

ERROR-CORRECTION MECHANISM TESTS FOR COINTEGRATION IN A SINGLE-EQUATION FRAMEWORK

BY ANINDYA BANERJEE, JUAN J. DOLADO AND RICARDO MESTRE

*Wadham College and Institute of Economics and Statistics, University of Oxford,
Universidad Carlos III de Madrid and Research Department, Bank of Spain*

First version received July 1996

Abstract. A new test is proposed for cointegration in a single equation framework where the regressors are weakly exogenous for the parameters of interest. The test is denoted as an error correction mechanism (ECM) test and is based upon the ordinary least squares coefficient of the lagged dependent variable in an autoregressive distributed lag model augmented with leads of the regressors. The limit distributions of the standardized coefficient and t ratio versions of the ECM tests are obtained and critical values are provided. These limit distributions do not depend upon nuisance parameters but they depend on the number of regressors. Finally, we compare their power properties with those of other cointegration tests available in the literature and find the circumstances under which the ECM tests have a better performance.

Keywords. Cointegration tests; error correction models; power properties; common factor restrictions.

1. INTRODUCTION

A new test for cointegration in a single-equation framework is proposed. The new test is based on the coefficient of the lagged dependent variable in an autoregressive distributed lag (ADL) model advocated by Hendry and Richard (1982) and Hendry (1987). This class of models has traditionally been used in the empirical literature to seek a tentatively adequate data characterization that encompasses rival models, displays parameter constancy, has martingale difference errors with respect to a selected information set and parsimoniously orthonormalizes the regressors. As proved by Engle *et al.* (1983), weak exogeneity of the regressors for the parameters of interest is a sufficient condition for ordinary least squares (OLS) to provide asymptotically efficient estimates of the parameters in the conditional ADL model.

Recent papers by Phillips and Loretan (1991), Saikkonen (1991) and Hendry (1995) have extended the previous analysis to the case where regressors are $I(1)$ processes. A feature common to all of these papers is that they concentrate on the case of cointegration among the variables, proposing new methods to achieve asymptotically efficient estimates of the elements of the cointegrating vector. This strategy, which consists of adding leads of the regressors and the error-correction term to the conditional model, has proved quite successful

since, in contrast to the fully modified estimator of Phillips and Hansen (1990) where a semiparametric correction is needed, the correction of the OLS estimator and the computation of the final estimator are accomplished simultaneously in the time domain.

In this paper, instead of operating under the alternative hypothesis of cointegration, we choose to work under the null hypothesis of non cointegration. We derive a test for cointegration denoted the error correction mechanism (ECM) test, which benefits from some of the advantages of ADL models described above. The procedure depends upon the significance of the lagged dependent variable since this is equivalent to testing the significance of the error correction terms in the ECM reparameterization of the model. This type of test has been previously suggested by Banerjee *et al.* (1986, 1993) and Boswijk (1991). However, an extensive study of its properties is not yet available in the literature.

The ECM test, in both its normalized bias and t ratio versions, has a limit distribution that does not depend on nuisance parameters. However, it is not dimension invariant since its limit distribution shifts with the number of regressors. Alternatively, Hansen (1990) has proposed a cointegration test based upon the Cochrane Orcutt (1949) estimation procedure whose limit distribution is dimension invariant and follows the unit root distribution simulated by Fuller (1976). Nevertheless, the latter test, along with other well known cointegration statistics such as the Engle and Granger (1987) test, suffers in finite samples from imposing potentially invalid *common factor* restrictions. Consequently, if these restrictions are not satisfied, the two latter types of test may have poor power properties. Since the ECM test does not suffer from this problem, there may be large advantages in its use.

Furthermore, as a by product of the power analysis undertaken under a sequence of local alternatives of cointegration, we show that the t ratio form of the ECM test may have better power properties than the normalized bias form, particularly when the common factor restrictions are grossly violated.

Lastly, it is important to note that, although the framework of reference is restricted to single equation conditional error correction models with a potentially unique cointegrating relationship, empirical studies abound where this is shown to be the case, e.g. money demand equations, consumption equations etc. (cf. Hendry *et al.* 1984). Therefore, we believe that the applicability of the ECM test in applied work may be quite important. We therefore provide critical values based upon the limit distribution of the test for a large range of the number of regressors.

The outline of the rest of the paper is as follows. In Section 2 we present the data generation process (DGP) for the simplified case where the regressors are assumed to be strictly exogenous, and we derive the limit distributions of the ECM test statistics under the null hypothesis of non cointegration. Section 3 gives the corresponding limit distribution under a sequence of alternatives representing near non cointegration. Section 4 offers a comparison of the ECM test with other cointegration tests often used in applied work, stressing the

problem of imposing possibly invalid common factor restrictions. In Section 5 we consider generalizations of the ECM tests to more realistic cases where the regressors are only assumed to be weakly exogenous. Section 6 provides Monte Carlo finite sample evidence about the relative performance of the ECM tests with respect to the other cointegration tests discussed in the paper. Concluding remarks are given in Section 7.

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 $\text{BM}(\Omega)$ refers to a Brownian motion with long run covariance matrix Ω ; $x_T = o_p(\phi_T)$ denotes that the sequence of random variables $\{x_T\}$ is of smaller order in probability than ϕ_T . Arguments of functionals on the space $[0, 1]$ are frequently suppressed so that $\int_0^1 B^2(r) dr$ is written as $\int B^2$ to reduce notation. Proofs of important results are contained in the Appendix.

2. A SIMPLE DGP AND THE ECM TEST STATISTIC

By using a simple DGP, based upon a single equation ECM model, this section describes the ECM testing procedure.

This DGP has been used elsewhere (cf. Hendry and Richard, 1982; Kremers *et al.*, 1992; Banerjee *et al.* 1993) and has the form

$$\Delta y_t = \alpha' \Delta x_t + \beta(y_{t-1} - \lambda' x_{t-1}) + \epsilon_t \quad (1)$$

$$\Delta x_t = u_t \quad t = 1, \dots, T \quad (2)$$

where α , λ and x_t are $k \times 1$ vectors of parameters and explanatory variables. The regressand y_t is a univariate process and β is a scalar; the initial conditions are, without loss of generality, set to zero and T is the sample size. The elements of x_t correspond to different regressors. The more general case where lags of Δx_t and Δy_t are allowed will be considered below. For the time being, we will assume that x_t is *strictly exogenous*, so that

$$\frac{1}{k} \begin{pmatrix} \epsilon_t \\ u_t \end{pmatrix} \sim \text{i.i.d.} \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix} \begin{pmatrix} \sigma_\epsilon^2 & 0' \\ 0 & \Sigma_u \end{pmatrix} \right] \equiv \text{i.i.d.}(0, \Sigma)$$

where $\Sigma_u > 0$ to avoid cointegration among the regressors (for a brief discussion of the possibility of cointegrated regressors, see below). With this set up, the partial sum processes $S_T(r) = T^{-1/2} \Sigma_1^{\text{[Tr]}}(\epsilon_t, u_t)'$ satisfy the multivariate invariance principles (cf. Phillips and Durlauf 1986).

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

where $B_r = (B_\epsilon(r), B_u(r))'$ is a $k + 1$ vector standardized Brownian motion, i.e. $\text{BM}(\text{I})$.

We further assume that $-2 < \beta \leq 0$. In this DGP, y_t and x_t are cointegrated when $-2 < \beta < 0$, while they are non integrated when $\beta = 0$. Thus, tests of

cointegration must rely upon some estimate of the parameter β . Under the simplifying assumption that x_t is strictly exogenous, non linear least squares (NLS) can be applied to (1) yielding consistent and asymptotically efficient estimates of α , β and λ . The ECM test statistic for cointegration, as suggested by Banerjee *et al.* (1986) and Boswijk (1991), is based upon estimating (1) by NLS and testing $H_0: \beta = 0$. Alternatively, Banerjee *et al.* (1993), drawing upon results from Kiviet and Phillips (1992), show that a parameter free distribution for the estimator of β can be achieved if x_{t-1} is added to (1), which is then estimated by OLS. This is so since, under the alternative hypothesis of cointegration, the true cointegrating slope λ is implicitly estimated when x_{t-1} is included as an additional regressor. Hence, according to this procedure, β is estimated by OLS from the unrestricted dynamic model

$$\Delta y_t = \alpha' \Delta x_t + \beta y_{t-1} + \theta' x_{t-1} + \epsilon_t = \alpha' \Delta x_t + \pi' w_{t-1} + \epsilon_t \quad (1')$$

where $w'_t = (y_t, x'_t)$ and $\pi' = (\beta, \theta')$.

Since $\beta(1, \lambda') = \pi'$, the non cointegrating restriction $\beta = 0$ implies $\pi = 0$ and so the ECM test can be based upon the OLS estimator of β in (1') or on its t ratio, denoted $\hat{\beta}_E$ and t_E , respectively. Thus, letting y and Δy be $T \times 1$ vectors of observations on y_t and Δy_t , the ECM estimator and its t ratio are defined by

$$\hat{\beta}_E = (y'_{-1} M y_{-1})^{-1} y'_{-1} M \Delta y \quad (3)$$

and

$$t_E = (\hat{\sigma}_\epsilon^{-2} y'_{-1} M y_{-1})^{1/2} \hat{\beta}_E \quad (3')$$

where $M = I - V(V'V)^{-1}V'$ and V is a $T \times 2k$ matrix of observations on $v_t = (\Delta x'_t, x'_{t-1})$, $\hat{\sigma}_\epsilon^2 = T^{-1} \sum_1^T \hat{\epsilon}_t^2$ and $\hat{\epsilon}_t$ is the OLS residual in (1').

Then, the following proposition holds.

PROPOSITION 1. For DGP (1) (2) under the null hypothesis of non cointegration ($\beta = 0$)

$$T \hat{\beta}_E \Rightarrow \left(\int \overline{B_\epsilon^2} \right)^{-1} \int \overline{B_\epsilon} dB_\epsilon$$

and

$$t_E \Rightarrow \left(\int \overline{B_\epsilon^2} \right)^{-1/2} \int \overline{B_\epsilon} dB_\epsilon$$

with

$$\overline{B_\epsilon} = B_\epsilon \left(\int B_u B_\epsilon \right)' \left(\int B_u B_u' \right)^{-1} B_u.$$

Note that $\overline{B_\epsilon}$ is the residual from the continuous time regression of B_ϵ on B_u . Thus, although the previous limit distributions are free of nuisance parameters,

they depend upon the number of elements (k) in x_t , as reflected by the presence of B_u in \overline{B}_ϵ , implying that corresponding test statistics are not dimension invariant.

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

The alternative hypothesis is that of cointegration which, for DGP (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 (1987, 1988) for near integrated time series can be applied to sharpen the results on the asymptotic power of the ECM tests.

To proceed with the analysis of local power we consider the following parameterization of the β coefficient:

$$\beta = 1 - \exp(c/T) \approx T^{-1}c. \quad (4)$$

In (4), c is a fixed scalar. We call time series that are generated by DGP (1) (2), with β as in (4), near non cointegrated processes, following the terminology introduced by Phillips (1988) 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 $c > 0$, (4) represents a local alternative to H_0 , so that the rate of approach is controlled and the effect of the alternative hypothesis on the limit distribution of the test statistics, based on the previous DGP, is directly measurable in terms of the non centrality parameters c .

To develop the analysis of local power, it is also useful to define the disturbance

$$e_t = (\alpha - \lambda)'u_t + \epsilon_t \quad (5)$$

such that, under the previous assumptions about u_t and ϵ_t , $E(e_t^2) = \sigma_\epsilon^2 + (\alpha - \lambda)' \Sigma_u (\alpha - \lambda)$. Then use is made of the diffusion process

$$K(r) = \int_0^r \exp \{c(r - s)\} dB(s) \equiv B(r) + c \int_0^r \exp \{c(r - s)\} B(s) ds \quad (6)$$

associated with the standardized disturbances ϵ_t , u_t and e_t , denoted K_ϵ , K_u and K_e , respectively. Note that, if $c = 0$, then $K \equiv B$.

Using (5) and (6), it is possible to show the following result.

PROPOSITION 2. For DGP (1) (2) and (4), under the alternative hypothesis of near non cointegration ($c > 0$)

$$T\hat{\beta}_E \Rightarrow c + \frac{\sigma_\epsilon}{\sigma_e} \left(\int \bar{K}_e^2 \right)^{-1} \int \bar{K}_e dB_\epsilon \equiv \Phi_E$$

$$t_E \Rightarrow \left(\sigma_\epsilon^{-2} \sigma_e^2 \int \bar{K}_e^2 \right)^{1/2} \Phi_E$$

with $\bar{K}_e = K_e \left(\int B_u K_e \right)' \left(\int B_u B_u' \right)^{-1} B_u$.

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

Although the comparison of the asymptotic distributions under the local alternative hypothesis is cumbersome, given the complexity of the Wiener functionals derived above, some results can be obtained using the relationship in (5). To illustrate the main result, let us simplify the analysis by assuming that there is a single regressor, i.e. $k = 1$. Then, given the relationship between e_t , ϵ_t and u_t that we repeat for convenience

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

we will define a signal to noise ratio $q = (\alpha - \lambda)s$, with $s = \sigma_u/\sigma_\epsilon$, corresponding to the ratio of the (square root of the) variance of $(\alpha - \lambda)u$, 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 this definition, the following proposition holds.

PROPOSITION 3. For DGP (1) (2) and (4), with $k = 1$, under the alternative hypothesis of near non cointegration ($c > 0$)

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

and

$$t_E \Rightarrow c(1 + q^2)^{1/2} \left(\int \bar{K}_e^2 \right)^{1/2} + \left(\int \bar{K}_e^2 \right)^{-1/2} \int \bar{K}_e dB_\epsilon + o_p(q^{-1}).$$

Various interesting properties arise from Proposition 3. First, asymptotically, as $q \uparrow \infty$, i.e. $\alpha \neq \lambda$ and $s \uparrow \infty$, the ECM test based upon the normalized bias has a slope equal to (minus) unity; since the limit distribution of $T\hat{\beta}_E$ is independent of q under the null and degenerates around c under the local alternative, the lower 5% tail of the distribution under the null will tend to be to the left of c . Thus, we should observe very low power of the test based upon the coefficient when q is large. Second, the limit distribution of the ECM test based upon the t ratio has a stochastic slope that depends upon q and does not degenerate around a single value, as is the case of the test based upon the normalized bias. Thus, when q is sizeable, the power of the t ratio test will be greater than that of the normalized bias test. This is an interesting result since,

as shown by Phillips and Ouliaris (1990), under a fixed alternative hypothesis the normalized bias test has non centrality which grows at rate T while the non centrality parameter of the t ratio diverges at rate $T^{1/2}$. However, as shown in Section 6, for reasonable sample sizes the power of the t test may be greater if q is sufficiently large.

4. COMPARISON WITH OTHER TEST STATISTICS FOR COINTEGRATION

Among the already very large collection of cointegration tests available in the literature (cf. Banerjee *et al.*, 1993), we wish to compare the power properties of the ECM test statistics with those of two popular test statistics for cointegration in a single equation framework. These are the Engle Granger (1987) test statistic and Hansen's (1990) Cochrane Orcutt test statistic. In what follows, we will denote these tests as EG and CO, respectively.

As is well known, the EG test is based upon a two step procedure. In the first step a static OLS regression of y_t on x_t is implemented, yielding an estimate of λ , say $\hat{\lambda}$. Next, the cointegration test is based upon the normalized bias or the t ratio of β in the regression

$$\Delta y_t - \hat{\lambda}' \Delta x_t = \beta(y_{t-1} - \hat{\lambda}' x_{t-1}) + \tilde{e}_t \quad (7)$$

where \tilde{e}_t is an error term such that $\tilde{e}_t = e_t + o_p(1)$.

The CO test is similar in spirit to the EG test except that the estimation of λ and the cointegration test are accomplished simultaneously, by estimating

$$\Delta y_t - \lambda' \Delta x_t = \beta(y_{t-1} - \lambda' x_{t-1}) + e_t \quad (8)$$

by NLS and testing for the significance of the NLS estimate of β . Hansen (1990, Theorem 2) proves that the normalized bias and the t ratio, denoted $T\hat{\beta}_{CO}$, and t_{CO} , have the limit Dickey Fuller distributions under the null hypothesis of non cointegration. Thus, this test has the advantage over the ECM and EG statistics that its limit distribution is independent of the dimension of the vector x_t , a feature which according to Hansen (1990) may improve its relative power properties.

Nonetheless, as pointed out by Kremers *et al.* (1992) and Hansen (1995), both the EG and the CO test suffer from the problem of imposing possibly invalid common factor restrictions. This problem can be readily reviewed by considering the alternative representation of equation (1)

$$\Delta y_t = \alpha' \Delta x_t + \beta(y_{t-1} - \lambda' x_{t-1}) + \epsilon_t = \lambda' \Delta x_t + \beta(y_{t-1} - \lambda' x_{t-1}) + e_t \quad (9)$$

with e_t defined as in (5). As an extreme example, let $\sigma_e^2 \uparrow 0$ but $\alpha \neq \lambda$ and let $(\alpha - \lambda)' \Sigma_u(\alpha - \lambda)$ be '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 ECM test statistics will be (arbitrarily) large. However, since $\sigma_e^2 = (\alpha - \lambda)' \Sigma_u(\alpha - \lambda) + \sigma_e^2$, the estimation of β in the CO and EG procedures will be much more imprecise, having an adverse effect on the power of both tests. In

other words, where $\alpha \neq \lambda$, invalid common factor restrictions are imposed in the estimation procedure underlying the latter tests, a feature which could have serious adverse effects on their power properties.

5. GENERALIZATIONS OF THE ECM TEST STATISTICS

In the previous sections, we have assumed for simplicity that the vector of regressors x_t is strictly exogenous in the conditional model (1). However, this is a very strong assumption. As proved by Engle *et al.* (1983), all that is needed for OLS to be an asymptotically efficient estimation method for the parameters in (1') is that the x_t are weakly exogenous for the parameters of interest $\psi = (\alpha', \beta, \theta')$. This weaker assumption is fairly well used in practice (cf. Hendry *et al.*, 1984) and allows for the presence of lags of Δx_t and Δy_t in the conditional model (1). To extend the ECM test statistics to this more general set up, we will consider an extended DGP consisting of the ADL conditional model

$$\gamma(L)\Delta y_t = \alpha(L)\Delta x_t + \beta(y_{t-1} - \lambda'x_{t-1}) + \epsilon_t \quad (1'')$$

where $\gamma(L)$ and $\alpha(L)$ are polynomials in the lag operator L ; $\gamma(L)$ is a scalar polynomial of order m_y and $\alpha(L) = (\alpha_1(L), \dots, \alpha_k(L))$ is a vector polynomial order of (m_1, \dots, m_k) .

The marginal process for Δx_t is as in (2) with u_t being now a stationary process with zero mean and continuous spectral density $f_{uu}(\omega)$, whose covariance function is absolutely summable. In this more general framework, the partial sum process constructed from the $k+1$ vector $v_t = (\epsilon_t, u_t')$ will now converge to a vector Brownian process $\text{BM}(\Omega)$ with long run covariance matrix given by

$$\Omega = \begin{pmatrix} \omega_{\epsilon\epsilon} & \omega'_{uc} \\ \omega_{uc} & \Omega'_{uu} \end{pmatrix} = \Sigma + A + A' = \Delta + A'; \quad \Sigma = \begin{pmatrix} \sigma_\epsilon^2 & 0' \\ 0 & \Sigma_u \end{pmatrix}$$

where $\omega_{\epsilon\epsilon} = \sigma_\epsilon^2$, $\omega'_{uc} = E(\epsilon_0 u_0') + \sum_{s=1}^{\infty} E(\epsilon_0 u_s' + \epsilon_s u_0')$ and $\Omega_{uu} = E(u_0 u_0') + \sum_{s=1}^{\infty} E(u_0 u_s' + u_s u_0')$.

Under the assumption of weak exogeneity of x_t , we have that $E(\epsilon_t/x_{t-i}) = 0$ for all $i \geq 0$. Thus, $\Delta_{21} = \sum_{s=0}^{\infty} E(u_0 \epsilon_s) = 0$, but there is no guarantee that $\sum_{s=0}^{\infty} E(\epsilon_0 u_s') = 0$. This would be implied by the stronger assumption that x_t is strictly exogenous for equation (1). Since $\Delta_{21} = 0$, the 'second order' biases stressed by Phillips and Hansen (1990) will be absent in the distribution of the ECM test statistics, as in Proposition 1. However, note that in this more general case B_ϵ and B_u are no longer independent Brownian motions. To illustrate this feature, consider the following example. For $k=1$, let $\Delta x_t = u_t = \gamma \Delta y_{t-1} + \eta_t = \gamma(a u_{t-1} + \epsilon_{t-1}) + \eta_t$ with $E(\epsilon_t \eta_s) = 0$ for all t and s . Then, x_t will be weakly exogenous for the parameters of interest in (1), but the long run covariance between B_ϵ and B_u will be $\gamma(1 - \alpha\gamma)^{-1} \sigma_\epsilon^2$, under the null hypothesis of non cointegration. This is so since $\sum_{k=0}^{\infty} E(\epsilon_0 u_k') \neq 0$, implying that the limit distributions obtained in Proposition 1 will now depend on

nuisance parameters ($\omega_{u\epsilon}$). Hence the corresponding tests will not be asymptotically similar. Therefore, in principle, the computation of critical values in this more general case is problematic.

To overcome the problem of lack of similarity, we follow the strategy proposed by Phillips and Loretan (1991) and Saikkonen (1991) of correcting for serial correlation by augmenting the conditional model in (1') with future values of Δx_t . Given the stationarity of u_t one would expect that the very remote future values of Δx_t only have a negligible impact on Δy_t and can therefore be ignored.

Under the previous conditions on the error terms, we may write

$$\epsilon_t = \sum_1^{\infty} a'_j u_{t+j} + \xi_t \quad (10)$$

where $\sum_1^{\infty} \|a_j\| < \infty$ and ξ_t is a stationary process such that $E(\xi_t u'_s) = 0$ for all t and s .

Since the sequence $\{a_j\}$ in (10) is absolutely summable, we have that $a_j = 0$ for $|j| > S$ and S large enough. Thus, the ECM statistics may be computed from the model

$$\gamma(L)\Delta y_t = \alpha(L)'\Delta x_t + \beta y_{t-1} + \theta'x_{t-1} + \sum_1^S a'_j \Delta x_{t+j} + \dot{\xi}_t \quad (11)$$

with $\dot{\xi}_t = \xi_t + \sum_{|j| > S} a'_j \Delta x_{t+j}$ and in agreement with the assumption previously used by Said and Dickey (1984) we shall assume that

$$S^3/T \rightarrow 0 \quad \text{and} \quad T^{1/2} \sum_j \|a_j\| \rightarrow 0 \text{ for } |j| > S.$$

Using similar arguments to those in Saikkonen (1991, Theorem 4.1) it is straightforward to show that the limit distributions of $T\hat{\beta}_E$ and t_E , on the basis of regression model (11), are identical to those derived in Proposition 1. In practice the value of S should be large enough to ensure that the coefficients a_j are effectively zero for $|j| > S$ while, at the same time, respecting the constraint that the least squares estimation of (11) is not feasible if S is too large compared with the sample size. In empirical applications, some experimentation with a few values of S is advisable. Although a thorough discussion of this issue is beyond the scope of this paper, some experimentation along the lines of Stock and Watson (1993) seemed to suggest that the choice $S = 1$ or $S = 2$ for $T = 100$ had good size properties.

Next, it is important to note that, although deterministic terms have been ignored in the previous analysis for simplicity, the data may be demeaned, or demeaned and detrended, before applying the ECM tests for cointegration. The limit distributions of the various tests discussed in the paper in such cases are of the same form as in Proposition 1, except that Brownian motions are replaced by the appropriate Brownian bridges. Given the advantages of using the t ratio for the ECM test, as discussed in Section 4, the asymptotic critical

values for the ECM t ratio are reported in Table I up to five regressors. In order to analyse the finite sample distribution of those tests, critical values for four different sample sizes ($T = 25, 50, 100$ and 500) are also presented. Since, as discussed in Section 1, there are many examples in applied work of single equation conditional models with weakly exogenous regressors for the parameters of interest, we think that the above critical values may be widely applicable.

It is also noteworthy that the common factor problem of the EG and CO test statistics for cointegration remains in this more general set up. Furthermore, the augmented Dickey Fuller version of the EG test and the semiparametric version proposed by Phillips and Ouliaris (1990) do not solve the problem. Since this argument is similar to that given by Kremers *et al.* (1992) where the potential cointegrating vector is assumed to be known *a priori*, we will simply summarize it briefly in what follows.

TABLE I
CRITICAL VALUES OF THE (t -RATIO) ECM TEST WITH DIFFERENT NUMBERS OF
REGRESSORS

		Size			
	T	0.01	0.05	0.10	0.25
<i>A. With constant</i>					
$k = 1$	25	4.12	3.35	2.95	2.36
	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
$k = 2$	25	4.53	3.64	3.24	2.60
	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
$k = 3$	25	4.92	3.91	3.46	2.76
	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
$k = 4$	25	5.27	4.18	3.68	2.90
	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
$k = 5$	25	5.53	4.46	3.82	2.99
	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

TABLE I
(continued)

<i>T</i>	Size				
	0.01	0.05	0.10	0.25	
<i>B. With constant and trend</i>					
<i>k</i> = 1	25	4.77	3.89	3.48	2.88
	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
<i>k</i> = 2	25	5.12	4.18	3.72	3.04
	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
<i>k</i> = 3	25	5.42	4.39	3.89	3.16
	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
<i>k</i> = 4	25	5.79	4.56	4.04	3.26
	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
<i>k</i> = 5	25	6.18	4.76	4.16	3.31
	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

For example, if we consider the conditional model (1''), the error term in the CO and EG testing procedures will be

$$e_t = [\alpha(L) \quad \gamma(L)\lambda]' u_t + \epsilon_t$$

which obviously need not be white noise. Indeed, in general it will follow a moving average (MA) process, whose serial correlation could be accounted for by means of the semiparametric corrections proposed by Phillips and Ouliaris (1990). It is known, however, that when the roots of such MA processes are close to being on the unit circle, these tests may suffer from severe size distortions (cf. Schwert, 1989). This problem does not arise when using the ECM statistics.

Finally, the possibility of cointegrated regressors may arise as a practical matter (see Granger and Lee, 1990). If x_t is correcting the errors of cointegrating relationships involving only x_{t-1} , then weak exogeneity still holds (see Hunter, 1990). In that case, given that the cointegration vector does not

include y_t , the proposed test is still applicable, except that the dimension of B_{it} in the ECM tests will be smaller than in the unrestricted version of the tests. Thus, using the critical values for the latter type of tests will lead to a conservative test. If, on the other hand, the cointegrating vector linking the x_t is known, then to achieve similarity, future values of the $I(0)$ cointegrating error may have to be added to the regression model in (11), choosing a value of k corresponding to the number of non cointegrated regressors in Table I.

6. FINITE-SAMPLE EVIDENCE

To provide finite sample evidence on the advantages of the ECM test statistics in comparison with the CO and EG tests, a small set of Monte Carlo experiments was conducted with (1) and (2) as the DGP, using 25 000 replications generated in GAUSS. A single exogenous regressor, i.e. $k = 1$, is used for illustrative purposes. Data were generated with the normalization $\sigma_\epsilon = 1$, without loss of generality, with three parameters (s , α , β) and the sample size T as experimental design variables. In this study we chose

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

$$\alpha = (0.1, 0.9)$$

$$\beta = (0.05, 0.10 \text{ (cointegration in both cases)})$$

$$T = 100.$$

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, the value of the cointegrating slope λ was fixed equal to 1 under the alternative hypothesis of cointegration. The choices of the short run coefficient (α) attempt to capture a smaller value ($\alpha = 0.1$) and a similar value ($\alpha = 0.9$) to the one chosen for λ ; the closer α and λ are, the closer the common factor restriction is to being satisfied. Combining the values of α and λ with those for s we obtain a wide range of values q , ranging from 0.005 to 18.

Table II reports the power of the three tests for the selected range of values for α and s , when $\beta = 0.05$ and $\beta = 0.10$. To control for finite sample biases, critical values were simulated under the hypothesis $H_0: \beta = 0$, and the reported powers are size adjusted. The results seem to be consistent with the discussion in Section 3. When q is low and c is small relative to T , $c = 5$ ($\beta = 0.05$), the ECM test, both in its normalized bias and t ratio versions, seems to be slightly less powerful than the CO test, reflecting the problem of dimensionality stressed by Hansen (1990). However, as q increases, either because α becomes different from λ or because s rises, the ECM test becomes the most powerful. Furthermore, in agreement with the degeneration of the

TABLE II
SIZE-ADJUSTED POWERS OF 5% TESTS (PERCENTAGES)

	Test	$s = 0.05$	$s = 1.00$	$s = 5.00$	$s = 20.00$
$\beta = 0.05$ $\alpha = 0.1$	CO	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)
$\alpha = 0.9$	CO	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)
$\beta = 0.10$ $\alpha = 0.1$	CO	69 (68)	8 (8)	0 (0)	0 (0)
	ECM	53 (54)	44 (67)	8 (100)	0 (100)
	EG	36 (36)	30 (30)	18 (18)	17 (17)
$\alpha = 0.9$	CO	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)

Note: Rejection rates for the t -ratio version of the tests are given in parentheses.

limit distributions of the coefficient version of the tests, their absolute power decreases as q increases. This is clearly not the case when we examine the t ratio version of the tests, where the ECM test shifts its asymptotic 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 while the CO test almost does not reject at all. As regards the EG test, the results indicate that its power also decreases as q increases, though at a lower rate than the power of the CO 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, since in contrast to the CO test the EG test is not dimension invariant.

Finally, although our testing procedure is designed in a single equation framework, a comparison with Johansen's (1991) procedure would be helpful. Indeed, the ECM procedure is a special case of Johansen's for a system in which the cointegrating vectors appear only in the equation of interest. Although an extensive study on the performance of Johansen's test is beyond the scope of this paper, we have carried out a small Monte Carlo study for the case $\beta = 0.10$ in the bivariate system consisting of equations (8) and (2). The error variances have been normalized to unity, yielding a covariance $\rho = (1 + q^2)^{-1/2}q$. The trace statistic $LR(0)$ is asymptotically equal to $T\hat{\lambda}_1 + T\hat{\lambda}_2$ where $\hat{\lambda}_1$ and $\hat{\lambda}_2$ are eigenvalues computed from a characteristic equation such as (2.11) in Johansen (1991). Suppose $\hat{\lambda}_1 > \hat{\lambda}_2$. Johansen (1991) shows that, if there is a unique cointegrating vector, $T\hat{\lambda}_2$ has an asymptotic distribution while $\hat{\lambda}_1$ converges to a positive constant. It is easy to see that in our model

$$\hat{\lambda}_1 \rightarrow \lambda_1 = \{1 - (\rho^2)(2 + \beta)/\beta\}^{-1}. \quad (12)$$

This in turn implies that, for a fixed T , LR(0) will (correctly) reject the null hypothesis on non cointegration more easily if $|\rho|$ is larger and that rejection will be harder the smaller is β . Using the 5% critical value in Osterwald Lenum (1992) with $T = 100$, we find that, when $s = 1$, the rejection rate of LR(0) is 46% (16%) when $\alpha = 0.1$ (0.9). When $s = 5$, it turns out to be 91% (18%). Further experiments, available upon request, show that for high values of s the LR(0) test performs slightly worse than the ECM t test and that the behaviour is much worse for low values of s .

7. CONCLUSIONS

Testing for cointegration has become an important facet of empirical analysis of economic time series in recent years and various tests are being used. In this paper we propose a new test, denoted the ECM test, in a single equation framework. The limit distribution of this test, in both its ‘normalized version’ and t ratio version, does not depend upon nuisance parameters but does depend on the dimension of the system. Critical values are therefore provided. Its power properties are compared with those of other popular tests of cointegration. Specifically, we concentrate on the CO and EG testing procedures. The CO test is dimension invariant whereas the EG test is not. However, both tests impose possibly invalid common factor restrictions in the estimation underlying the tests. We show that when the restrictions are invalid, the power properties of the CO and EG tests may be very poor in comparison with the ECM test, which does not impose those restrictions.

Moreover, as a by product of the analysis, we show that the t ratio form of the ECM test may be preferable to the normalized bias form, under the alternative hypothesis of cointegration, when the common factor restrictions do not hold. The results are obtained for a simple DGP and are then shown to extend to more general cases.

APPENDIX

The analysis contained in this appendix draws on a number of well known results in Phillips (1987, 1988) and Phillips and Ouliaris (1990). Under the null hypothesis of non cointegration the DGP(H_0) is given by

$$\begin{aligned} \Delta y_t &= \alpha' \Delta x_t + \epsilon_t \\ \Delta x_t &= u_t \end{aligned}$$

with

$$\begin{pmatrix} \epsilon_t \\ u_t \end{pmatrix} \sim \text{i.i.d.} \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix} \begin{pmatrix} \sigma_\epsilon^2 & 0' \\ 0 & \Sigma_u \end{pmatrix} \right].$$

The results do not depend on the initialization, so let us define $S_{\epsilon t} = \sum_{i=1}^t \epsilon_i$ and $S_{u t} = \sum_{i=1}^t u_i$, where the sums run from 1 to t and S_{ϵ} and S_u are $T \times 1$ and $T \times k$ matrices of observations on $S_{\epsilon t}$ and $S_{u t}$, respectively. Note that $x = S_u \alpha$ and $y = S_u \alpha + S_{\epsilon}$. Let M_1 be the projection matrix $M_1 = I - x(x'x)^{-1}x'$.

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

- (a) $T^{-2}S'_{\epsilon-1}M_1S_{\epsilon-1} \Rightarrow \sigma_{\epsilon}^2 \int \bar{B}_{\epsilon}^2$
- (b) $T^{-1}S'_{\epsilon-1}M_1\epsilon \Rightarrow \sigma_{\epsilon}^2 \int \bar{B}_{\epsilon} dB_{\epsilon}$
- (c) $T^{-1}S'_{\epsilon-1}M_1\epsilon \rightarrow 0$
- (d) $T^{-1}\sum_1^T \epsilon_t u'_t \rightarrow 0$

where $\bar{B}_{\epsilon} = B_{\epsilon} (\int B_u B_u)' (\int B_u B_u)^{-1} B_u$.

Under the local alternative hypothesis of near non cointegration the DGP(H_{la}) is given by

$$\begin{aligned} \Delta z_t &= \beta z_{t-1} + \epsilon_t \\ \Delta x_t &= u_t \end{aligned}$$

with $\beta = T^{-1}c$; $z_t = y_t = \lambda'x_t + e_t = (\alpha - \lambda)'u_t + \epsilon_t$ and z and e are $T \times 1$ vectors of observations on z_t and e_t .

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

- (a) $T^{-2}z'z \Rightarrow \sigma_e^2 \int K_e^2$
- (b) $T^{-2}z'M_1z \Rightarrow \sigma_e^2 \int \bar{K}_e^2$
- (c) $T^{-1}z'_1M_1\epsilon \Rightarrow \sigma_e \sigma_{\epsilon} \int \bar{K}_e dB_{\epsilon}$

where $\bar{K}_e = K_e (\int B_u K_e)' (\int B_u B_u)^{-1} B_u$.

PROOF OF PROPOSITION 1. Let V be a $T \times 2k$ matrix of observations on $v_t = (\Delta x'_t, x'_{t-1})'$ and x_{-1} and Δx be $T \times k$ matrices of observations on x_t and x_{t-1} , respectively. Define the projection matrix $M = I - V(V'V)^{-1}V'$ such that, by partitioned inverses, $M = M_1 - M_1 \Delta x (\Delta x' M_1 \Delta x)^{-1} \Delta x' M_1$, where M_1 is defined above.

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

$$\begin{aligned} T\hat{\beta}_E &= (T^{-2}y'_{-1}M_1y_{-1})^{-1}(T^{-1}y'_{-1}M\Delta y) \\ &= (T^{-2}S'_{\epsilon-1}MS_{\epsilon-1})^{-1}(T^{-1}S'_{\epsilon-1}M\epsilon) \end{aligned}$$

since $y = S_u \alpha + S_{\epsilon}$, $\Delta y = u \alpha + \epsilon$ and M is orthogonal to x_{-1} and Δx . Using parts (a) to (d) of (R1) and the relationship between M and M_1 , we have

$$\begin{aligned} T^{-2}S'_{\epsilon-1}MS_{\epsilon-1} &= (T^{-2}S'_{\epsilon-1}M_1S_{\epsilon-1}) - T^{-1}(T^{-1}S'_{\epsilon-1}M_1u)(T^{-1}u'M_1u)^{-1}(T^{-1}u'M_1S_{\epsilon-1}) \\ &= T^{-2}S'_{\epsilon-1}M_1S_{\epsilon-1} + o_p(1) \end{aligned}$$

and

$$\begin{aligned} T^{-1}S'_{\epsilon-1}M\epsilon &= (T^{-1}S'_{\epsilon-1}M_1\epsilon) - (T^{-1}S'_{\epsilon-1}M_1u)(T^{-1}u'M_1u)^{-1}(T^{-1}u'M_1\epsilon) \\ &= T^{-1}S'_{\epsilon-1}M_1\epsilon + o_p(1) \end{aligned}$$

given (e) in (R1).

Next, using the limit distributions in (a) and (b) in (R1) yields the required result

$$T\hat{\beta}_E = (S'_{\epsilon-1}M_1S_{\epsilon-1})^{-1}S'_{\epsilon-1}M_1\epsilon + o_p(1) \Rightarrow \left(\int \bar{B}_{\epsilon}^2 \right)^{-1} \int \bar{B}_{\epsilon} dB_{\epsilon}. \quad (A1)$$

To prove that $\hat{\sigma}_\epsilon^2 \rightarrow \sigma_\epsilon^2$, define P as the $T \times (2k + 1)$ matrix of observations on $(\Delta x'_t, x'_{t-1}, y_{t-1})'$ and the projection matrix $M_p = I - P(P'P)^{-1}P'$. Then

$$\begin{aligned} \hat{\sigma}_\epsilon^2 &= T^{-1}\epsilon' M_p \epsilon = T^{-1}\epsilon'\epsilon - T^{-1}(T^{-1}\epsilon'P)(T^{-2}P'P)^{-1}(T^{-1}P'\epsilon) \\ &= T^{-1}\epsilon'\epsilon + o_p(1) \rightarrow \sigma_\epsilon^2. \end{aligned} \quad (\text{A2})$$

From (A1) and (A2), the distribution of the t ratio follows along the same lines, leading to the required results.

PROOF OF PROPOSITION 2. Let $z_t = y_t - \lambda'x_t$ and $\hat{z}_t = y_t - \hat{\alpha}'x_t$ where $\hat{\alpha}$ is the least squares estimator of α in (1'). Then

$$z_t = \hat{z}_t + (\hat{\alpha} - \alpha)'x_t = \hat{z}_t + o_p(1)$$

since $\hat{\alpha} \rightarrow \alpha$ at rate $T^{-1/2}$. Then

$$\begin{aligned} T\hat{\beta}_E &= (T^{-2}y'_{-1}My_{-1})^{-1}(T^{-1}y'_{-1}M\Delta y) = (T^{-2}z'_{-1}Mz_{-1})^{-1}(T^{-1}z'_{-1}M\Delta y) \\ &= T\beta + (T^{-2}z'_{-1}Mz_{-1})^{-1}T^{-1}z'_{-1}M_1\epsilon + o_p(1) \end{aligned} \quad (\text{A3})$$

since M is orthogonal to x_{-1} and Δx , and the limit distribution of $(T^{-2}z'_{-1}Mz_{-1})$ is equal to the limit distribution of $(T^{-2}z'_{-1}M_1z_{-1})$, following the same arguments as in the proof for Proposition 1.

Finally, using $T\beta \rightarrow c$ and substituting results (a) to (c) of (R2) into (A3) yields the required results. Since $\hat{\sigma}_\epsilon \rightarrow \sigma_\epsilon$ and $\hat{\sigma}_e \rightarrow \sigma_e$, the proof for the limit distribution of the t ratio follows along similar lines, leading to the required results.

PROOF OF PROPOSITION 3. For $k = 1$, from the limit distributions in Proposition 2 we have

$$T\hat{\beta}_E \Rightarrow c + \frac{\sigma_\epsilon}{\sigma_e} \left(\int \bar{K}_e^2 \right)^{-1} \int \bar{K}_e dB_\epsilon.$$

Since $\sigma_\epsilon/\sigma_e = (1 + q^2)^{-1/2}$, as $q \uparrow \infty$, we have that $\sigma_\epsilon/\sigma_e \uparrow 0$ and

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

Furthermore, since $\hat{\sigma}_\epsilon \rightarrow \sigma_\epsilon$ and $\hat{\sigma}_e \rightarrow \sigma_e$, the proof for the limit distribution of the t ratio proceeds along similar lines.

ACKNOWLEDGEMENTS

We are grateful to L. Álvarez, P. Boswijk, J. Gonzalo, N. Haldrup, D. Hendry, J. MacKinnon, P.C.B. Phillips and a referee for helpful comments on previous versions of the paper. The first author also wishes to thank the Department of Economics at Queen's University, Canada, for their hospitality while a first version of this paper was being written, and the UK Economic and Social Research Council for funding his research. The second author acknowledges support from Spanish Dirección General de Investigación Científica y Técnica (DGICYT), reference number SEC 96 0738.

REFERENCES

- BANERJEE, A., DOLADO, J., HENDRY, D. and SMITH, G. (1986) Exploring equilibrium relationships in econometrics through static models: some Monte-Carlo evidence. *Oxford Bull. Econ. Stat.* 48, 253–77.
- , —, GALBRAITH J. and HENDRY, D. (1993) *Cointegration, Error Correction and the Econometric Analysis of Non-stationary Data*. Oxford: Oxford University Press.
- BOSWIJK, H. (1991) Testing for cointegration in structural models. Report AE7/91, University of Amsterdam.
- COCHRANE, D. and ORCUTT, G. (1949) Application of least-squares regression to a relationship containing autocorrelated error terms. *J. Am. Stat. Assoc.* 44, 32–61.
- ENGLER, R. and GRANGER, C. (1987) Cointegration and error correction: representation, estimation and testing. *Econometrica* 55, 251–76.
- , HENDRY, D. and RICHARD, J. F. (1983) Exogeneity. *Econometrica* 51, 277–304.
- FULLER, W. (1976) *Introduction to Statistical Time Series*. New York: Wiley.
- GRANGER, C. and LEE, T. (1990) Multicointegration. In *Advances in Econometrics* (eds G. Rhodes and T. Fomby). Greenwich, CT: JAI Press.
- HANSEN, B. (1990) A powerful, simple test for cointegration using Cochrane–Orcutt. Working Paper 230, University of Rochester.
- (1995) Rethinking the univariate approach to unit root testing: using covariates to increase power. *Economet. Theory* 11, 1148–72.
- HENDRY, D. (1987) Econometric methodology: a personal perspective. In *Advances in Econometrics* (ed. T. Bewley). Cambridge: Cambridge University Press.
- (1995) On the interaction of unit roots and exogeneity. *Economet. Rev.* 14, 383–419.
- and ERICSSON, N. (1991) Modeling the demand for narrow money in the United Kingdom and the United States. *Eur. Econ. Rev.* 35, 833–81.
- and RICHARD, J. F. (1982) On the formulation of empirical models in dynamic econometrics. *J. Economet.* 20, 3–23.
- , PAGAN, A. and SARGAN, J. D. (1984) Dynamic specification. In *Handbook of Econometrics*, Vol. 1 (eds Z. Griliches and M. Intriligator). Amsterdam: North-Holland.
- HUNTER, J. (1990) Cointegrating exogeneity. *Econ. Lett.* 34, 33–35.
- KADANE, J. (1970) Testing overidentifying restrictions when the disturbances are small. *J. Am. Stat. Assoc.* 65, 182–85.
- (1971) Comparison of k -class estimators when the disturbances are small. *Econometrica* 39, 723–37.
- KIVIET, J. and PHILLIPS, G. (1992) Exact similar tests for unit roots and cointegration *Oxford Bull. Econ. Stat.* 54, 349–68.
- KREMERS, J., ERICSSON, N. and DOLADO, J. (1992) The power of cointegration tests. *Oxford Bull. Econ. Stat.* 325–48.
- OSTERWALD-LENUM M. (1992) A note with quantities of the asymptotic distribution of the maximum likelihood cointegration rank test statistics. *Oxford Bull. Econ. Stat.* 54, 461–71.
- PHILLIPS, P. (1987) Towards a unified asymptotic theory of autoregression. *Biometrika* 74, 535–47.
- (1988) Regression theory for near-integrated time series. *Econometrica* 56, 1021–43.
- and DURLAUF, S. (1986) Multiple time series with integrated variables. *Rev. Econ. Stud.* 53, 473–96.
- and HANSEN, B. (1990) Statistical inference in instrumental variables regressions with I(1) processes. *Rev. Econ. Stud.* 57, 99–125.
- and LORETAN, M. (1991) Estimating long-run economic equilibria. *Rev. Econ. Stud.* 58, 407–36.
- and OULIARIS, S. (1990) Asymptotic properties of residual based tests for cointegration. *Econometrica* 58, 165–93.
- SAID, S. and DICKEY (1984) Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika* 71, 599–607.
- SAIKKONEN, P. (1991) Asymptotically efficient estimation of cointegration regressions. *Economet. Theory* 7, 1–21.
- SCHWERT, G. (1989) Tests for unit roots: a Monte-Carlo investigation. *J. Bus. Econ. Stat.* 7, 147–59.
- STOCK, J. and WATSON, M. (1993) A simple estimator of cointegrating vectors in higher order integrated systems. *Econometrica* 61, 783–820.