statistic that has the same asymptotic distribution as the locally best test statistic. Hence we greatly improve the generality of the model, but at the expense of relying on asymptotic theory.

4.1 A Generalised Model

Since the tests for models 1 and 2 are identical, we concentrate on model 2. Using $\zeta_t = [u_t, v_t']'$ and defining $S_t = [S_{ut}, S_{vt}']' = \left(\sum_{i=1}^{t} \zeta_i\right)$, we assume that ζ_t satisfies the multivariate invariance principle as set out in Theorem 2.1 of Phillips and Durlauf (1987), so that

$$T^{-\frac{1}{2}} S_{[Tr]} \Rightarrow B(r) \text{ as } T \to \infty$$

where $B(r)$ is a $k+1$ dimensional Brownian motion defined on $r \in [0, 1]$ with covariance matrix Ω. The Brownian motion can be partitioned conformably with ζ_t as $B(r) = [B_1(r), B_2(r)']'$, and similarly the covariance matrix is

$$\Omega = \lim_{T \to \infty} T^{-1} E(S_T S_T') = \begin{bmatrix} \omega_1^2 & \omega_{12} \\ \omega_{21} & \Omega_{22} \end{bmatrix}.$$

We can regard Ω as the long run covariance matrix of ζ_t in the sense that it captures both contemporaneous variances and covariances as well as covariances at all lags. It can be expressed as a sum of the contemporaneous and lagged components by writing $\Omega = \Sigma + \Lambda + \Lambda'$ where

$$\Sigma = E(\zeta_t \zeta_t') = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \Sigma_{22} \end{bmatrix}$$

and

$$\Lambda = \sum_{j=1}^{\infty} E(\zeta_t \zeta_{t-j}') = \begin{bmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \Lambda_{22} \end{bmatrix}.$$

The conditions imposed by Phillips and Durlauf (1987) on ζ_t allow for some non-stationarity, but in defining the preceding covariance matrices we have added the assumption of stationarity. This is necessary so that we can interpret u_t as a stationary error term in the cointegrating regression. These conditions on ζ_t generalise model 2 considerably, since x_t is now endogenous, and the error terms u_t and v_t may be any of a wide range of stationary processes, including all stationary ARMA processes. The LBUI test from section 2.2 is no longer exactly LBUI for this generalised model, but we will show that the only difference to its asymptotic distribution is the introduction of correlation between $B_1(r)$ and $B_2(r)$, which can be removed with a non-parametric correction due to Phillips and Hansen (1990). This is described in section 4.3.

4.2 Asymptotic Distribution of the Test

In this section, we derive the asymptotic distribution of the test given by (3.6) when applied to model 2 (that is, not including the generalisations described above). Since u_t and v_t are mutually independent iid processes in model 2, ζ_t is a special case of the general class of processes described above. Thus the convergence of $T^{-1/2}S_{[Tr]}$ to a Brownian motion is assured. The lagged covariance matrix Λ is zero, and the covariance matrices Σ and Ω are block diagonal and equal, so the two components of the Brownian motion $B(r)$ are uncorrelated. To reflect this in the notation, we write the partition of $B(r)$ as $B(r) = [B_{1.2}(r), B_2(r)]$, so $B_{1.2}(r)$ is independent of $B_2(r)$.

The asymptotic distribution is found by considering the form of the test statistic given by equation (3.2). Under the null hypothesis, we know that $z_t = u_t$ and so

$$\hat{z}_i = u_i - [1\ x_i'](X'X)^{-1}X'u, \tag{4.1}$$

where

$$X'X = \begin{bmatrix} T & \sum_{t=1}^{T} x_t' \\ \sum_{t=1}^{T} x_t & \sum_{t=1}^{T} x_t x_t' \end{bmatrix} = \begin{bmatrix} E & F \\ G & H \end{bmatrix}$$

and

$$X'u = \begin{bmatrix} \sum_{t=1}^{T} u_t \\ \sum_{t=1}^{T} x_t u_t \end{bmatrix} = \begin{bmatrix} B \\ C \end{bmatrix}.$$

Then

$$(X'X)^{-1} = \begin{bmatrix} D^{-1} & -D^{-1}FH^{-1} \\ -H^{-1}GD^{-1} & H^{-1} + H^{-1}GD^{-1}FH^{-1} \end{bmatrix}$$

where $D = E - FH^{-1}G$. This leads to the following expression for $K_t = \left(\sum_{i=1}^{t} \hat{z}_i\right)$:

$$K_t = S_{ut} - D^{-1}(B - FH^{-1}C)t - \left(\sum_{i=1}^{t} x_i'\right)(H^{-1}C - H^{-1}GD^{-1}(B - FH^{-1}C)) \tag{4.2}$$

by summing (4.1). We can now consider the behaviour of each of these terms as $T \to \infty$.

The notation will be simplified in the following results by assuming the integrals are taken from 0 to 1 and with respect to r unless stated otherwise, and by omitting the argument r from the Brownian motions. Lemma 2.1 of Park and Phillips (1988) provides a number of useful intermediate results.

$$T^{-\frac{3}{2}}F \Rightarrow \int B_2' \quad T^{-\frac{3}{2}}G \Rightarrow \int B_2 \quad T^{-2}H \Rightarrow \int B_2B_2'$$

$$T^{-\frac{1}{2}}B \Rightarrow B_{1.2}(1) \quad T^{-1}C \Rightarrow \int B_2 dB_{1.2}$$

We will also make use of the following two results:

$$T^{-1}t \Rightarrow r \qquad T^{-\frac{3}{2}}\sum_{i=1}^{t} x_i' \Rightarrow \int_0^r B_2(s)'ds.$$

Now we can obtain the limiting behaviour of each of the terms in equation (4.2).

$$T^{-\frac{1}{2}}S_{ut} \Rightarrow B_{1.2}(r)$$

$$T^{-1}D \Rightarrow 1-\int B_2'\left(\int B_2B_2'\right)^{-1}\int B_2$$

$$T^{-\frac{1}{2}}(B-FH^{-1}C) \Rightarrow B_{1.2}(1)'-\int B_2'\left(\int B_2B_2'\right)^{-1}\left(\int B_2 dB_{1.2}\right)$$

$$T^{\frac{1}{2}}H^{-1}G \Rightarrow \left(\int B_2B_2'\right)^{-1}\int B_2$$

$$TH^{-1}C \Rightarrow \left(\int B_2B_2'\right)^{-1}\left(\int B_2 dB_{1.2}\right)$$

These limits can be combined to obtain the limiting distribution of K_t. However, these limits contain implicit nuisance parameters in the form of the variances of the Brownian motions $B_{1.2}$ and B_2 given by $\omega_{1.2}^2$ and Ω_{22} respectively. To show this, we define $W(r)=\Omega^{-\frac{1}{2}}B(r)$ to be standard Brownian motion, since its covariance matrix is the identity. Since Ω is block diagonal for this case, we can also define the elements of the partitioned standard Brownian motion to be $W_{1.2}(r)=\omega_1^{-1}B_{1.2}(r)$ and $W_2(r)=\Omega_{22}^{-\frac{1}{2}}B_2(r)$.

For convenience, we define the following functions of standard Brownian motions:

$$P=1-\int W_2'\left(\int W_2W_2'\right)^{-1}\int W_2$$

$$Q=W_{1.2}(1)'-\int W_2'\left(\int W_2W_2'\right)^{-1}\int W_2 dW_{1.2}$$

$$R=\left(\int W_2W_2'\right)^{-1}\int W_2$$

$$S=\left(\int W_2W_2'\right)^{-1}\int W_2 dW_{1.2}$$

Combining the limits from the previous paragraph, and replacing Brownian motions by standard Brownian motions gives the following limit for K_t:

$$T^{-\frac{1}{2}}K_t \Rightarrow \omega_1 V(r), \tag{4.3}$$

where

$$V(r) = W_{1.2}(r) - P^{-1}Qr - \left(\int_0^r W_2(s)'ds\right)(S - RP^{-1}Q)].$$

Now we can evaluate the limit of the numerator of (3.2) as follows:

$$T^{-2}\sum_{t=1}^{T} K_t^2 = T^{-1}\int_0^1 K_{[Tr]}^2 dr$$

$$\Rightarrow \omega_1^2 \int_0^1 V(r)^2 dr. \tag{4.4}$$

The bottom line of (3.2) is the estimate of the error variance $\hat{\sigma}_z^2 = T^{-1}\left(\sum_{t=1}^{T} \hat{z}_t^2\right)$, which will be a consistent estimator of σ_1^2 since $z_t = u_t$ under the null hypothesis. Also $\omega_1^2 = \sigma_1^2$ since u_t is iid, so we can say

$$\hat{\sigma}_z^2 \Rightarrow \omega_1^2. \tag{4.5}$$

Now the asymptotic distribution of the test statistic can be found by the ratio of (4.4) and (4.5).

$$s_2 = \frac{T^{-2}\sum_{t=1}^{T} K_t^2}{\hat{\sigma}_z^2} \Rightarrow \int V(r)^2 dr. \tag{4.6}$$

4.3 Correction for Endogenous Regressors in the Generalised Model

When applied to the generalised model described in section 4.1, the LBUI test no longer has the same finite sample optimal power properties. In this section, we consider the asymptotic behaviour of the test under the more general assumptions that u_t and v_t are correlated stationary processes.

Following the derivation in section 4.2, equation (4.2) is unchanged as the expression for K_t. Since the elements of the Brownian motion $B(r) = [B_1(r), B_2(r)']'$ are no longer independent, we have new limits for B and C.

$$T^{-\frac{1}{2}}B \Rightarrow B_1(1) \quad T^{-1}C \Rightarrow \int B_2 dB_1 + \delta_{21},$$

where $\Delta = \Sigma + \Lambda$ is a matrix partitioned conformably with Ω, Σ, Λ. Thus we have two differences in the limits introduced by the generalised model - correlation between the Brownian motions and a related nuisance parameter δ_{21}. To remove these effects, we make use of the method of fully modified OLS suggested by Phillips and Hansen (1990). Notice that both of the problems are due to the endogeneity of x_t, not to the serial correlation in u_t and v_t. The effect of the serial correlation will be discussed later in the section.

To deal with the endogenous regressors, we consider the modified error process $u_t^+ = u_t - \omega_{12}\Omega_{22}^{-1}v_t$. If $\zeta_t^+ = [u_t^+, v_t]$ and $S_t^+ = \left(\sum_{i=1}^{t} \zeta_i^+\right)$ then the long run covariance matrix of ζ_t^+ is

$$\Omega^+ = \lim_{T \to \infty} T^{-1} E(S_T^+ S_T^{+\prime}) = \begin{bmatrix} \omega_{1.2}^2 & 0 \\ 0 & \Omega_{22} \end{bmatrix},$$

where $\omega_{1.2}^2 = \omega_1^2 - \omega_{12}\Omega_{22}^{-1}\omega_{21}$. Now $T^{-\frac{1}{2}} S_{[Tr]} \Rightarrow B(r)$ where $B(r) = [B_{1.2}(r), B_2(r)']'$. We use $B_{1.2}(r)$ since the covariance matrix of the Brownian motion $B(r)$ is block diagonal. Also, the following limits containing u_t^+ hold.

$$T^{-\frac{1}{2}} \sum_{t=1}^{T} u_t^+ \Rightarrow B_{1.2}(1)$$

$$T^{-1} \sum_{t=1}^{T} x_t u_t^+ \Rightarrow \int B_2 dB_{1.2} + \delta_{21}^+,$$

where $\delta_{21}^+ = \delta_{21} - \Delta_{22}\Omega_{22}^{-1}\omega_{21}$.

Even with endogenous regressors, OLS estimation of the cointegrating regression (under the null hypothesis) is consistent. Taking the estimated residuals $\hat{z}_t$ from this regression, we can form $\hat{\zeta}_t = [\hat{z}_t, \Delta x_t']'$ and obtain consistent estimates of Ω and Δ by

$$\hat{\Omega} = T^{-1} \sum_{t=1}^{T} \hat{\zeta}_t \hat{\zeta}_t' + T^{-1} \sum_{k=1}^{l} w(k,l) \sum_{t=k+1}^{T} (\hat{\zeta}_{t-k}\hat{\zeta}_t' + \hat{\zeta}_t \hat{\zeta}_{t-k}')$$

$$\hat{\Delta} = T^{-1} \sum_{k=0}^{l} \sum_{t=k+1}^{T} \hat{\zeta}_t \hat{\zeta}_{t-k}',$$

where $w(k,l)$ are a set of weights to ensure the postive-definiteness of the estimated matrix (see Newey and West (1987) for example), and the truncation parameter l increases at a slower rate than T (say $O(T^{1/2})$). Then if we define the transformed variable $y_t^+ = y_t - \hat{\omega}_{12}\hat{\Omega}_{22}^{-1}\Delta x_t$, we can write the cointegrating regression as

$$y_t^+ = \mu_0 + x_t'\beta_0 + (u_t - \hat{\omega}_{12}\hat{\Omega}_{22}^{-1}v_t), \tag{4.7}$$

and the error term in this regression is asymptotically equivalent to u_t^+. The Phillips-Hansen fully modified OLS estimator is then defined to be

$$\beta^{+} = (X'X)^{-1}(X'y^{+} - e_k T\hat{\delta}_{21}^{+}), \tag{4.8}$$

where $e_k = [0, I_k]'$ and $\hat{\delta}_{21}^{+} = \hat{\delta}_{21} - \hat{\Delta}_{22}\hat{\Omega}_{22}^{-1}\hat{\omega}_{21}$. The extra term involving $\hat{\delta}_{21}^{+}$ in included to remove the nuisance parameter δ_{21}^{+} from the asymptotic distribution of β^{+}. This correction will also remove the nuisance parameter from the asymptotic distribution of our test statistic.

We can now derive the asymptotic distribution of the test in equation (3.2) using the residuals obtained from the fully modified OLS regression. The estimated residuals will be given by

$$\hat{u}^{+} = u^{+} - X(X'X)^{-1}(X'u^{+} - e_k T\hat{\delta}_{21}^{+}), \tag{4.9}$$

where

$$X'u^{+} - e_k T\hat{\delta}_{21}^{+} = \begin{bmatrix} \sum_{t=1}^{T} u_t^{+} \\ \sum_{t=1}^{T} x_t u_t^{+} - T\hat{\delta}_{21}^{+} \end{bmatrix} = \begin{bmatrix} B^{+} \\ C^{+} \end{bmatrix}.$$

The limits of these terms are easily found and they are not dependent of the nuisance parameter δ_{21}^{+} and involve only independent Brownian motions.

$$T^{-\frac{1}{2}}B^{+} \Rightarrow B_{1.2}(1)$$

$$T^{-1}C^{+} = T^{-1}\sum_{t=1}^{T} x_t u_t^{+} - \hat{\delta}_{21}^{+} \Rightarrow \int B_2 dB_{1.2}$$

Notice the these terms have the same limiting distributions as B and C for model 2 (see section 4.2). Now we can define $K_t^{+} = \left(\sum_{i=1}^{t} \hat{u}_i^{+}\right)$, and following the steps leading to equation (4.2) we obtain

$$K_t^{+} = S_u^{+} - D^{-1}(B^{+} - FH^{-1}C^{+})t - \left(\sum_{i=1}^{t} x_i'\right)(H^{-1}C^{+} - H^{-1}GD^{-1}(B^{+} - FH^{-1}C^{+})). \tag{4.10}$$

Those terms without a "+" superscript are left unchanged by fully modified OLS, since they involve only functions of x_t. We can find the limits of each term in (4.10), and those of interest are

$$T^{-\frac{1}{2}}S_u^{+} \Rightarrow B_{1.2}(r)$$

$$T^{-\frac{1}{2}}(B^{+} - FH^{-1}C^{+}) \Rightarrow B_{1.2}(r) - \int B_2'\left(\int B_2 B_2'\right)^{-1}\left(\int B_2 dB_{1.2}\right)$$

$$TH^{-1}C^{+} \Rightarrow \left(\int B_2 B_2'\right)^{-1}\left(\int B_2 dB_{1.2}\right).$$

Hence the limit of each term of K_t^{+} is the same as the limit of each term of K_t when constructed from the restricted model 2. Thus the limit of the top line of (3.2) using K_t^{+} will be the same as given in equation (4.4) for model 2.

$$T^{-2}\sum_{t=1}^{T} K_t^{+2} \Rightarrow \omega_{1.2}^2 \int_0^1 V(r)^2 dr. \qquad (4.11)$$

The one difference between (4.11) and (4.4) is that in equation (4.11) $\omega_{1.2}^2$ is not equal to $\sigma_{1.2}^2$ since u_t^+ is not iid in the general case. If, however, we define the transformed test statistic

$$s_2^+ = \frac{T^{-2}\sum_{t=1}^{T} K_t^{+2}}{\hat{\omega}_{1.2}^2}, \qquad (4.12)$$

where $\hat{\omega}_{1.2}^2$ is a consistent estimator of $\omega_{1.2}^2$, then the asymptotic distribution of s_2^+ is

$$s_2^+ \Rightarrow \int_0^1 V(r)^2 dr. \qquad (4.13)$$

That is, s_2^+ for the general model has the same asymptotic distribution as s_2 for model 2. Hence we have a test which is asymptotically equivalent to the LBUI test, and is applicable to a wide range of data generating processes.

5 Consistency

In this section, we consider the asymptotic behaviour of the test under the alternative hypothesis. We refer to equation (2.9) as the model under the alternative hypothesis, where $\theta \neq 0$. Defining $u_t^a = \theta S_{ut} + u_t$ we obtain the following regression model

$$y = X\beta + u^a, \qquad (5.1)$$

from which the test statistic s_2 can be constructed. The OLS residuals from equation (5.1) are

$$\hat{u}^a = u^a - X(X'X)^{-1}X'u^a, \qquad (5.2)$$

where

$$X'u^a = \begin{bmatrix} \theta \sum_{t=1}^{T} S_{ut} + S_{uT} \\ \theta \sum_{t=1}^{T} x_t S_{ut} + \sum_{t=1}^{T} x_t u_t \end{bmatrix} = \begin{bmatrix} K + B \\ L + C \end{bmatrix}.$$

If we define $K_t^a = \left(\sum_{i=1}^{t} \hat{u}_i^a\right)$ to be the partial sum process under the alternative hypothesis, then

$$K_t^a = S_{ut} - D^{-1}(B - FH^{-1}C)t - \left(\sum_{i=1}^{t} x_i'\right)(H^{-1}C - H^{-1}GD^{-1}(B - FH^{-1}C))$$

$$+\theta\left[\sum_{i=1}^{t} S_{ui} - D^{-1}(K - FH^{-1}L)t - \left(\sum_{i=1}^{t} x_i'\right)(H^{-1}L - H^{-1}GD^{-1}(K - FH^{-1}L))\right]$$

$$= K_t + \theta M_t.$$

From section 4, we know that K_t is $O_p(T^{1/2})$. Thus we are only concerned with finding the order of probability of M_t. The orders of the terms of M_t can be found by noting that K, F, G and $\sum_{i=1}^{t} x_i$ are $O_p(T^{3/2})$, L and H are $O_p(T^2)$, and D and t are $O_p(T)$. Then $\sum_{i=1}^{t} S_{ui}$, $D^{-1}(K - FH^{-1}L)t$, and $\left(\sum_{i=1}^{t} x_i'\right)(H^{-1}L - H^{-1}GD^{-1}(K - FH^{-1}L))$ are $O_p(T^{3/2})$. Thus M_t, and hence K_t^a, are $O(T^{3/2})$. The top line of s_2 under the null hypothesis is $T^{-2}\sum_{t=1}^{T} K_t^{a2}$ which is $O_p(T^2)$. From Phillips (1987) or KPSS, we know that the long run variance estimator $\hat{\omega}_{1.2}^2$ is $O_p(lT)$, and so under the alternative hypothesis s_2 is $O_p(T/l)$. Since $l \to \infty$ at a lesser rate than $T \to \infty$ (typically $l = O(T^{1/2})$), we deduce that $s_2 \to \infty$ under the alternative hypothesis, and hence the test is consistent.

6 Critical Values

As can be seen from equation (4.13), the asymptotic distribution of the test statistic s_2^+ is a function of $V(r)$ only, which in turn is dependent only on two independent standard Brownian motions. In particular, we have been able to produce a test which does not depend on nuisance parameters such as $\omega_{1.2}$, Ω_{22} or δ_{21}. The only way in which the asymptotic distribution depends on the model specification is that the number of explanatory variables in the cointegrating regression determines the dimensions of $W_2(r)$. Hence we can obtain sets of asymptotic critical values for various numbers of regressors. These are found by simulation of (4.13) using a GAUSS program with a sample size of 4000 and 50000 replications for 1 to 5 explanatory variables. The results of this simulation are reported in Table 1.

Table 1 Asymptotic Critical Values

Number of Regressors (Excluding Constant)	10%	5%	1%
1	0.2335	0.3202	0.5497
2	0.1617	0.2177	0.3727
3	0.1203	0.1590	0.2756
4	0.0929	0.1204	0.1983
5	0.0764	0.0972	0.1560

7 Summary

In this section we give a summary of the testing procedure that has been derived in the paper. First apply OLS to the regression

$$y_t = \mu_0 + x_t'\beta_0 + z_t, \tag{7.1}$$

and obtain the estimated residuals $\hat{z}_t$. Form $\zeta_t = [\hat{z}_t, \Delta x_t']'$ and calculate the estimated covariance matrices

$$\hat{\Omega} = \begin{bmatrix} \hat{\omega}_1^2 & \hat{\omega}_{12} \\ \hat{\omega}_{21} & \hat{\Omega}_{22} \end{bmatrix} = T^{-1}\sum_{t=1}^{T}\zeta_t\zeta_t' + T^{-1}\sum_{k=1}^{l} w(k,l)\sum_{t=k+1}^{T}(\zeta_{t-k}\zeta_t' + \zeta_t\zeta_{t-k}') \tag{7.2}$$

$$\hat{\Delta} = \begin{bmatrix} \delta_{11} & \delta_{12} \\ \delta_{21} & \hat{\Delta}_{22} \end{bmatrix} = T^{-1}\sum_{k=0}^{l}\sum_{t=k+1}^{T}\zeta_t\zeta_{t-k}', \tag{7.3}$$

where $w(k,l) = 1 - k/(l+1)$. Then calculate

$$y_t^+ = y_t - \hat{\omega}_{12}\hat{\Omega}_{22}^{-1}\Delta x_t$$

and

$$\delta_t^+ = \delta_{21} - \hat{\Delta}_{22}\hat{\Omega}_{22}^{-1}\hat{\omega}_{21}$$

and re-estimate the cointegrating regression using the fully modified OLS estimator

$$\beta^+ = (X'X)^{-1}(X'y^+ - e_k T\hat{\delta}_{21}^+), \tag{7.4}$$

where $e_k = [0, I_k]'$. Obtain the estimated residuals $\hat{u}_t^+$ from this regression and construct the test statistic

$$s_2^+ = \frac{T^{-2}\sum_{t=1}^{T} K_t^{+2}}{\hat{\omega}_{1.2}^2}, \tag{7.5}$$

where $K_t^+ = \sum_{i=1}^{t}\hat{u}_i^+$ and $\hat{\omega}_{1.2}^2 = \hat{\omega}_1^2 - \hat{\omega}_{12}\hat{\Omega}_{22}^{-1}\hat{\omega}_{21}$.

Critical values for this test statistic are provided in Table 1 above. This procedure provides a test for the null hypothesis of cointegration which is consistent and is asymptotically equivalent to the locally best invariant test.

8 Monte Carlo Comparison

In this section we examine briefly the power properties of our test by comparing its ability to distinguish between null and alternative hypothesis with that of the Dickey-Fuller test. A power comparison between these two tests is difficult, as they are derived from different assumed data generating processes (DGP's), and have opposite null and alternative hypotheses. To resolve these difficulties, the Monte Carlo study is undertaken as follows.

We first specify a DGP as in equations (2.6) to (2.8), as designed for the s_2^+ test, with $T = 100$ and $k = 1$. A 10% critical value for the s_2^+ test is taken from Table 1 (0.2335). For the Dickey-Fuller test, we generate 10000 sample under the hypothesis of cointegration ($\theta = 0$), and compute Dickey-Fuller t statistics for cointegration. A "critical value" of -8.89 is found where 10% of statistics greater than this value (leading to a conclusion of no cointegration). Both tests thus have a 10% chance of concluding that there is no cointegration when $\theta = 0$. "Powers" are then calculated for a range of values of θ under the alternative, with 10000 replications. These results are given in the first half of Table 2.

To ensure a fair opportunity is given to each test, further results were obtained with a DGP for which the Dickey-Fuller test is designed, namely

$$y_t = x_t'\beta_0 + \mu_t \tag{8.1}$$

$$\mu_t = \rho\mu_{t-1} + u_t. \tag{8.2}$$

For the Dickey-Fuller test, a 10% critical value for testing the null of no cointegration is found in Phillips and Ouliaris (1990) (-3.0657), and for the s_2^+ test we use 10000 replications under the hypothesis of no cointegration ($\rho = 1$) to find a "critical value" of 0.7543 where 10% of the s_2^+ statistics are less than this value (leading to a conclusion of cointegration). Both test thus have a 10% chance of concluding there is cointegration when $\rho = 1$. "Powers" are then calculated for values of ρ under the alternative, with 10000 replications. These results are given in the second half of Table 2.

It is difficult to draw definitive conclusions from such a limited Monte Carlo study, but Table 2 does suggest a couple of points worthy of note. Firstly, it is not surprising that each test performs at its best when the DGP follows that for which the test is designed. The s_2^+ test shows clear domination over Dickey-Fuller in the error components DGP, particularly for small values of θ. Likewise, the autoregressive DGP favours the Dickey-Fuller test, especially as ρ moves further from one. The s_2^+ test actually dominates for ρ close to one, but this is reversed with smaller ρ.

Given that we do not know which DGP is more realistic, it is inappropriate to "choose" between the two tests on the basis of power. What we can say, though, is that deriving a test which reverses the null and alternative hypotheses has not obviously hurt the capacity to distinguish between cointegration and no cointegration.

Table 2 Powers of Tests[a]

DGP: Equations (2.6), (2.7), (2.8) Null Hypothesis: Cointegration						
	$\theta = 0$	$\theta = 0.05$	$\theta = 0.1$	$\theta = 0.15$	$\theta = 0.2$	$\theta = 0.25$
s_2^+	0.103	0.281	0.519	0.673	0.764	0.839
Dickey-Fuller	0.100	0.164	0.344	0.567	0.718	0.837
DGP: Equations (7.1), (7.2) Null Hypothesis: No Cointegration						
	$\rho = 1$	$\rho = 0.95$	$\rho = 0.9$	$\rho = 0.85$	$\rho = 0.8$	$\rho = 0.75$
s_2^+	0.100	0.235	0.386	0.528	0.643	0.720
Dickey-Fuller	0.105	0.181	0.375	0.642	0.866	0.966

[a] In both DGP's the innovations u_t are iid $N(0,1)$, so the Dickey-Fuller procedure is not augmented by any lags of the differenced residuals, and the s_2^+ test has a lag truncation parameter (l) of zero.

9 Application

Inder and Silvapulle (1992) used cointegration methods to test if the Fisher effect applies in Australia. One question related to this is whether nominal interest rates and inflation are cointegrated. Consider the equation

$$i_t = \beta_0 + \beta_1 \pi_t + u_t \tag{8.1}$$

where i_t is the nominal interest rate and π_t is the inflation rate. Inder and Silvapulle (1992) treat (8.1) as a long run relationship which implies that the interest rate and the inflation rate are cointegrated. They use the augmented Dickey-Fuller test to test for cointegration between the series as represented by the change in the log of CPI for inflation and the bank accepted bill rate for interest rates. The estimated parameters for equation (8.1) were $\beta_0 = 7.2275$ and $\beta_1 = 0.4387$, and the Dickey-Fuller t statistic was -2.5921. Compared with a 10% critical value of -2.84, it can be seen that the null hypothesis of no cointegration could not be rejected.

However, it is possible that the failure to reject the null hypothesis was due only to lack of power of the Dickey-Fuller test, rather than true non-cointegration. To check this, we computed s_2^+ to test the null hypothesis of cointegration. We chose the lag length l by following Inder (1991), who suggested choosing the largest significant lag in the autocorrelation function of the OLS residuals. For this application, this gave a choice of $l = 4$. We obtained $s_2^+ = 1.6674$, which when compared to the 1% critical value of 0.5497 allows rejection of the null hypothesis of cointegration. This provides further evidence for the results of Inder and Silvapulle (1992) rejecting the cointegration hypothesis.

References

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

Inder B. (1991) "Some Strong Evidence for a Unit Root in Economic Time Series" Mimeo, Monash University.

Inder B. and Silvapulle P. (1992) "Does the Fisher Effect Apply in Australia?" Monash University Working Paper No. 5/92.

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

Kwiatkowski D., Phillips P.C.B., Schmidt P. and Shin Y. (1992) "Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root: How sure are we that economic time series have a unit root?" *Journal of Econometrics* 54, 159-178.

Park J.Y. and Phillips P.C.B. (1988) "Statistical Inference in Regressions with Integrated Processes: Part 1" *Econometric Theory* 4, 468-497.

Phillips P.C.B. and Durlauf S.N. (1986) "Multiple Time Series Regression with Integrated Processes" *Review of Economic Studies* 53, 473-495.

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

Phillips P.C.B. and Ouliaris S. (1990) "Asymptotic Properties of Residual Based Tests for Cointegration" *Econometrica* 58, 165-193.

Snyder R.D. (1985) "Recursive Estimation of Dynamic Linear Statistical Models" *Journal of the Royal Statistical Society B*, 47, 272-276.

Watson M.W. (1986) "Univariate Detrending Methods with Stochastic Trends" *Journal of Monetary Economics* 18, 49-75.