Queen's Economics Department Working Paper No. 922

On the Power of Cointegration Tests: Dimension Invariance vs. Common Factors

Anindya Banerjee Juan J. Dolado Ricardo Mestre

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

3-1995

Discussion Paper #922

On the Power of Cointegration Tests: Dimension Invariance vs. Common Factors

by

Anindya Banerjee
University of Oxford
Queen's University

Juan J. Dolado
Bank of Spain

Ricardo Mestre
Bank of Spain

March 1995

Discussion Paper #922

On the Power of Cointegration Tests: Dimension Invariance vs. Common Factors

by

Anindya Banerjee
University of Oxford
Queen's University

Juan J. Dolado
Bank of Spain

Ricardo Mestre
Bank of Spain

March 1995

Discussion Paper #922

On the Power of Cointegration Tests: Dimension Invariance vs. Common Factors

by

Anindya Banerjee
University of Oxford
Queen's University

Juan J. Dolado
Bank of Spain

Ricardo Mestre
Bank of Spain

March 1995

ON THE POWER OF COINTEGRATION TESTS: DIMENSION INVARIANCE VS. COMMON FACTORS (*)

by

Anindya Banerjee
Wadham College
University of Oxford
and Queen's University

Juan J. Dolado
Research Dpt.
Bank of Spain
and CEMFI

Ricardo Mestre
Research Dpt.
Bank of Spain

January 1994
January 1995 (Revised Version)

ABSTRACT

This paper considers the trade-off, for cointegration tests, between dimension and power: that is, we compare the power performance of test-statistics which are dimension-invariant but impose common-factor restrictions with tests which are not dimension free but do not impose those restrictions. As a byproduct of the analysis, we consider cases where the t-ratio form of the tests have better power properties than the coefficient form, in spite of the latter diverging at rate $O_p(T)$ and the former at $O_p(T^{1/2})$, under the alternative hypothesis of cointegration.

Address for Correspondence: Juan J. Dolado, CEMFI, Casado del Alisal 5, 28014 Madrid. Tel.: (341)4290551.

(*) We are grateful to Luis Alvarez, Manuel Arellano, Peter Boswijk, Jesús Gonzalo, Niels Haldrup, David Hendry, Svend Hylleberg, Soren Johansen, James MacKinnon, Enrique Sentana, Gregor Smith, two anonymous referees and participants in seminars at the University of Oxford, University of Barcelona, CEMFI and ESEM'93 Uppsala for helpful discussions and to Araceli Requerey for her skill in typing the manuscript. The first author also wishes to thank the Dpt. of Economics at Queen's for their hospitality while a revised version of this paper was being written. The usual caveat applies.

1. INTRODUCTION

This paper compares the performance of a number of cointegration tests. The tests considered fall into two classes. The first consists of dimension-invariant tests, as recently proposed by Hansen (1990), based on the Cochrane-Orcutt estimation procedure (henceforth denoted as C-O test). The advantage of these tests is that they converge to the Dickey-Fuller distribution regardless of the number of variables in the cointegrating regression. The second class consists of tests based on the estimated error-correction coefficient in the ECM representation of the model, as proposed by Banerjee *et al.* (1986) and Boswijk (1991) (henceforth denoted as ECM test). In comparing both types of tests, we extend the argument made in Kremers *et al.* (1992), about imposing potentially invalid common-factor restrictions in residual based cointegration tests with known cointegrating vectors, to the case where cointegrating vectors have to be estimated, and argue that the dimension-invariant class of tests may suffer from the same problem. Consequently, if these restrictions are not satisfied the tests may have poor power properties.

Furthermore, as a byproduct of the analysis under the alternative hypothesis of cointegration, we show that cointegration tests based upon the t-ratio form may sometimes have better power properties than tests based upon the estimated coefficient itself. This is an interesting result, since as shown by Phillips and Ouliaris (1990), under the alternative hypothesis, the former test has non-centrality which grows at rate $T^{1/2}$ while the non-centrality parameter of the latter test diverges at rate T. However, we show that, for finite samples, as a local result, the t-form may be superior.

To examine the asymptotic and finite-sample properties of the various test procedures, we use a very simple, but illustrative, data generating process (DGP), and later show that the reason for the lack of power of the dimension-invariant tests may remain in more general cases.

The basic framework of reference is that of single equation conditional error correction models where OLS is an asymptotically efficient estimation procedure. For expository purposes, in most of the paper we use the simplifying assumption that the regressors are *strongly exogenous* for the parameters of interest. Later on in the analysis we generalise the results to the more realistic case where the regressors are only *weakly exogenous* but the use of OLS is still asymptotically optimal. We show that the ECM test in that case is not similar, i.e. depends on nuisance parameters. To overcome that problem, a "modified" ECM test is suggested which turns out to be asymptotically similar.

The rest of the paper is organised as follows. Section 2 presents the data generation process (DGP) of interest, briefly describes both the C-O and the ECM test procedures and compares their asymptotic distributions under the null hypothesis of non-cointegration. Section 3 gives the corresponding limiting distributions under the alternative hypothesis of cointegration, using both a fixed alternative and a near non-cointegrated alternative. Section 4 provides Monte-Carlo finite-sample evidence about the illustrative DGP. Section 5 considers generalisations to more realistic cases. Concluding remarks are given in section 6.

In common with most of the literature in this field, we follow some notational conventions. The symbol "$\Rightarrow$" denotes weak convergence of probability measures; "$\rightarrow$" denotes convergence in probability; "$\equiv$" denotes equality in distribution; $BM(\Omega)$ refers to a Brownian motion with long-run covariance matrix Ω. Arguments of functionals on the space [0,1] are frequently suppressed and integrals with respect to Lebesgue measure on the space [0,1] such as $\int_0^1 B_{\varepsilon}^2(r)dr$ are written as $\int B_{\varepsilon}^2$ to reduce notation. Emboldened symbols represent vectors or matrices of variables. Proofs of important results are relegated to the appendix.

2. A SIMPLE DGP AND THE C-O AND ECM TEST STATISTICS

By using a simple DGP, based upon a multivariate dynamic process, this section describes the C-O and ECM testing procedures.

This bivariate DGP has been used elsewhere [c.f. Davidson *et al.* (1978), Banerjee *et al.* (1986) and Kremers *et al.* (1992)] and has the form:

$$\Delta y_t = \alpha'\Delta\mathbf{x}_t + \beta(y_{t-1}-\lambda'\mathbf{x}_{t-1}) + \varepsilon_t \tag{1}$$

$$\Delta\mathbf{x}_t = \mathbf{u}_t \tag{2}$$

where $\Delta y_t = y_t - y_{t-1}$, $\Delta\mathbf{x}_t = \mathbf{x}_t - \mathbf{x}_{t-1}$, and

$$\begin{array}{c} 1 \\ k \end{array}\begin{pmatrix} \varepsilon_t \\ \mathbf{u}_t \end{pmatrix} \sim IN\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{\varepsilon}^2 & 0' \\ 0 & \Sigma_u \end{pmatrix}\right) \equiv N(0,\Sigma)$$

and α, λ and $\mathbf{x}_t$ are (k x 1) vectors of parameters and explanatory variables. With this set-up, the partial sum processes $S_T(r) = T^{-1/2}\sum_1^{[Tr]}(\varepsilon_t, \mathbf{u}_t')'$ satisfy the multivariate invariance principles [c.f. Phillips and Durlauf (1986)]:

$$S_T(r) \Rightarrow \Sigma^{1/2}\mathbf{B}(r) \equiv BM(\Omega)$$

where $\mathbf{B}(r)=(B_\varepsilon(r),\mathbf{B}_u(r)')'$ is a $(k+1)$ vector standard Brownian motion, i.e. BM(I).

We further assume that $-2<\beta\leq 0$. In this DGP, y_t and $\mathbf{x}_t$ are cointegrated when $-2<\beta<0$, while they are not cointegrated when $\beta=0$. Thus, tests of cointegration must rely upon some estimate of the parameter β. Under the simplifying assumption that $\mathbf{x}_t$ is strongly exogenous (c.f. Hendry and Richard (1982)) for the parameters of the conditional model (1), non-linear least squares (NLS) can be applied to (1) yielding consistent estimates of α, β and λ. Either the coefficient estimate $\hat{\beta}$ or the t-ratio based upon $\hat{\beta}$ can be used to test the null hypothesis that y_t and $\mathbf{x}_t$ are cointegrated with cointegrating vector $(1,-\lambda')$. However, it is well known that the asymptotic distributions of these test-statistics shift away from the origin as the dimensionality of the vector $\mathbf{x}_t$ increases (see, for example, Banerjee and Hendry (1992) and MacKinnon (1994)). Thus, larger test statistics are needed for rejection, reflecting the fact that this class of test-statistics is not dimension-invariant. This has often been called the "curse of dimensionality".

Hansen's C-O procedure corrects for this problem by using an iterative estimation variant of the Engle and Granger (1987) procedure. According to the C-O procedure [c.f. Cochrane and Orcutt (1949)], the iterated method is equivalent to estimating the parameters of the following model by NLS

$$\Delta y_t-\lambda'\Delta\mathbf{x}_t = \beta'[y_{t-1}-\lambda'\mathbf{x}_{t-1}] + e_t \tag{3}$$

where substitution of (2) into (1) implies that $e_t = (\alpha-\lambda)'\mathbf{u}_t+\varepsilon_t$.

Denoting the NLS estimators of λ and β in (3) by $\hat{\lambda}_{co}$ and $\hat{\beta}_{co}$, respectively, Hansen (1990, Theorem 2) proves the following result.

Proposition 1. For DGP (1)-(2), under the null hypothesis of no-cointegration (β=0)

$$T^{1/2}(\hat{\lambda}_{co}-\alpha) \Rightarrow N(\mathbf{0},\sigma^2_{\varepsilon}\Sigma^{-1}_{u})$$

$$T\hat{\beta}_{co} \Rightarrow \int B_{\varepsilon}dB_{\varepsilon}/\int B^2_{\varepsilon}$$

and

$$t_{co} \Rightarrow \int B_{\varepsilon}dB_{\varepsilon}/(\int B^2_{\varepsilon})^{1/2}$$

where t_{co} is the t-ratio defined as

$$t_{co} = (\hat{\sigma}^{-2}_{e}T^{-2}\hat{\mathbf{z}}'_{co-1}\hat{\mathbf{z}}_{co-1})^{1/2}T\hat{\beta}_{co};\ \hat{\sigma}^2_e = T^{-1}\sum_1^T\hat{e}^2_t$$

with $\hat{z}_{cot}=y_t-\hat{\lambda}'_{co}\mathbf{x}_t$ and $\hat{\mathbf{z}}_{co}$ is a (Tx1) vector of observations on $\hat{z}_{cot}$.■

The previous results show that both the C-O coefficient and t-ratio tests have the limiting Dickey-Fuller distributions. Note also that these tests are asymptotically similar, i.e invariant to the nuisance parameters of the DGP $(\alpha,\sigma^2_{\varepsilon},\Sigma_u)$ under the null hypothesis.

Although the previous test statistics are dimension-invariant, i.e. their distributions are independent of the dimension of the vector $\mathbf{x}_t$, it is important to note that the equation (3), as compared to equation (1), ignores part of the information contained in $\Delta\mathbf{x}_t$. Equivalently, (3) imposes the restriction that $\alpha=\lambda$, i.e. a common-factor restriction. The transformation of (1) to (3) provides several interesting insights. First, (1) and (3) are equivalent representations of the DGP, given the relationship between ε_t and e_t noted above, but the two errors are not equal unless $\alpha=\lambda$. Second, the same condition ($\alpha=\lambda$) is required for the common-factor restriction to be valid. This follows from noting that

$$[1-(1+\beta)L]y_t = [\alpha-(\alpha+\lambda\beta)L]'\mathbf{x}_t + \varepsilon_t \tag{4}$$

where L is the lag operator ($\Delta= 1-L$).

Interestingly, in this case, even if the common-factor restriction is invalid, e_t remains white noise, although not an innovation with respect to lagged $\mathbf{x}_t$ and y_t.

The ECM test statistic for cointegration, as suggested by Banerjee *et al.* (1986) and Boswijk (1991), can be based upon estimating (1) by NLS and testing H_0: $\beta=0$. Alternatively, Banerjee *et al.* (1993), using the results of Kiviet and Phillips (1992), show that a parameter free distribution for the estimator of β can be achieved if $\mathbf{x}_{t-1}$ is added to (1), setting $\lambda=0$ without loss of generality. This is so, since, under the alternative hypothesis, the true cointegrating slope λ is implicitly estimated when $\mathbf{x}_{t-1}$ is included as an additional regressor under the assumption of strong (weak) exogeneity of $\mathbf{x}_t$ with respect to β and λ. Therefore, according to this procedure, β is estimated by OLS from the unrestricted dynamic model

$$\Delta y_t = \alpha'\ \Delta\mathbf{x}_t + \beta y_{t-1} + \theta'\mathbf{x}_{t-1} + \varepsilon_t = \alpha'\Delta\mathbf{x}_t + \pi'\mathbf{w}_{t-1} + \varepsilon_t \tag{1'}$$

where $w_t'=(y_t, x_t')'$ and $\pi'=(\beta,\theta')$.

Since $\beta(1,-\lambda')=\pi'$, then the non-cointegration restriction $\beta=0$ implies $\pi=0$ and so the ECM test can be based on the significance of the OLS estimator of β, denoted by $\hat{\beta}_E$ in (1)'. Thus letting $\mathbf{y}$ and $\Delta\mathbf{y}$ to be (Tx1) vectors of observations on y_t and Δy_t, the ECM estimator, is defined by

$$\hat{\beta}_E = [\mathbf{y}'_{-1}\mathbf{M}\mathbf{y}_{-1}]^{-1}\mathbf{y}_{-1}\mathbf{M}\Delta\mathbf{y} \tag{5}$$

where $\mathbf{M} = \mathbf{I}-\mathbf{V}(\mathbf{V}'\mathbf{V})^{-1}\mathbf{V}'$ and $\mathbf{V}$ is a (Tx2k) matrix of observations on $\mathbf{v}_t=(\Delta\mathbf{x}_t', \mathbf{x}_{t-1}')'$. The following proposition holds.

Proposition 2. For DGP(1)-(2), under the null hypothesis of no cointegration ($\beta=0$)

$$T\hat{\beta}_E \Rightarrow [\int B_\varepsilon^2-(\int \mathbf{B}_u B_\varepsilon)'(\int \mathbf{B}_u \mathbf{B}_u')^{-1}(\int \mathbf{B}_u B_\varepsilon)]^{-1}[\int B_\varepsilon dB_\varepsilon-(\int \mathbf{B}_u B_\varepsilon)'(\int \mathbf{B}_u \mathbf{B}_u')^{-1}(\int \mathbf{B}_u dB_\varepsilon)]$$
$$\equiv [\int \bar{B}_\varepsilon^2]^{-1}\int \bar{B}_\varepsilon dB_\varepsilon$$

with $\bar{B}_\varepsilon = B_\varepsilon-(\int \mathbf{B}_u B_\varepsilon)'(\int \mathbf{B}_u \mathbf{B}_u')^{-1}\mathbf{B}_u$

and

$$t_E \Rightarrow [\int B_\varepsilon^2-(\int \mathbf{B}_u B_\varepsilon)'(\int \mathbf{B}_u \mathbf{B}_u')^{-1}\int \mathbf{B}_u B_\varepsilon]^{-1/2}[\int B_\varepsilon dB_\varepsilon-(\int \mathbf{B}_u B_\varepsilon)'(\int \mathbf{B}_u \mathbf{B}_u')^{-1}\int \mathbf{B}_u dB_\varepsilon]$$
$$\equiv [\int \bar{B}_\varepsilon^2]^{-1/2}\int \bar{B}_\varepsilon dB_\varepsilon$$

where t_E is the t-ratio defined as

$$t_E = [\hat{\sigma}_\varepsilon^{-2}T^{-2}\mathbf{y}'_{-1}\mathbf{M}\mathbf{y}_{-1}]^{1/2}T\hat{\beta}_E; \quad \hat{\sigma}_\varepsilon^2 = T^{-1}\sum_1^T \hat{\varepsilon}_t^2 \ \blacksquare$$

Note that $\bar{B}_\varepsilon$ is the residual from the continuous time regression of B_ε on $\mathbf{B}_u$. Thus the above distributions depend upon the number of elements (k) in $\mathbf{x}_t$, as reflected by the presence of $\mathbf{B}_u$ in $\bar{B}_\varepsilon$, implying that the corresponding test-statistics are not dimension-invariant.

Remark. As a curiosity, it can be noted from the prevous limiting distributions that there is a transformation of $T\hat{\beta}_E$ or t_E which has the Dickey-Fuller distribution and is also dimension-free. We will denote such estimator as EDF estimator (ECM cum Dickey-Fuller distribution). It is immediate to see that $\hat{\beta}_E$ in (5) is identical to

$$\hat{\beta}_E = (\hat{\mathbf{z}}'_{E-1}\mathbf{M}\hat{\mathbf{z}}_{E-1})^{-1}\hat{\mathbf{z}}'_{E-1}\mathbf{M}\Delta\hat{\mathbf{z}}_E \tag{6}$$

where $\hat{z}_{Et}=y_t-\hat{\alpha}'_E\mathbf{x}_t$ and $\hat{\mathbf{z}}_E$ is a (Tx1) vector of observations on $\hat{z}_{Et}$, where $\hat{\alpha}_E$ is the OLS estimator of α in (1'). Now, since the projection matrix $\mathbf{M}$ anihilates both $\mathbf{x}_{-1}$ and $\Delta\mathbf{x}$, tests based upon the following estimator

$$\hat{\beta}_{EDF} = (\hat{\mathbf{z}}'_{E-1}\hat{\mathbf{z}}_{E-1})^{-1}\hat{\mathbf{z}}'_{E-1}\Delta\hat{\mathbf{z}}_E \tag{7}$$

such as $T\hat{\beta}_{EDF}$ or the t-ratio, have the Dickey-Fuller asymptotic distributions. Thus, this is the only test which does not impose common-factor restrictions and is dimension invariant. However, in a sense, this is the free-lunch case, i.e it seems to buy dimension-invariance at zero cost, and should not be expected to exist to play a useful role. This is indeed what happens. The EDF test-statistics test for the stationarity of $\hat{z}_{Et}$ which is an inconsistent estimator of the true ECM term, $y_t-\lambda'\mathbf{x}_t$, because $\hat{\alpha}_E \to \alpha$ and not to λ, under the alternative hypothesis of cointegration.

Hence, the EDF test has zero power under fixed alternatives of cointegration and we do not discuss this test further.[1]

3. DISTRIBUTION OF THE STATISTICS UNDER THE ALTERNATIVE HYPOTHESIS OF COINTEGRATION

The alternative hypothesis is that of cointegration which, for (1)-(2), is given by $-2<\beta<0$. Because the error-correction term in (1) is stationary under the alternative hypothesis, distributional results from conventional central limit theorems, instead of functional central limit theorems, apply for fixed alternatives. In contrast, under a suitable sequence of local alternatives, the non-conventional asymptotic theory developed by Phillips (1987b, 1988) for near-integrated time series can be applied to sharpen the results on the relative asymptotic power functions for the C-O and ECM tests. In order to provide some intuition for the results obtained under near-no cointegration we first discuss briefly the fixed alternative case.

The main result here is that the ECM test tends to have higher power than the C-O test when $\mathrm{var}(e_t)$ is large relative to $\mathrm{var}(\varepsilon_t)$. The intuition behind these results is as follows. The ECM regression conditions on $\Delta\mathbf{x}_t$, $\mathbf{x}_{t-1}$ and y_{t-1}, whereas the C-O regression conditions on the three sets of variables subject to restrictions. This results in a loss of potentially valuable information. Consider again the alternative representations of equation (1):

$$\Delta y_t = \alpha'\Delta\mathbf{x}_t + \beta(y_{t-1}-\lambda'\mathbf{x}_{t-1}) + \varepsilon_t = \lambda'\Delta\mathbf{x}_t + \beta(y_{t-1}-\lambda'\mathbf{x}_{t-1}) + e_t$$

As an extreme example, let $\varepsilon_t \uparrow 0$ but $\alpha\neq\lambda$ and $(\alpha-\lambda)'\Sigma_u(\alpha-\lambda)$ is "substantial". In that case, the ECM regression has a near perfect fit with α, β and λ being estimated with near exact precision, and the t-ratio for $\hat{\beta}_E$ is (arbitrarily) large. However, since the variance of e_t is

$$\sigma_e^2 = (\alpha-\lambda)'\Sigma_u(\alpha-\lambda) + \sigma_\varepsilon^2$$

the estimates of λ and β in the C-O procedure will be much more imprecise, affecting the power of the test based upon $\hat{\beta}_{co}$.

3.1 Distributions under a Local Alternative Hypothesis

To formalise the intuition for the case of local alternatives, we use the distribution theory for local alternatives discussed in Phillips (1987b, 1988) *inter alia*. These non-central distributions help in the analysis of the local asymptotic power properties of the various tests and, as a limiting case, they allow us to obtain the distribution under fixed alternatives. A similar method was used by Kremers *et al.* (1992) for the case where the potential cointegrating vector was assumed to be known and did not need to be estimated. Consider the following parameterisation

$$\beta = 1-\exp(c/T) \simeq -c/T \tag{8}$$

In (8), c is a fixed scalar. We call time series that are generated by (1)-(2), with β as in (8), near-no cointegrated, following the terminology introduced by Phillips (1987b) for univariate processes. The scalar c represents a non-centrality parameter which may be used to measure deviations from the null hypothesis $H_0:\beta=0$. When $c>0$, (8) represents a local alternative to H_0, so that the rate of approach is controlled and the effect of the alternative hypothesis on the limiting distribution of the statistics, based on the DGP (1)-(2)-(8), is directly measurable in terms of the non-centrality parameter c.

To proceed to the analysis of local power, use is made of the following diffusion process

$$K(r) = \int_0^r \exp[c(r-s)]dB(s) \equiv B(r) + c\int_0^r \exp[c(r-s)]Bds \qquad (9)$$

associated with the standardised disturbances ε, $\mathbf{u}$ and e, denoted as K_ε, K_u and K_e, respectively. Note that if $c=0$ then $K=B$.

Using (9) it is possible to show the following result:

Proposition 3. For DGP (1)-(2) and (8), under the alternative hypothesis of near-no cointegration ($c>0$)

$$T\hat{\beta}_{co} \Rightarrow \left\{ \sigma_e^2 \int K_e^2 - 2\sigma_e(\alpha-\lambda)'\Sigma_u^{1/2}\int \mathbf{B}_u K_e + (\alpha-\lambda)'\Sigma_u^{1/2}\int \mathbf{B}_u \mathbf{B}_u' \Sigma_u^{1/2}(\alpha-\lambda) \right\}^{-1}$$

$$\left\{ -c[\sigma_e^2\int K_e^2 - \sigma_e(\alpha-\lambda)'\Sigma_u^{1/2}\int \mathbf{B}_u K_e] + \sigma_e\sigma_\varepsilon \int K_e dB_\varepsilon - \sigma_\varepsilon(\alpha-\lambda)'\Sigma_u^{1/2}\int \mathbf{B}_u dB_\varepsilon \right\} \equiv \Phi_{co}$$

$$T\hat{\beta}_E \Rightarrow -c + \left\{ \sigma_e^2[\int K_e^2 - (\int \mathbf{B}_u K_e)'(\int \mathbf{B}_u \mathbf{B}_u')^{-1}(\int \mathbf{B}_u' K_e) \right\}^{-1}$$

$$\sigma_e\sigma_\varepsilon\left\{ \int K_e dB_\varepsilon - (\int \mathbf{B}_u K_e)'(\int \mathbf{B}_u \mathbf{B}_u')^{-1}(\int \mathbf{B}_u dB_\varepsilon) \right\} \equiv \Phi_E$$

whereas in the case of the t-ratios,

$$t_{co} \Rightarrow \left\{\sigma_\varepsilon^{-2}[\sigma_e^2\int K_e^2-2\sigma_e(\alpha-\lambda)'\Sigma_u^{-1/2}\int B_uK_e+(\alpha-\lambda)'\Sigma_u^{1/2}\int B_uB_u'\Sigma_u^{1/2}(\alpha-\lambda)]\right\}^{1/2}\Phi_{co}$$

$$t_E \Rightarrow \left\{\sigma_\varepsilon^{-2}\sigma_e^2[\int K_e^2-(\int B_uK_e)'(\int B_uB_u')^{-1}(\int B_uK_e)]\right\}^{1/2}\Phi_E \quad \blacksquare$$

Since $\sigma_e K_e = (\alpha-\lambda)'\Sigma_u^{1/2}K_u+\sigma_\varepsilon K_\varepsilon$, note that, when c=0, the non-centrality parameters of the three statistics are zero, $K\equiv B$ and the distributions under the null in Propositions 1 & 2 are recovered, i.e. power equals size.

Although the comparison of the asymptotic distributions under the alternative local hypothesis is cumbersome, given the complexity of the Wiener functional derived above, we expect that the power of the C-O test will be higher than the power of the ECM test under those assumptions wherein the "curse of dimensionality" is strongest, i.e. when the common factor restrictions are valid and the number of regressors is large. However, when these restrictions do not hold, the relative ranking in power may be totally altered. To illustrate this case, let us simplify the analysis by assuming that there is a single regressor, that is k=1. Then, given the existing relationship between the disturbances e_t, ε_t and u_t:

$$e_t = (\alpha-\lambda)u_t + \varepsilon_t$$

we will define a "signal-to-noise" ratio $q=(\alpha-\lambda)s$ with $s=\sigma_u/\sigma_\varepsilon$, corresponding to the ratio of the (square root of the)variance of $(\alpha-\lambda)\Delta x_t$ relative to ε_t. This ratio will play a prominent role in the analysis since, as $q\uparrow\infty$, it allows for "small-σ" approximations, i.e $s^{-1}\uparrow 0$; cf. Kadane (1970,1971). Making use of these results, the following proposition holds.

Proposition 4. For DGP (1)-(2) and (8), when k=1, under the alternative hypothesis of near-no cointegration (c>0):

$$T\hat{\beta}_{co} \Rightarrow -c[\int(K_e-B_u)^2]^{-1}\int K_e(K_e-B_u) + o_p(q^{-1})$$
$$T\hat{\beta}_E \Rightarrow -c+o_p(q^{-1})$$

whereas in the case of the t-ratios,

$$t_{co} \Rightarrow -c[\int(K_e-B_u)^2]^{-1/2}\int K_e(K_e-B_u) + o_p(q^{-1})$$

$$t_E \Rightarrow -c(1+q^2)^{1/2}[\int K_e^2-(\int K_eB_u)^2/(\int B_u^2)]^{1/2} + [\int K_e^2-(\int K_eB_u)^2/(\int B_u^2)]^{-1/2}$$
$$[\int K_e dB_\varepsilon-(\int K_eB_u)(\int B_u dB_\varepsilon)/(\int B_u^2)] + o_p(q^{-1})$$

Various interesting properties arise from Proposition 4. In what follows it will be convenient to divide the discussion between those properties pertaining to the coefficient-test statistics and those relating to the t-ratios.

a) Coefficient Test-Statistics

First, asymptotically as $q\uparrow\infty$, i.e. $\alpha\neq\lambda$ and $s\uparrow\infty$, the non-centrality parameter in the C-O test has a stochastic slope given by zero whereas the ECM test has a slope equal to (minus) unity, i.e. there is a degenerate distribution centred on (-c). The intuition behind this result is that the variance of the denominator in the distribution of $T\hat{\beta}_{co}$ tends to overcome the variability of the numerator, leading to low power in the C-O test. Hence, the ECM test will tend to be more powerful than the

C-O test when q is sizeable, for a given sample size.

Second, and most importantly, from Proposition 2, since the limiting distribution of $T\hat{\beta}_E$ is independent of q under the null hypothesis, and degenerates around (-c) under the local alternative, for small values of c, the lower 5% tail of the distribution under the null will tend to be to the left of (-c). Therefore, we should also observe very low power of the test based on the ECM coefficient, although higher than that pertaining to the C-O test as q gets larger. Notice that this problem does not arise with the t-ratio version of the ECM test, as will be seen below. For this reason, the tests based upon the t-ratios might be preferable to those based directly on the scaled coefficients.

b) t-ratio Test-Statistics

In this case the limiting distribution of the ECM test has a stochastic slope which depends upon q and does not degenerate around a single value as in the case of the tests based on the coefficient. In the case of the C-O test-statistic, using arguments similar to those employed previously, Proposition 4 shows that the limiting distribution does not depend on q. Thus when q is sizeable, its power will be lower than that of the ECM test, tending towards zero even though the limiting distribution under the local alternative tends to be less skewed to the left than that under the null hypothesis.

3.2 Distributions under a Fixed Alternative Hypothesis

For the case of the fixed alternative ($c\uparrow\infty$ and $T\uparrow\infty$), the deviation from equilibrium ($y-\lambda'\mathbf{x}$) is stationary and the limiting distributions are as follows:

Proposition 5. For DGP (1)-(2), under a fixed alternative hypothesis ($\beta<0$), with $q^2=(\alpha-\lambda)'\Sigma_u(\alpha-\lambda)/\sigma_\varepsilon^2$

$$T\hat{\beta}_{co} \equiv T^{1/2}N[0,1-(1+\beta)^2] + T\beta = O_p(T)$$
$$t_{co} \equiv N(0,1) + [T/(1-(1+\beta)^2)]^{1/2}\beta = O_p(T^{1/2})$$
$$T\hat{\beta}_E \equiv T^{1/2}N[0,(1-(1+\beta)^2)/1+q^2] + T\beta = O_p(T)$$
$$t_E \equiv N(0,1) + [T(1+q^2)/(1-(1+\beta)^2)]^{1/2}\beta = O_p(T^{1/2}) \blacksquare$$

Remark. Note that, for given q, the coefficient tests are $O_p(T)$ whilst the t-ratio tests are $O_p(T^{1/2})$, as shown by Phillips and Ouliaris (1990, Theorem 5.1). However, notice that the non-centrality parameters of the t-ratio test will be larger than that of the coefficient test if $T<1/1-(1+\beta)^2$ (in the C-O case) and $T<(1+q^2)/[1-(1+\beta)^2]$ (in the ECM case). In terms of power itself, it is easy to see that, for both classes of tests, the power of the t-test will be larger than that of the coefficient test if $(1+\beta)^2<[1-(cv_c/cv_t)^2T^{-1}]$ where cv_c and cv_t are the critical values at the chosen significance level[2]. Equally, since q^2 is $\mathbf{R}^2/1-\mathbf{R}^2$ where $\mathbf{R}^2$ is the population R^2 with $\beta=0$, then the coefficient test will have larger non-centrality parameter than the t-ratio test if $\mathbf{R}^2>1-[T(1+\beta)^2]^{-1}$ (e.g. if $\beta=-0.01$ and $T=100$, then the cutt-off point will be such that $\mathbf{R}^2>0.497$). Finally, note that for $q>0$, the non-centrality parameter of t_E is larger than that of t_{co}, while the variance of the $T\hat{\beta}_{co}$ is larger than that of $T\hat{\beta}_E$.

5. FINITE SAMPLE EVIDENCE

To examine the size and power of the C-O and ECM statistics in finite samples, a set of Monte-Carlo experiments were conducted with (1) and (2) as the DGP, using simulations based on 25,000 replications generated in GAUSS386. A single exogenous regressor, k=1, was used for illustrative purposes. Data were generated with the normalization $\sigma_{\varepsilon}=1$, without loss of generality, leaving three parameters (s,α,β) and the sample size T as experimental design variables. In this study we choose

$s = (0.05, 1, 5, 20)$

$\alpha = (0.1, 0.9)$

β = (0 [no cointegration], -0.05, -0.10 [cointegration in both cases])

$T = (100)$

The implied range of the "signal-to-noise" ratio is broad, including values potentially favourable and unfavourable for the relative power comparisons among the different tests. In order to simplify the analysis, under the alternative hypothesis, the value of the cointegrating slope, λ, was fixed equal to 1. Similarly, the values of the short-run elasticity, α, attempt to capture a smaller (α=0.1) and a similar value (α=0.9) relative to the one chosen for λ. Combining the values of α and λ with those for s, we obtain a wide range of values for q, ranging from 0.005 to 18.

In order to compute the non-linear estimators in the C-O procedure, we have followed Hansen's advice in using a bias adjusted estimator of β in the initial iteration. Thus, if $\hat{\beta}^{1}_{co}$ is the estimator in the initial iteration, let us define

$$\hat{\beta}^{+}_{co} = \hat{\beta}^{1}_{co} + a/T$$

where a>0 is a fixed constant which Hansen suggests selecting equal to 10. Eight iterations were performed with this procedure and, at the final stage, a/T was subtracted from $\hat{\beta}^{+}_{H}$, in order to use the standard Dickey-Fuller tables.

Finally, in order to enlarge the range of the comparisons, we have also included the well known Engle and Granger (1987) test (henceforth denoted as EG test), based of the static regression model. This test suffers both from the "curse of dimensionality" and the "common-factor restriction" problem (c.f. Phillips and Ouliaris (1990) and Kremers *et al.* (1992)), so that it is useful to see how it performs relative to the other tests discussed in this paper.

Under the null of no cointegration (β=0) Table 1 presents a summary of the critical values at both tails of the distributions. In order to afford comparisons with the standard Dickey-Fuller distribution, to which the C-O test should correspond under the null, the first row in Table 1 offers the cumulative distribution of the DF test (c.f. Fuller (1976)). The empirical distributions were computed under the different choices of s, turning out to be highly invariant to the chosen value of that ratio, in agreement with analytical results contained in Proposition 1. Given this degree of invariance, the reported figures correspond to the averages of the critical values across the chosen range of values for s. On the one hand, it can be observed that the empirical distribution of the C-O test is close to the DF unit root distribution, although there seems to be more divergence in the case of the coefficient version than in the t-ratio version of the tests, where the deviations seem to be small. On the other hand, as expected, the empirical distribution of the ECM differs from the unit root distribution, reflecting the fact that it is not dimension-free. Similarly, the EG test, being a residual based test, also differs from the standard DF distribution.

Next, in order to examine the dependence of the test on the dimension of the system, we compare, in Table 2, the evolution of the critical values as the number of regressor is extended from one, as before, to five exogenous variables when T=100. The results here confirm our earlier finding. The t-tests seem to be more immune to the "curse of dimensionality" than the coefficient tests and, as expected, the ECM and EG tests shift to the left in distribution as the dimension of the system increases.

Finally, Table 3 reports size adjusted powers for the selected range of values for α and s, when $\beta=-0.05$ and $\beta=-0.10$. Since only negative values are consistent with the stability of the system, a one-sided 5% test was used. The results seem to be consistent with the asymptotic results derived in the previous section. First, when q is low and c is small, e.g. c=-5 when $\beta=-0.05$ and T=100, the ECM test, both in its t-ratio and coefficient versions, seems to be slightly less powerful than the C-O test, reflecting the "curse of dimensionality". However, as q increases, either because α becomes different from λ or because s rises, the ECM test becomes the most powerful. Also, in agreement with the degeneration of the asymptotic distributions of the coefficient version of the tests, their absolute power decreases as q increases. This is clearly not the case with the t-ratio versions where the ECM tests shifts its complete distribution to the left so as to achieve maximum power. For example, an extreme case is when c=-5 ($\beta=-0.05$), $\alpha=0.1$ and s=20, where the t-ratio version of the ECM test rejects 100% of the time, the C-O test almost does not reject at all.

As regards the power of the EG test, the results indicate that its power also decreases as q increases, though at a lower rate than the power of the C-O test. In agreement with the results in Banerjee *et al.* (1986), it turns out to have lower power than the ECM test, even when q is small.

5. GENERALISATIONS

The common factor problem of the C-O statistic remains when (1) includes additional lags. Furthermore, the use of the ADF test-statistic, or the non-parametric corrections suggested by Phillips (1987a), on the C-O residuals do not resolve the problem. Since this argument is similar to that given by Kremers et al. (1992) where the potential cointegration vector is assumed to be known *a priori*, we will not discuss it further. However, we note a few issues. First, if (1) is generalized to

$$\gamma(L)\Delta y_t = \alpha(L)'\Delta \mathbf{x}_t + \beta(y-\lambda'\mathbf{x})_{t-1} + \varepsilon_t \qquad (1'')$$

then e_t in the C-O procedure given by

$$e_t = (\alpha(L) - \gamma(L)\lambda')'\mathbf{u}_t + \varepsilon_t$$

need not be white noise. Indeed, in general, it will follow a moving average (MA) process, which are accounted by means of the Phillips and Perron's (1988) Z statistics. It is known than when the roots of such MA processes are close to be on the unit circle, these tests may suffer from severe size distortions; (c.f. Schwert (1989)).

Second, the simplifying assumption that $\mathbf{u}_t$ is white noise in

(1) can be replaced by being I(0), but long-run independent of ε_t, that is $\mathbf{B}_u$ will still be an independent Brownian motion of B_ε, with long-run variance (for k=1) given by:

$$\omega^2 = \sigma_u^2 + 2\phi_u$$

where $\sigma_u^2 = E(u_1^2)$ and $\phi_u = \sum_{j=2}^{\infty} E(u_1 u_j)$. In this case, the limiting distributions in Proposition remain similar, except that the "signal-to-noise" ratio q now becomes

$$q = (\alpha-\lambda)\omega_u/\sigma_\varepsilon$$

Third, there is the more realistic case where $\mathbf{x}_t$ is not strongly exogenous, as assumed in DGP(1)-(2), but only weakly exogenous for the parameters of interest $\psi=(\beta,\lambda')'$, in the Engle *et al.* (1983) sense. In this case, the more general DGP will consist of (1) and the following marginal process for $\Delta\mathbf{x}_t$

$$\Delta\mathbf{x}_t = \mathbf{u}_t = \sum_0^{\infty}\mathbf{A}_j\mathbf{a}_{t-j}, \quad \sum_0^{\infty} ||\mathbf{A}_j|| < \infty \qquad (2')$$

where the (k+1) innovation vector $\mathbf{a}_t=(\varepsilon_{t-1},\eta_t')'$ is assumed to be a strictly stationary and ergodic with zero mean and finite covariance matrix $\Sigma_a=\mathrm{diag}(\sigma_\varepsilon^2,\Sigma_\eta)>0$. Thus, in this more general case, the partial sum process constructed form the (k+1) vector $\mathbf{v}_t=(\varepsilon_t,\mathbf{u}_t')'$ will now converge to BM(Ω) where

$$\Omega = \begin{pmatrix} \omega_{\varepsilon\varepsilon} & \omega'_{u\varepsilon} \\ \omega_{u\varepsilon} & \Omega_{uu} \end{pmatrix} = \Sigma + \Lambda + \Lambda' = \Delta + \Lambda'; \ \Sigma = \begin{pmatrix} \sigma^2_\varepsilon & 0' \\ 0 & \Sigma_{uu} \end{pmatrix}$$

with $\omega_{\varepsilon\varepsilon}=\sigma^2_\varepsilon$, $\omega_{u\varepsilon}=\sum_{-\infty}^{\infty}E(\mathbf{u}_0\varepsilon_t)$ and $\Lambda=\sum_{1}^{\infty}E(\mathbf{v}_0\mathbf{v}'_t)$.

Let $I_t=\sigma(\mathbf{a}_t,\mathbf{a}_{t-1},\dots)$.Then, since $E(\varepsilon_t/I_t)=0$, OLS will yield efficient estimates of β in (1') and the ECM tests can be implemented in a single equation framework, yielding limiting distributions as those obtained in Proposition 2.

Notice that this follows because, under the previous assumptions, the correlation between ε_t and present and lagged values of Δx_t is zero, i.e $\Delta_{21}=\sum_{0}^{\infty}E(\mathbf{u}_0\varepsilon_t)=0$. Thus, second-order bias effects capturing the "one-sided long-run covariance" will be absent in the limiting distributions of the ECM test-statistics, as in Proposition 2. However, note that in this case, B_ε and $\mathbf{B}_u$ are no longer independent Brownian motions. To illustrate that feature take the following simple example. For k=1, let $\Delta x_t=u_t=\gamma\Delta y_{t-1}+\eta_t=\gamma(\alpha u_{t-1}+\varepsilon_{t-1})\eta_t$ with $E(\varepsilon_t\eta_s)=0$ for all t and s, then x_t will be weakly exogenous for ψ in (1'), but the long-run covariance between B_ε and $\mathbf{B}_u$ will be $\gamma(1-\alpha\gamma)^{-1}\sigma^2_\varepsilon$ under the null hypothesis of non-cointegration. This is so since $\sum_{1}^{\infty}E((\varepsilon_0\mathbf{u}_t)\neq0$, implying that the limiting distributions obtained in Proposition 2, will now depend on nuisance parameters $(\omega_{u\varepsilon})$ and the corresponding tests will not be asymptotically similar. Thus, in principle, the computation of critical values in this more general case is problematic. Consequently, the appropriate critical values will differ form those tabulated in Table 1 where B_ε and B_u were assumed to be BM(I).

To overcome the problem of lack of similarity, we follow in part the seminal work of Hansen and Phillips (1990) on the derivation of "modified" estimators which, under cointegration, achieve mixed Gaussian asymptotic distribution. Since our work is under the null of non-cointegration, the idea here is slightly different, namely to reformulate functionals of Brownian motions $\mathbf{B}(r)\equiv BM(\Omega)$ into distributionally equivalent functionals of independent standard Brownian motions BM(I), which will now be denoted as $\mathbf{W}(r)$. The mapping between the unstandardised $\mathbf{B}(r)$ and the standardised $\mathbf{W}(r)$ is given by $\mathbf{W}(r)=\mathbf{K}'\mathbf{B}(r)$ where $\Omega^{-1}=\mathbf{KK}'$; c.f. Phillips and Ouliaris (1990). Then

$$\mathbf{K}' = \begin{bmatrix} \omega_{\varepsilon\varepsilon.u}^{-1/2} & -\omega_{\varepsilon\varepsilon.u}^{-1/2}\omega_{\varepsilon u}\Omega_{uu}^{-1} \\ 0 & \Omega_{uu}^{-1/2} \end{bmatrix} = \begin{bmatrix} h & -h\mathbf{d}' \\ 0 & \Omega_{uu}^{-1/2} \end{bmatrix}$$

with $\omega_{\varepsilon\varepsilon.u}=\omega_{\varepsilon\varepsilon}-\omega_{\varepsilon u}\Omega_{uu}^{-1}\omega_{u\varepsilon}$, $\omega_{\varepsilon\varepsilon}=\sigma_\varepsilon^2$ and $\mathbf{d}'=\omega_{\varepsilon u}\Omega_{uu}^{-1}$. Thus, $W_1=h(B_\varepsilon-\mathbf{d}'\mathbf{B}_u)$ and $\mathbf{W}_2=\Omega_{uu}^{-1/2}\mathbf{B}_u$.

Let $\Delta y_t^+=\Delta y_t-\hat{\mathbf{d}}'\Delta\mathbf{x}_t$ where $\hat{\mathbf{d}}=\hat{\Omega}_{uu}^{-1}\hat{\omega}_{u\varepsilon}$. Note that, using residuals from the NLS estimation of (1') we can estimate Ω_{uu}, $\omega_{u\varepsilon}$ and h^{-1} consistently by

$$\hat{\Omega}_{uu} = T^{-1}\sum_{1}^{T} \Delta\mathbf{x}_t\Delta\mathbf{x}_t'$$

$$\hat{\omega}_{u\varepsilon} = \sum_{i=1}^{l} w_i\tilde{\gamma}_i; \quad \tilde{\gamma}_i = T^{-1}\sum_{i+1}^{l}\Delta\mathbf{x}_t\hat{\varepsilon}_{t-i}$$

$$\hat{h}^{-1} = \hat{\omega}^{1/2}_{\varepsilon\varepsilon.u} = \left(\hat{\sigma}^2_{\varepsilon} - \hat{\omega}_{\varepsilon u}\hat{\Omega}^{-1}_{uu}\hat{\omega}_{u\varepsilon}\right)^{1/2}$$

where the weights w_i are usually selected so that for each i, $l\uparrow\infty$ as $T\uparrow\infty$ so that $l=O(T^{1/4})$; c.f. Newey and West (1987). One simple choice for the weights is the Barlett window $w_i=1-|i|/(l+1)$ with $l=\text{int}[1+5T^{-.7/4}]$ as suggested by Bierens (1993). Then, the following result holds:

Proposition 6. For DGP (1)-(2'), under the null hypothesis ($\beta=0$), the ECM modified estimator and its t-ratio (denoted as ME) are defined as

$$\hat{\beta}_{ME} = [\mathbf{y}'_{-1}\mathbf{M}_1\mathbf{y}_{-1}]^{-1}\mathbf{y}'_{-1}\mathbf{M}_1\Delta\mathbf{y}^{+}$$

$$t_{ME} = [\hat{\omega}^{-1}_{\varepsilon\varepsilon.u}\mathbf{y}'_{-1}\mathbf{M}_1\mathbf{y}_{-1}]^{1/2}\hat{\beta}_{ME}$$

and have the following limiting distributions

$$T\hat{\beta}_{ME} \Rightarrow [\int\bar{W}^2_{\varepsilon}]^{-1}\int\bar{W}_{\varepsilon}dW_{\varepsilon}$$

$$t_{ME} \Rightarrow [\int\bar{W}^2_{\varepsilon}]^{-1/2}\int\bar{W}_{\varepsilon}dW_{\varepsilon}$$

with $\bar{W}_{\varepsilon}=W_{\varepsilon}-(\int\mathbf{W}_u W_{\varepsilon})'(\int\mathbf{W}_u\mathbf{W}'_u)^{-1}\mathbf{W}_u$.

Thus, form the previous result we can observe that the modified estimator $\hat{\beta}_{ME}$ can be interpreted as the OLS estimator of the coefficient on y_{t-1} in the regression of Δy^{+}_t on x_{t-1} and y_{t-1}. Note that $\Delta\mathbf{x}_t$ is excluded from the set of regressors, since in that case the use of the projection matrix $\mathbf{M}$ in the partitioned regression would imply $\mathbf{M}\Delta\mathbf{y}^{+}=\mathbf{M}\Delta\mathbf{y}$ given that $\mathbf{M}\Delta\mathbf{x}=0$.

Finally, though deterministic terms have been ignored in the previous analysis for the sake of simplicity, the data may be demeaned, or demeaned and detrended, before applying the various tests for cointegration. The limiting distributions of the various tests discussed in the paper in such cases are of the same form, except that the Brownian motions are replaced by the appropriate Brownian bridges. The asymptotic critical values for the ECM test in its t-ratio version, which are only available for k=1 (Banerjee *et al.* (1993)) are extended in Table 4. In order to analyse the finite sample distribution of those tests the critical values for the lower tail of the distribution in Table 4 up to five regressors for four different sample sizes (T=25,50,100 and 500) are also presented. Since there are many examples in applied work of single equation conditional models with weakly exogenous regressors for the parameters of interest (see, e.g. Hendry, 1987), we think that the above critical values may be widely applicable.

6. CONCLUSIONS

Testing for cointegration has become an important facet of empirical analysis of economic time series over the last several years and various tests are being used. In this paper, we compare the relative performance of dimension-invariant tests whose distributional theory does not depend upon the dimensionality of the system but impose common-factor restrictions and tests which are not dimension-free but do not impose such restrictions. Using the argument of Kremers *et al.* (1992), we show that in realistic cases the former may have poor power properties. Moreover, as a byproduct of the analysis, we show that, in spite of coefficient tests being $O_p(T)$ and t-ratio tests being $O_p(T^{1/2})$ under the alternative hypothesis of cointegration, the latter may have better power properties. The results are obtained for a simple DGP and then shown to extend to more general cases.

NOTES

1. However, if $\alpha=\lambda$ in the DGP, then the non-centrality parameters of the EDF and C-O tests are identical since now $\hat{\alpha}_E$ is a consistent estimate of λ.

2. The following example taken from Gregory and Hansen (1993) serves as an illustration of the previous result. In their Table 2, they generate simulations of the DGP, $y_t=1+2x_t+\varepsilon_t$ and $x_t=y_t+\eta_t$ with $\varepsilon_t=(1-\rho L)^{-1}v_t$ and $\Delta\eta_t=\omega_t$ where ω_t and v_t are orthogonal nid(0,1) processes. The critical values at 5% level of the Z_α (coefficient) and Z_t (t-ratio) tests (with constant term) are -40.48 and -4.61 respectively. Using the asymptotic distributions for $T\hat{\beta}_{co}$ and t_{co} in Proposition 5, for $\beta=-0.5$ and T=50, the rejection probabilities are $Pr(\phi<0.30)$ (t-ratio test) and $Pr(\phi<-2.53)$ (coefficient test) where ϕ is a standarised normal variate. Similarly for $\beta=-1$, they are $Pr[\phi<2.46]$ and $Pr[\phi<1.35]$ respectively. Thus, the results in Proposition 5 provide a simple explanation of their finding that the Z_t test has better power properties than the Z_α test in those instances.

REFERENCES

Anderson, T.W. (1991) *The Statistical Analysis of Time Series*, New York: Wiley.

Banerjee, A., J. Dolado, J. Galbraith and D.F. Hendry (1993) *Co-integration, Error-Correction and the Econometric Analysis of Non-Stationary Data*, Oxford, Oxford University Press.

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.

Banerjee, A., and D.F. Hendry (1992) "Testing Integration and Cointegration: An Overview", *Oxford Bulletin of Economics and Statistics*, 5448, 225-255.

Bierens, H.J. (1993) "High Order Second Autocorrelations and the Unit Root Hypothesis", *Journal of Econometrics* 57, 137-160.

Boswijk, H.P. (1991) "Testing for Cointegration in Structural Models", Report AE 7/91, University of Amsterdam.

Cochrane, D. and G.H. Orcutt (1949) "Application of Least Squares Regression to Relationships Containing Autocorrelated Error Terms", *Journal of the American Statistical Association*, 44, 32-61.

Davidson, J.E.H., D.F. Hendry, F. Srba and S. Yeo (1978) "Econometric Modelling of the Aggregate Time-Series Relationships between Consumers' Expenditure and Income in the United Kingdom", *Economic Journal*, 88, 661-692.

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

Fuller, W.A. (1976) *Introduction to Statistical Time Series*, New York, John Wiley and Sons.

Gregory, A.W. and B.E. Hansen (1993) "Residual-Based Tests for Cointegration in Models with Regime Shifts", mimeo.

Hansen, B.E. (1990) "A Powerful, Simple Test for Cointegration Using Cochrane-Orcutt" University of Rochester. Working Paper No. 230.

Hansen, B.E. and P.C.B. Phillips (1990) "Estimation adn Inference in Models of Cointegration: A Simulation Study" in Fomby, T. and Rhodes, G. (Eds.), *Advances in Econometrics* 8: *Co-Integration, Spurious Regressions and Unit Roots* (Greenwich, CT: JAI Press), pp. 225-248.

Hendry, D.F. (1987) "Econometric Methodology: A Personal Perspective" in T. Bewley (ed.) *Advances in Econometrics*, vol. 2, Cambridge: CUP.

Hendry, D.F. and J.F. Richard (1982) "On the Formulation of Empirical Models in Dynamic Econometrics", *Journal of Econometrics*, 20, 3-23.

Kadane, J.B. (1970) "Testing Overidentifying Restrictions When the Disturbances Are Small", *Journal of the American Statistical Association* 65, 182-185.

Kadane, J.B. (1971) "Comparison of k-Class Estimators When the Disturbances Are Small", *Econometrica* 39, 723-37.

Kiviet, J.F. and G.D.A. Phillips (1992) "Exact Similar Tests for Unit Roots and Cointegration", *Oxford Bulletin of Economics and Statistics* 54, 349-368.

Kremers, J.J.M. Ericsson, N.R. and J.Dolado (1992) "The Power of Cointegration Tests", *Oxford Bulletin of Economics and Statistics*, 54, 325-348.

Mac Kinnon, J.G. (1994) "Approximate Asymptotic Distribution Functions for Unit-Root and Cointegration Tests", *Journal of Business & Economic Statistics* 12, 167-176.

Neuvey, W. and K. West (1987) "A Simple Positive Semi-Definite Heteroskedasticity and Autocorrelation Consisntent Covariance Matrix", *Econometrica* 55, 703-708.

Phillips, P.C.B. (1986) "Understanding Spurious Regressions in Econometrics", *Journal of Econometrics* 33, 311-340.

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

Phillips, P.C.B. (1987b) "Towards a Unified Asymptotic Theory of Autoregression", *Biometrika*, 74, 535-547.

Phillips, P.C.B. (1988) "Regression Theory for Near-Integrated Time Series", *Econometrica*, 56, 1021-1043.

Phillips, P.C.B. and S.N. Durlauf (1986) "Multiple Times Series with Integrated Variables", *Review of Economic Studies* 53, 473-496.

Phillips, P.C.B. and P. Perron (1988) "Testing for a Unit Root in Time Series Regression", *Biometrika* 75, 335-346.

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

Schwert, G.W. (1989) "Tests for Unit Roots: A Monte Carlo Investigation", *Journal of Business and Economic Statistics* 7, 147-159.

APPENDIX

The analysis contained in this appendix draws on a number of well known results in Phillips (1987b, 1988) and Phillips and Ouliaris (1990). Under the null hypothesis of no-cointegration, the DGP (H_0) is given by

$$\Delta y_t = \alpha' \Delta \mathbf{x}_t + \varepsilon_t; \qquad \begin{pmatrix} \varepsilon_t \\ \mathbf{u}_t \end{pmatrix} \sim IN\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix} \begin{pmatrix} \sigma_\varepsilon^2 & 0' \\ 0 & \Sigma_u \end{pmatrix} \right)$$

$$\Delta \mathbf{x}_t = \mathbf{u}_t;$$

The results do not depend on the inialization, so let us define $S_{\varepsilon t} = \sum_1^T \varepsilon_i$ and $\mathbf{S}_{ut} = \sum_1^T \mathbf{u}_i$ and let $\mathbf{S}_\varepsilon$ and $\mathbf{S}_u$ be (Tx1) and (Txk) matrices of observations on $S_{\varepsilon t}$ and $\mathbf{S}_{ut}$, respectively. Note that $\mathbf{x}=\mathbf{S}_u$ and $y=\mathbf{S}_u\alpha+S_\varepsilon$. Let $\mathbf{M}_1$ be the projection matrix $\mathbf{M}_1 = I - \mathbf{x}_{-1}(\mathbf{x}'_{-1}\mathbf{x}_{-1})^{-1}\mathbf{x}'_{-1}$.

Then, the following set of asymptotic results (R1) are used in the proofs:

(a) $T^{-2}\mathbf{S}'_{\varepsilon-1}\mathbf{M}_1\mathbf{S}_{\varepsilon-1} \Rightarrow \sigma_\varepsilon^2 \int \bar{B}_\varepsilon^2$

(b) $T^{-1}\mathbf{S}'_{\varepsilon-1}\mathbf{M}_1\varepsilon \Rightarrow \sigma_\varepsilon^2 \int \bar{B}_\varepsilon dB_\varepsilon$

(c) $T^{-1}\mathbf{u}'\mathbf{M}_1\mathbf{u} = O_p(1)$

(d) $T^{-1}\mathbf{S}'_{\varepsilon-1}\mathbf{M}_1\mathbf{u} = O_p(1)$

(e) $T^{-1}\sum_1^T \varepsilon_t \mathbf{u}_t = o_p(1)$

where $\bar{B}_\varepsilon = B_\varepsilon - (\int \mathbf{B}_u B_\varepsilon)'(\int \mathbf{B}_u \mathbf{B}'_u)^{-1}\mathbf{B}_u$.

Under the local alternative hypothesis of near-no cointegration the DGP (H_{1a}) is given by

$$\Delta z_t = \beta z_{t-1} + e_t$$

$$\Delta \mathbf{x}_t = \mathbf{u}_t$$

with $\beta=-c/T$; $z_t=y_t-\lambda'\mathbf{x}_t$, $e_t=(\alpha-\lambda)'\mathbf{u}_t+\varepsilon_t$ and $\mathbf{z}$ and $\mathbf{e}$ are (Tx1) vectors of observations on z_t and e_t.

In this case the following additional asymptotic results (R2) are used

(a) $T^{-2}\mathbf{z}'\mathbf{z} \Rightarrow \sigma_e^2\int K_e^2$

(b) $T^{-2}\mathbf{z}'\mathbf{M}_1\mathbf{z} \Rightarrow \sigma_e^2\int \bar{K}_e^2$

(c) $T^{-2}\mathbf{z}'_{-1}\mathbf{M}_1\varepsilon \Rightarrow \sigma_e\sigma_\varepsilon\int \bar{K}_e dB_\varepsilon$

(d) $T^{-1}\mathbf{z}'_{-1}\mathbf{e} \Rightarrow \sigma_e^2\int K_e dB_e$

(e) $T^{-2}\mathbf{x}'\mathbf{z} \Rightarrow \sigma_e\Sigma_u^{1/2}\int \mathbf{B}_u K_e$

where $\bar{K}_e=K_e-(\int \mathbf{B}_u K_e)'(\int \mathbf{B}_u \mathbf{B}_u')^{-1}\mathbf{B}_u$

and

$$\sigma_e^2\int K_e^2 = (\alpha-\lambda)'\Sigma_u^{1/2}(\int \mathbf{K}_u \mathbf{K}_u')\Sigma_u^{1/2}(\alpha-\lambda)+2(\alpha-\lambda)'\sigma_\varepsilon\Sigma_u^{1/2}\int \mathbf{K}_u K_\varepsilon+\sigma_\varepsilon^2\int K_\varepsilon^2$$

Proof of Proposition 1

See Hansen (1990, Theorem 2)

Proof of Proposition 2

Let $\mathbf{V}$ be a (Tx2k) matrix of observations on $\mathbf{v}_t=(\Delta\mathbf{x}'_t,\mathbf{x}'_{t-1})'$ and $\mathbf{x}_{-1}$ and $\Delta\mathbf{x}$ be (Txk) matrices of observations on $\Delta\mathbf{x}_t$ and $\mathbf{x}_{t-1}$, respectively. Define the projection matrices $\mathbf{M=I-V(V'V)V'}$ such that, by partitioned inverses, $\mathbf{M}=\mathbf{M}_1-\mathbf{M}_1\Delta\mathbf{x}(\Delta\mathbf{x}'\mathbf{M}_1\Delta\mathbf{x})^{-1}\Delta\mathbf{x}'\mathbf{M}_1$.

Then, $\hat{\beta}_E$ is computed such that

$$T\hat{\beta}_E = (T^{-2}\mathbf{y}'_{-1}\mathbf{M}\mathbf{y}_{-1})^{-1}(T^{-1}\mathbf{y}'_{-1}\mathbf{M}\Delta\mathbf{y})$$

$$= (T^{-2}\mathbf{S}'_{\varepsilon-1}\mathbf{M}\mathbf{S}_{\varepsilon-1})^{-1}(T^{-1}\mathbf{S}'_{\varepsilon-1}\mathbf{M}\varepsilon)$$

since $\mathbf{y}=\mathbf{S}_u\alpha+\mathbf{S}_\varepsilon$, $\Delta\mathbf{y}=\mathbf{u}\alpha+\varepsilon$ and $\mathbf{M}$ is orthogonal to $\mathbf{x}_{-1}$ and $\Delta\mathbf{x}$. Using parts (a) to (d) of (R1) and the relationship between $\mathbf{M}$ and $\mathbf{M}_1$ we have

$$T^{-2}\mathbf{S}'_{\varepsilon-1}\mathbf{M}\mathbf{S}_{\varepsilon-1}=(T^{-2}\mathbf{S}'_{\varepsilon-1}\mathbf{M}_1\mathbf{S}_{\varepsilon-1})-T^{-1}(T^{-1}\mathbf{S}'_{\varepsilon-1}\mathbf{M}_1\mathbf{u})(T^{-1}\mathbf{u}'\mathbf{M}_1\mathbf{u})^{-1}(T^{-1}\mathbf{u}'\mathbf{M}_1\mathbf{S}_{\varepsilon-1})$$

$$=T^{-2}\mathbf{S}'_{\varepsilon-1}\mathbf{M}_1\mathbf{S}_{\varepsilon-1}+T^{-1}O_p(1)O_p(1)O_p(1)=T^{-2}\mathbf{S}'_{\varepsilon-1}\mathbf{M}_1\mathbf{S}_{\varepsilon-1}+o_p(1)$$

and

$$T^{-1}\mathbf{S}'_{\varepsilon-1}\mathbf{M}\varepsilon=(T^{-1}\mathbf{S}'_{\varepsilon-1}\mathbf{M}_1\varepsilon)-(T^{-1}\mathbf{S}'_{\varepsilon-1}\mathbf{M}_1\mathbf{u})(T^{-1}\mathbf{u}'\mathbf{M}_1\mathbf{u})^{-1}(T^{-1}\mathbf{u}'\mathbf{M}_1\varepsilon)=$$

$$T^{-1}\mathbf{S}'_{\varepsilon-1}\mathbf{M}_1\varepsilon-O_p(1)O_p(1)o_p(1)=T^{-1}\mathbf{S}'_{\varepsilon-1}\mathbf{M}_1\varepsilon+o_p(1)$$

since $T^{-1}\mathbf{u}'\mathbf{M}_1\varepsilon=(T^{-1}\mathbf{u}'\varepsilon)-T^{-1}(T^{-1}\mathbf{u}'\mathbf{x}_{-1})(T^{-2}\mathbf{x}'_{-1}\mathbf{x}_{-1})^{-1}(T^{-1}\mathbf{x}'_{-1}\varepsilon)=o_p(1)$, given (e) in R(1).

Next, using the limiting distributions in (a) and (b) in R(1), yields the required results

$$T\hat{\beta}_E=(S'_{\varepsilon-1}M_1S_{\varepsilon-1})^{-1}(S'_{\varepsilon-1}M_1\varepsilon)+o_p(1) \Rightarrow (\int\bar{B}^2_\varepsilon)^{-1}\int\bar{B}_\varepsilon dB_\varepsilon \tag{A.1}$$

To prove that $\hat{\sigma}^2_\varepsilon \to \sigma_\varepsilon$, just write

$$\hat{\sigma}^2_\varepsilon=T^{-1}\varepsilon' M_p\varepsilon=T^{-1}\varepsilon'\varepsilon-T^{-1}(T^{-1}\varepsilon'\mathbf{P})(T^{-2}\mathbf{P}'\mathbf{P})^{-1}(T^{-1}\mathbf{P}'\varepsilon)$$

$$=T^{-1}\varepsilon'\varepsilon-T^{-1}O_p(1)O_p(1)O_p(1)=\sigma^2_\varepsilon+o_p(1) \tag{A.2}$$

where $\mathbf{P}$ is the (Tx2k+1) matrix of observations on $(\Delta\mathbf{x}'_t,\mathbf{x}'_{t-1},y_{t-1})$ and $\mathbf{M}_p=I-\mathbf{P}(\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}'$.

From (A.1) and (A.2) the distribution of the t-ratio follows along the same lines, leading to the required results.

Proof of Proposition 3

Define $z_t=y_t-\lambda'\mathbf{x}_t$, $\hat{z}_{cot}=y_t-\hat{\lambda}'_{co}\mathbf{x}_t$ and $\hat{z}_{Et}=y_t-\hat{\alpha}'_E\mathbf{x}_t$.

Then

$$z_t=\hat{z}_{cot}+(\hat{\lambda}_{co}-\lambda)'\mathbf{x}_t=\hat{z}_{cot}+(\alpha-\lambda)'\mathbf{x}_t+o_p(1) \tag{A.3}$$

and

$$z_t=\hat{z}_{Et}+(\hat{\alpha}_E-\lambda)'\mathbf{x}_t = \hat{z}_{Et}+(\alpha-\lambda)'\mathbf{x}_t+o_p(1) \tag{A.4}$$

since $\hat{\lambda}_{co}\to\infty$ and $\hat{\alpha}_e\to\alpha$ at rate $O_p(T^{-1/2})$. Let $\tilde{z}_t=z_t-(\alpha-\lambda)'\mathbf{x}_t$ and $\tilde{\mathbf{z}}$ is the (Tx1) vector of observations on $\tilde{z}_t$.

Then, from (A.3) and (A.4), the C-O and ECM estimators can be written as

$$T\hat{\beta}_{co}=(T^{-2}\hat{\mathbf{z}}'_{co-1}\hat{\mathbf{z}}_{co-1})^{-1}(T^{-1}\hat{\mathbf{z}}'_{co-1}\Delta\hat{\mathbf{z}}_{co-1})=$$

$$T\beta+(T^{-2}\tilde{\mathbf{z}}'_{-1}\tilde{\mathbf{z}}_{-1})^{-1}T^{-1}\tilde{\mathbf{z}}'_{-1}(\varepsilon+\beta\mathbf{x}_{-1}(\alpha-\lambda))+o_p(1) \tag{A.5}$$

and

$$T\hat{\beta}_E=(T^{-2}\mathbf{y}'_{-1}\mathbf{M}\mathbf{y}_{-1})(T^{-1}\mathbf{y}'_{-1}\mathbf{M}\Delta\mathbf{y})=(T^{-2}\mathbf{z}'_{-1}\mathbf{M}\mathbf{z}_{-1})^{-1}(T^{-1}\mathbf{z}'_{-1}\mathbf{M}\Delta\mathbf{y})$$

$$=T\beta+(T^{-2}\mathbf{z}'_{-1}\mathbf{M}_1\mathbf{z}_{-1})^{-1}(T^{-1}\mathbf{z}'_{-1}\mathbf{M}_1\varepsilon)+o_p(1) \tag{A.6}$$

since $\mathbf{M}$ is orthogonal to $\mathbf{x}_{-1}$ and $\Delta\mathbf{x}$ and the limiting distribution of $(T^{-2}\mathbf{z}'_{-1}\mathbf{M}\mathbf{z}_{-1})$ is equal to the limiting distribution of $(T^{-2}\mathbf{z}'_{-1}\mathbf{M}_1\mathbf{z}_{-1})$, as shown in the proof of Propositon 2.

Then using $T\beta$=-c and substituting results (a) to (e) in (R2) into (A.5) and (A.6) yields the required results. Since $\hat{\sigma}_\varepsilon\to\sigma_\varepsilon$ and $\hat{\sigma}_e\to\sigma_e$, the proofs for the t-ratio statistics follow along similar lines, leading to the required results.■

Proof of Proposition 4

For k=1, from the limiting distributions in Proposition 3, we have

$$T\hat{\beta}_{co}\Rightarrow(\int\tilde{B}_e^2)^{-1}[-c\int K_e\tilde{B}_e+(\sigma_\varepsilon/\sigma_e)\int\tilde{B}_e dB_\varepsilon]$$

where $\tilde{B}_e=[K_e-(\alpha-\lambda)'(\sigma_u/\sigma_e)B_u]$.

Since $q=(\alpha-\lambda)\sigma_u/\sigma_\varepsilon$, it follows that $(\sigma_\varepsilon/\sigma_e)=(1+q^2)^{-1/2}$ and $(\alpha-\lambda)\sigma_u/\sigma_e=q/(1+q^2)^{1/2}$. Thus, as $q\uparrow\infty$, $(\sigma_\varepsilon/\sigma_e)\uparrow 0$ and

$(\alpha-\lambda)(\sigma_u/\sigma_e)\uparrow 1$, i.e $\tilde{B}_e \equiv K_e$, implying the required result for $T\hat{\beta}_{co}$, namely

$$T\hat{\beta}_{co} \Rightarrow -c\int(K_e^2-K_eB_u)/\int(K_e-B_u)^2+o(q^{-1})$$

As regards $T\hat{\beta}_E$, we have

$$T\hat{\beta}_E \Rightarrow -c+(\sigma_\varepsilon/\sigma_e)(\int\bar{B}_e^2)^{-1}(\int\bar{B}_e dB_\varepsilon)$$

where $\bar{B}_e = K_e-(\int B_uK_e)(\int B_u^2)^{-1}B_u$. Since $(\sigma_\varepsilon/\sigma_e)=(1+q^2)^{-1/2}$, as $q\uparrow\infty$ $(\sigma_\varepsilon/\sigma_e)\uparrow 0$ and

$$T\hat{\beta}_E \Rightarrow -c+o(q^{-1})$$

as required. Since $\hat{\sigma}_\varepsilon \to \sigma_\varepsilon$ and $\hat{\sigma}_e \to \sigma_e$ the proofs for the limiting distribution t-ratios proceed along similar lines.■

Proof of Proposition 5

Under a fixed alternative, the DGP(H_a) is given by

$$\Delta y_t = \alpha'\Delta\mathbf{x}_t + \beta(y_{t-1}-\lambda'\mathbf{x}_{t-1}) + \varepsilon_t$$
$$\Delta\mathbf{x}_t = \mathbf{u}_t$$

with $\beta<0$. Then, $z_t(=y_t-\lambda'\mathbf{x}_t)$ is governed by the following AR(1) process

$$z_t = [1-(1+\beta)L]^{-1}[(\alpha-\lambda)'\mathbf{u}_t+\varepsilon_t] = [1-(1+\beta)L]^{-1}e_t$$

where L is the lag operator and the variance of z_t is $V(z_t)=\sigma_e^2[1-(1+\beta)^2]^{-1}$.

Then,

$$T^{1/2}(\hat{\beta}_{co}-\beta) = (T^{-1}\mathbf{z}'_{-1}\mathbf{z}_{-1})^{-1}T^{-1/2}\mathbf{z}'_{-1}\Delta\mathbf{z} + o_p(1) \Rightarrow N[0,1-(1+\beta)^2]$$

by standard theory (e.g. Anderson, 1971, Chapter 5).

Thus,

$$T\hat{\beta}_{co} = T(\hat{\beta}_{co}-\beta) + T\beta = T^{1/2}N[0,1-(1+\beta)^2] + T\beta = O_p(T)$$

and

$$t_{co} = (T^{-1}\hat{\sigma}_e^{-2}\mathbf{z}'_{-1}\mathbf{z}_{-1})^{1/2}[T^{1/2}(\hat{\beta}_{co}-\beta)+T^{1/2}\beta] \Rightarrow N(0,1) + [T/1-(1+\beta)^2]^{1/2}\beta$$
$$= O_p(T^{1/2})$$

Similarly,

$$T^{1/2}(\hat{\beta}_E-\beta) \Rightarrow N[0,\sigma_\varepsilon^2[1-(1+\beta)^2]/\sigma_e^2] \equiv N[0,(1-(1+\beta)^2)/1+q^2]$$

where $q^2=(\alpha-\lambda)'\Sigma_u(\alpha-\lambda)/\sigma_\varepsilon^2$.

Thus,

$$T\hat{\beta}_E = T^{1/2}N[0,(1-(1+\beta)^2)/1+q^2] + T\beta = O_p(T)$$

and

$$t_E = (T^{-1}\hat{\sigma}_\varepsilon^{-2}\mathbf{z}'_{-1}\mathbf{M}\mathbf{z}_{-1})^{-1/2}[T^{1/2}(\hat{\beta}_E-\beta) + T^{1/2}\beta] \Rightarrow N[0,1] +$$
$$[T(1+q^2)/(1-(1+\beta)^2)]^{1/2} = O_p(T^{1/2})$$

as required.∎

Proof of Proposition 6

Let

$$\bar{W}_1 = W_1-(\int \mathbf{W}_2 W_1)'(\int \mathbf{W}_2 \mathbf{W}_2')^{-1}\mathbf{W}_2$$

$$\bar{B}_\varepsilon = B_\varepsilon-(\int \mathbf{B}_u B_\varepsilon)'(\int \mathbf{B}_u \mathbf{B}_u')^{-1}\mathbf{B}_u$$

Then, since $W_1=h(B_\varepsilon-\mathbf{d}'\mathbf{B}_u)$ and $\mathbf{W}_2=\Omega_{uu}^{-1/2}\mathbf{B}_u$, it follows that

$$\bar{W}_1 = h\bar{B}_\varepsilon$$

Therefore,

$$\int \bar{W}_1^2 = h^2\int \bar{B}_\varepsilon^2$$

and

$$\int \bar{W}_1 dW_1 = h^2\int \bar{B}_\varepsilon dB_\varepsilon - h^2\mathbf{d}'\int B_\varepsilon d\mathbf{B}_u$$

Hence,

$$[\int \bar{W}_1^2]^{-1}\int \bar{W}_1 dW_1 = [\int \bar{B}_\varepsilon^2]^{-1}[\int \bar{B}_\varepsilon dB_\varepsilon - \mathbf{d}'\int \bar{B}_\varepsilon d\mathbf{B}_u] \qquad (A.7)$$

Therefore, the "modified" ECM estimator defined by

$$T\hat{\beta}_{EM} = [T^{-2}\mathbf{y}_{-1}'\mathbf{M}_1\mathbf{y}_{-1}][T^{-1}\mathbf{y}_{-1}'\mathbf{M}_1(\Delta y-\Delta\mathbf{x}\hat{\mathbf{d}}]$$

has the limiting distribution given in (A.7) since

(i) $\hat{\mathbf{d}} \to \mathbf{d}$

(ii) $T^{-2}\mathbf{y}'_{-1}\mathbf{M}_1\mathbf{y}_{-1} \Rightarrow \sigma^2_\varepsilon \int \bar{B}^2_\varepsilon$ and $T^{-1}\mathbf{y}'_{-1}\mathbf{M}_1\Delta\mathbf{y} \Rightarrow \sigma^2_\varepsilon \int \bar{B}_\varepsilon dB_\varepsilon$

(as shown in Proposition 1)

(iii) $T^{-1}\mathbf{y}'_{-1}\mathbf{M}_1\Delta\mathbf{x}=T^{-1}\mathbf{y}'_{-1}\mathbf{M}_1\mathbf{u} =T^{-1}\mathbf{S}'_{\varepsilon-1}\mathbf{M}_1\mathbf{u}\Rightarrow\int\bar{B}_\varepsilon d\mathbf{B}_u$ (by (d) in R1)

leading to the required results. The proof for the "modified" ratio t_{ME} follows along similar lines.■

Table 1

Critical Values for the Coefficient (t-ratio) Version of the Tests

Size (k=1)

Test	0.05	0.10	0.90	0.95
DF	-7.90 (-1.95)	-5.60 (-1.61)	0.95 (0.90)	1.31 (1.29)
C-O	-8.69 (-2.05)	-6.16 (-1.69)	0.97 (0.92)	1.37 (1.34)
ECM	-12.38 (-2.60)	-9.66 (-2.27)	0.66 (0.44)	1.21 (0.88)
EG	-15.24 (-2.77)	-12.40 (-2.47)	-0.54 (-0.27)	0.24 (0.15)

Note: The number of replications (N) is 25,000; the sample size (T) is 100. The first figure corresponds to the critical value of the coefficient version of the test; the second figure (in parenthesis) corresponds to the critical value of the t-ratio version of the test.
The notation associated with the tests is the following: i) DF: Dickey-Fuller standard unit root test; C-O: Hansen's Cochrane-Orcutt tests (computed after eight iterations with a correction factor equal to 10/T); ECM: ECM coefficient test; EG: Engle and Granger test (computed form the OLS residuals of the static regression of y_t on x_t).

Table 2

Critical Values for the Coefficient (t-ratio) Version of the Tests
(different number of regressors)

Size

Test	0.05	0.10	0.90	0.95
C-O (k=1)	-8.69 (-2.05)	-6.16 (-1.69)	0.97 (0.92)	1.37 (1.34)
(k=2)	-9.76 (-2.15)	-6.94 (-1.77)	0.98 (0.86)	1.34 (1.28)
(k=3)	-9.50 (-2.19)	-6.17 (-1.81)	0.99 (0.90)	1.36 (1.31)
(k=4)	-9.73 (-2.28)	-6.47 (-1.91)	0.98 (0.88)	1.37 (1.32)
(k=5)	-10.37 (-2.38)	-7.22 (-1.98)	0.94 (0.88)	1.36 (1.34)
ECM (k=1)	-12.38 (-2.60)	-9.66 (-2.27)	0.66 (0.44)	1.21 (0.88)
(k=2)	-16.38 (-3.03)	-13.14 (-2.68)	0.09 (0.04)	0.93 (0.51)
(k=3)	-19.72 (-3.36)	-16.24 (-3.01)	-0.80 (-0.31)	0.33 (0.14)
(k=4)	-22.99 (-3.63)	-19.28 (-3.26)	-1.95 (-0.59)	-0.50 (-0.17)
(k=5)	-26.13 (-3.87)	-22.24 (-3.50)	-3.03 (-0.80)	-1.32 (-0.39)
EG (k=1)	-15.24 (-2.77)	-12.40 (-2.47)	-0.54 (-0.27)	0.24 (0.15)
(k=2)	-21.07 (-3.37)	-17.74 (-3.04)	-2.83 (-1.03)	-1.58 (-0.68)
(k=3)	-26.31 (-3.80)	-22.72 (-3.49)	-5.29 (-1.49)	-3.76 (-1.21)
(k=4)	-31.04 (-4.19)	-27.10 (-3.87)	-7.86 (-1.87)	-6.25 (-1.59)
(k=5)	-35.56 (-4.55)	-31.68 (-4.23)	-10.72 (-2.24)	-8.67 (-1.96)

Note: See note to Table 1; k denotes the number of I(1) regressors.

Table 3

Size Adjusted Powers of 5% Tests
(percentages)

		s=0.05	s=1.00	s=5.00	s=20.00
	β=-0.05				
Test					
α=0.1	C-O	30 (30)	8 (7)	0 (0)	0 (0)
	ECM	22 (18)	14 (23)	0 (88)	0 (100)
	EG	14 (15)	11 (11)	5 (5)	4 (4)
α=0.9	C-O	30 (30)	28 (28)	16 (16)	1 (1)
	ECM	21 (17)	21 (17)	18 (19)	5 (48)
	EG	14 (14)	13 (14)	12 (12)	7 (7)
	β=-0.10				
α=0.1	C-O	69 (68)	8 (8)	0 (0)	0 (0)
	ECM	53 (54)	44 (67)	8 (100)	0 (199)
	EG	36 (36)	30 (30)	18 (18)	17 (17)
α=0.9	C-O	70 (70)	67 (67)	27 (26)	1 (1)
	ECM	53 (54)	53 (55)	51 (53)	30 (94)
	EG	37 (37)	37 (38)	34 (35)	22 (23)

<u>Note</u>: Rejection rates for the t-ratio version of the tests are given in parenthesis.

Table 4

Critical values of the (t-ratio) ECM Test
Different number of regressors

	T	Size 0.01	0.05	0.10	0.25
		A. (with constant)			
	25	-4.12	-3.35	-2.95	-2.36
(k=1)	50	-3.94	-3.28	-2.93	-2.38
	100	-3.92	-3.27	-2.94	-2.40
	500	-3.82	-3.23	-2.90	-2.40
	∞	-3.78	-3.19	-2.89	-2.41
	25	-4.53	-3.64	-3.24	-2.60
(k=2)	50	-4.29	-3.57	-3.20	-2.63
	100	-4.22	-3.56	-3.22	-2.67
	500	-4.11	-3.50	-3.10	-2.66
	∞	-4.06	-3.48	-3.19	-2.65
	25	-4.92	-3.91	-3.46	-2.76
(k=3)	50	-4.59	-3.82	-3.45	-2.84
	100	-4.49	-3.82	-3.47	-2.90
	500	-4.47	-3.77	-3.45	-2.90
	∞	-4.46	-3.74	-3.42	-2.89
	25	-5.27	-4.18	-3.68	-2.90
(k=4)	50	-4.85	-4.05	-3.64	-3.03
	100	-4.71	-4.03	-3.67	-3.10
	500	-4.62	-3.99	-3.67	-3.11
	∞	-4.57	-3.97	-3.66	-3.10
	25	-5.53	-4.46	-3.82	-2.99
(k=5)	50	-5.04	-4.43	-3.82	-3.18
	100	-4.92	-4.30	-3.85	-3.28
	500	-4.81	-4.39	-3.86	-3.32
	∞	-4.70	-4.27	-3.82	-3.29

B. (with constant and trend)

	25	-4.77	-3.89	-3.48	-2.88
(k=1)	50	-4.48	-3.78	-3.44	-2.92
	100	-4.35	-3.75	-3.43	-2.91
	500	-4.30	-3.71	-3.41	-2.91
	∞	-4.27	-3.69	-3.39	-2.89
	25	-5.12	-4.18	-3.72	-3.04
(k=2)	50	-4.76	-4.04	-3.66	-3.09
	100	-4.60	-3.98	-3.66	-3.11
	500	-4.54	-3.94	-3.64	-3.11
	∞	-4.51	-3.91	-3.62	-3.10
	25	-5.42	-4.39	-3.89	-3.16
(k=3)	50	-5.04	-4.25	-3.86	-3.25
	100	-4.86	-4.19	-3.86	-3.30
	500	-4.76	-4.15	-3.84	-3.31
	∞	-4.72	-4.12	-3.82	-3.29
	25	-5.79	-4.56	-4.04	-3.26
(k=4)	50	-5.21	-4.43	-4.03	-3.39
	100	-5.07	-4.38	-4.02	-3.46
	500	-4.93	-4.34	-4.02	-3.47
	∞	-4.89	-4.30	-4.00	-3.45
	25	-6.18	-4.76	-4.16	-3.31
(k=5)	50	-5.37	-4.60	-4.19	-3.53
	100	-5.24	-4.55	-4.19	-3.66
	500	-5.15	-4.54	-4.20	-3.69
	∞	-5.11	-4.52	-4.18	-3.67

Notes: The asymptotic (∞) critical values are based on 5,000 simulations; k denotes the number of I(1) regressors in the cointegrating regression.