BANCO DE ESPAÑA

ON SOME SIMPLE TEST FOR COINTEGRATION: THE COST OF SIMPLICITY

Anindya Banerjee, Juan J. Dolado and Ricardo Mestre

SERVICIO DE ESTUDIOS
Documento de Trabajo nº 9302

BANCO DE ESPAÑA

ON SOME SIMPLE TESTS FOR COINTEGRATION: THE COST OF SIMPLICITY

Anindya Banerjee, Juan J. Dolado and Ricardo Mestre (*)

(*) The affiliation of the first author is Wadham College and I.E.S., University of Oxford. The affiliation of the second and third authors is Rescarch Dpt., Bank of Spain. This paper was prepared in part while the first two authors were visiting the Institute of Economics at Aarhus University, who we thank for generous hospitality. We are grateful to luis Alvarez, Manuel Arellano, Jesús Gonzalo, Niels Haldrup, Svend Hylleberg and participants in a seminar at the University of Oxford for helpful discussions and to Carmina Arellano for her skill in typing the manuscript. The usual caveat applies.

SERVICIO DE ESTUDIOS
Documento de Trabajo n.° 9302

In publishing this series the Bank of Spain seeks to disseminate studies of interest that will help acquaint readers better with the Spanish economy.

The analyses, opinions and findings of these papers represent the views of their authors; they are not necessarily those of the Bank of Spain.

ISBN: 84-7793-203-4
Depósito legal: M-40255-1992
Imprenta del Banco de España

1. INTRODUCTION

In a recent paper, Hansen (1990) provides a powerful, simple test for cointegration which avoids a common problem affecting most available tests of cointegration in a multivariate context. This problem is known as the 'curse of dimensionality'. It is associated with the fact that the distributions of existing test statistics depend upon the number of variables in the system, so that there is a need to simulate new distributions and there might also be a decrease in power as the size of the system increases. The cause of this dependence is that these tests estimate <u>all</u> the unit roots under the null, a procedure which might be superfluous given that the alternative hypothesis most commonly entails a low dimensional cointegrating space. Hansen provides a GLS approach that avoids this problem, which reduces in a single equation setting to iterated Cochrane-Orcutt regression (henceforth denoted as the C-O test). He shows that the asymptotic distribution of the AR (1) coefficient is the same as the distribution of a univariate Dickey-Fuller test- statistic, either in its coefficient or t-ratio form, irrespectively of the number of regressors in the regression equation. Henceforth we denote this class of test as <u>dimension-invariant</u>, with the use of dimension referring to the number of regressors in the regression model.

However, the C-O test implicitly imposes <u>common-factor restrictions</u> on the model which may or may not be valid. In this paper we concentrate on the consequences of this imposition and question whether there is a test-statistic which does not impose such a restriction and yet is still dimension invariant.

The C-O test is most powerful when the common-factor restriction is valid and remains consistent under a properly defined local alternative hypothesis. If the restrictions are violated grossly, the C-O test loses power relative to other tests of cointegration where no a priori restrictions are imposed. Thus, more broadly, the paper investigates the trade-off, for power of tests, which exists between possibly estimating unneccessary parameters (if the common-factor restriction is valid) and estimating a model where an invalid common-factor restriction has been imposed (i.e. the restriction is invalid). In large enough sample sizes the trade-off should favour unrestricted estimation and we show this to be the case. Thus, in this sense there is a 'cost of simplicity' which the applied econometrician should take into account when making use of these tests.

Among the alternative tests considered which are not dimension invariant, we concentrate on a simple test based on the estimated error-correction coefficient in the error correction representation (henceforth denoted as ECM test), as proposed by Banerjee et al. (1986), which, however, does not impose the common-factor restriction. We therefore present critical values up to five regressors. We then proceed to compare the power of the ECM test with the C-O test. The gain in power, for violations of the common-factor restriction, in some cases, is quite impressive. We also show that there is a unique transformation of the ECM test which, in the spirit of the C-O approach, retains the property of being dimension-invariant. This test (henceforth denoted as EDF test: "ECM cum Dickey-Fuller"), however, is powerless against the alternative hypothesis of cointegration but there are some directions, related to local alternatives, in which it can be more powerful than the ECM test. Moreover, although it has the same asymptotic properties as the C-O test-statistic under the null hypothesis, it offers a slightly better approximation of the Dickey-Fuller distribution in finite samples.

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 C-O procedure may remain in more general cases.

The rest of the paper is organised as follows. Section 2 describes the data generation process (DGP) of interest and briefly describes Hansen's C-O procedure. Section 3 describes the ECM and EDF test statistics together with their asymptotic distributions under the null hypothesis of no cointegration, while Section 4 gives the corresponding asymptotic distributions under the alternative hypothesis of cointegration, using both a fixed alternative and a near non-cointegrated alternative. Section 5 provides some Monte-Carlo finite-sample evidence about the illustrative DGP. Section 6 considers generalisations. Finally, Section 7 concludes.

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; B_ϵ (similar for the disturbances u and e defined below) denotes a unit variance Brownian motion process associated to the standarised disturbance $(\epsilon/\sigma_\epsilon)$. Arguments of functionals on the space [0,1] are frequently suppressed and integrals on the space [0,1] such as $\int_0^1 B_\epsilon^2(r)\,dr$ are written as $\int B_\epsilon^2$ to reduce notation. Proofs are left to the appendix.

2. A SIMPLE DGP AND THE C-O TEST STATISTIC

By using a simple DGP, based upon a multivariate dynamic process, this section focuses on the merits and limitations of Hansen's C-O procedure.

The illustrative bivariate DGP has been used elsewhere for expository purposes; c.f. Davidson et al. (1978), Banerjee et al. (1986) and Kremers et al. (1992):

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

$$\Delta x_t = u_t \tag{2}$$

where

$$\begin{matrix} 1 \\ k \end{matrix} \begin{pmatrix} \epsilon_t \\ u_t \end{pmatrix} \sim NI \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_\epsilon^2 & 0 \\ 0 & \Sigma_u \end{pmatrix} \right)$$

and α, λ and x_t are (k x 1) vectors of parameters and explanatory variables.

We further assume that $-1 < \beta \leq 0$. Typically, in many empirical examples, $\alpha \approx 0.5$ and $\beta = -0.1$ with $\sigma_u > \sigma_\epsilon$, σ_u being a representative element of the diagonal of Σ_u ; that is, the short-run effect is smaller than the long-run effect, the adjustment process to the equilibrium relationship is slow, and the variability of the regressors tends to be larger than the variability of the error term. In this DGP, y_t and the vector of x_t's are cointegrated when $-1 < \beta < 0$, while they are not cointegrated when $\beta = 0$. Thus, tests of cointegration must rely upon some estimate of ß. Under the convenient assumption that the x_t's are strongly exogenous (see Hendry and Richard (1982)) for the parameters of the conditional model (1), non-linear least squares (NLS) can be applied to (1) as follows:

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

where either the coefficient or the t-ratio based upon ß can be used to test the null hypothesis that y_t and x_t are cointegrated with cointegrating vector $(1, -\lambda')'$. However, it will be shown that the asymptotic distributions of these test-statistics shift away from the origin as the dimensionality of the vector x_t increases. Thus larger test statistics are needed for rejection reflecting the fact that this class of test-statistics is not dimension invariant.

Hansen's procedure corrects this problem by performing a slightly different regression, based upon the assumption that $\alpha = \lambda$, i.e. using a common root formulation given by:

$$y_t = \lambda' x_t + v_t \qquad (3)$$

$$v_t = (1+\beta) v_{t-1} + e_t \qquad (4)$$

or

$$\Delta y_t = \lambda' \Delta x_t + \beta\,(y_{t-1} - \lambda' x_{t-1}) + e_t \qquad (5)$$

where

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

To estimate (3)-(4), iterated least squares can be used. Denote the one-step iterated estimators by $(\hat{\lambda}_H, \hat{\beta}_H)$. This is the procedure chosen by Hansen.

According to the C-O (1949) classic procedure, the iterated method works as follows. Apply OLS to the transformed equation:

$$\Delta y_t - \hat{\beta}_H y_{t-1} = \lambda_H' [\Delta x_t - \hat{\beta}_H x_{t-1}] + e_t \tag{6}$$

where $\hat{\beta}_H$ is obtained from applying OLS to the autoregression (4) using the OLS residuals, $\hat{v}_t$'s, generated from regression (3).

Under the null hypothesis of no cointegration, $H_o : \beta = 0$, Hansen's Theorem 1 shows that $\hat{\beta}_H \rightarrow 0$ and $\hat{\lambda}_H \rightarrow \alpha$ where:

$$\alpha \equiv [E(\Delta x_t \Delta x_t')]^{-1} E[\Delta x_t \Delta y_t]$$

Thus, under the null hypothesis, the C-O estimate of $\hat{\lambda}_H$ converges to a vector of constants corresponding to the short-run coefficients in equation (1) of the DGP when $\beta = 0$. From this interesting result, Hansen suggests a test for cointegration using residuals from the C-O estimates which is dimension invariant. Consider now the following regression:

$$\Delta y_t - \hat{\lambda}_H' \Delta x_t = \tilde{\beta}_H' [y_{t-1} - \hat{\lambda}_H' x_{t-1}] + \tilde{e}_t \tag{7}$$

Theorem 2 in Hansen (1990) shows that, since $\hat{\lambda}_H \rightarrow \alpha$, the regressand and regressor in (7) are asymptotically equivalent to ϵ_t and $S_{\epsilon t-1}$

$(=\sum_{1}^{t-1}\epsilon_1)$ and hence

$$T\,\tilde{\beta}_{B}' \rightarrow \frac{\int B_\epsilon\, dB_\epsilon}{\int B_\epsilon^2}$$

which is the distribution of the univariate Dickey-Fuller unit root test based upon the coefficient of the first order autorregresive model.

It can be noted that Hansen's procedure can also be applied directly in a single stage, estimating (5) by NLS. Denote these estimators by $(\hat{\lambda}_{NH}, \hat{\beta}_{NH})$. Since this estimation procedure is asymptotically equivalent to the iterated C-O method, again the following theory holds.

Proposition 1 Under the null hypothesis of no cointegration ($\beta = 0$),

$$T^{\frac{1}{2}}(\hat{\lambda}_{NH}-\alpha) \rightarrow N\,(0,\sigma_\epsilon^2\,\Sigma_u^{-1})$$

$$T\hat{\beta}_{NH} \rightarrow \int B_\epsilon\, dB_\epsilon / \int B_\epsilon^2$$

and

$$t_{NH} \rightarrow \int B_\epsilon\, d\,B_\epsilon \,/\, (\int B_\epsilon^2)^{1/2}$$

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

$$t_{NH} = (\hat{\sigma}_e^{-2}\, T^{-2}\, \hat{z}_{NH}'\ \hat{z}_{NH}\,)^{1/2}\, T\, \hat{\beta}_{NH}$$

with

$$\hat{z}_{NHt} = y_t - \hat{\lambda}_{NH}' \; x_t$$

That is, both the coefficient and the t-test once more have the Dickey-Fuller distributions. Hansen notes that multiple iterations of the C-O procedure, equivalent to the non-linear method, improve the approximation to the Dickey-Fuller distribution, and this is what it is achieved by applying NLS. Note also that these tests are <u>similar</u>, i.e. invariant to the parameters of the DGP (α, σ^2_ϵ, Σ_u) under the null hypothesis and, therefore, such tests have the same null rejection frequency independently of the values of the nuisance parameters.

Although the previous test statistics are dimension invariant, i.e. their distributions are independent of the dimension of the vector x_t, note that equation (5), as compared to equation (1), ignores part of the information contained in Δx_t. Equivalently, (5) imposes the restriction that $\alpha=\lambda$, i.e. a common factor restriction. The transformation of (1) to (5) provides several interesting insights. First, (1) and (5) are equivalent representations of the DGP, given the above relationship between the disturbances ϵ_t and e_t, 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,

$$[1-(1+\beta)L] \; y_t = [\alpha-(\alpha+\lambda\beta)L]' \; x_t + \epsilon_t \qquad (8)$$

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 x_t

and y_t. Since, as mentioned above, it is a common empirical finding that $\alpha \neq \lambda$, imposing $\alpha=\lambda$ is rather arbitrary and the more general case remains the case of interest in this paper.

3. THE ECM AND EDF TEST STATISTICS

Drawing upon the previous ideas, we suggest the following alternative testing procedures which do not impose the common factor restriction but remain similar. The first test-statistic is not dimension-invariant, whilst the second is a transformation of the first one which achieves this type of invariance.

3.1. The ECM Test Statistic

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 of $H_o: ß=0$. Alternatively, (1) can be estimated by OLS in a simpler way, rewriting it as an unrestricted dynamic model of the form

$$\Delta y_t = \alpha_E' \Delta x_t + ß_E (y_{t-1} - c' x_{t-1}) + \theta_E' x_{t-1} + \epsilon_t \qquad (1'')$$

where c is a kx1 unit vector

When $ß_E \neq 0$, the cointegrating vector can be obtained from solving (1") for $(1:c'-\theta_E'/ß_E)'=(1:\lambda')'$. The relevant cointegration test has been shown by Banerjee and Hendry (1992) to be either a coefficient test or a t-test of $H_o: ß_E=0$, which is asympotically equivalent to test $H_o: ß=0$ in (1'). We denote the OLS estimator of $ß_E$ in (1") as $\hat{ß}_E$, defined by:

$$\hat{\beta}_E = \left[(y-xc)'_{-1}\ M\ (y-xc)_{-1}\right]^{-1}(y-xc)_{-1}\ M\ \Delta y \tag{9}$$

where M is the projection matrix orthogonal to the subspace spanned by Δx and x_{-1} (T x k). Then the following proposition holds.

Proposition 2. Under the null hypothesis of no cointegration ($\beta = 0$),

$$T\ \hat{\beta}_E \Rightarrow [\int B_\epsilon^2 - \int B_\epsilon B_u'(\int B_u B_u')^{-1} \int B_u B_\epsilon]^{-1}[\int B_\epsilon\ dB_\epsilon - \int B_\epsilon B_u'(\int B_u B_u')^{-1}\int B_u dB_\epsilon]$$

and

$$t_E \Rightarrow [\int B_\epsilon^2 - \int B_\epsilon B_u'(\int B_u B_u')^{-1} \int B_u B_\epsilon]^{-1/2}[\int B_\epsilon\ dB_\epsilon - \int B_\epsilon B_u'(\int B_u B_u')^{-1}\int B_u dB_\epsilon]$$

where t_E is the t-ratio defined as

$$t_E = [\hat{\sigma}_\epsilon^{-2}\ T^{-2}(y-xc)'_{-1}\ M\ (y-xc)_{-1}\ ]^{1/2}\ T\ \hat{\beta}_E$$

Note that these distributions depend upon the limiting distribution of the x_t's, as reflected by the presence of B_u, and therefore the corresponding test-statistics are not dimension invariant.

3.2. The EDF Test Statistic

It is important to note from the previous asymptotic distribution that there is a transformation of $T\hat{\beta}_E$ which has the Dickey-Fuller distribution and therefore is also dimension invariant. This transformation is obtained by solving out for the ratio between the first elements in the numerator and denominator in the previous distribution and gives rise to a new test statistic, denoted as $T\hat{\beta}_{EDF}$, given by:

$$T\,\hat{\beta}_{EDF} = T\,\hat{\beta}_E\,(1-\hat{\psi}_1'\hat{\psi}_3) + T\,\hat{\psi}_1'\hat{\psi}_2 \qquad (10)$$

where

$$\hat{\psi}_1 = \left[\frac{\hat{z}_{E-1}'\,\hat{z}_{E-1}}{T^2}\right]^{-1} \frac{\hat{z}_{E-1}'\,x_{-1}}{T^2} \Rightarrow (\sigma_\epsilon^2 \int B_\epsilon^2)^{-1}\,(\Sigma_u^{\frac{1}{2}}\,\sigma_\epsilon \int B_u\,B_\epsilon)$$

$$T\,\hat{\psi}_2 = \left(\frac{x_{-1}'\,x_{-1}}{T^2}\right)^{-1} \frac{x_{-1}'\,\Delta\,\hat{z}_E}{T} \Rightarrow (\Sigma_u^{\frac{1}{2}} \int B_u\,B_u'\,\Sigma_u^{\frac{1}{2}})^{-1}\,(\Sigma_u^{\frac{1}{2}}\,\sigma_\epsilon \int B_u dB_\epsilon)$$

$$\hat{\psi}_3 = \left(\frac{x_{-1}'x_{-1}}{T^2}\right)^{-1} \frac{x_{-1}'\,\hat{z}_{E-1}}{T^2} \Rightarrow (\Sigma_u^{\frac{1}{2}} \int B_u B_u'\,\Sigma_u^{\frac{1}{2}})^{-1}\,(\Sigma_u^{\frac{1}{2}}\,\sigma_\epsilon \int B_u B_\epsilon)$$

with $\hat{z}_{Et}=y_t-\hat{\alpha}_E{}'x_t$

That is, $\hat{\psi}_1$ ($\hat{\psi}_3$) is the (kxl) vector of parameter estimates in the regression of x_{-1} on $\hat{z}_{E-1}$ ($\hat{z}_{E-1}$ on x_{-1}) and $\hat{\psi}_2$ is the corresponding vector in the regression of $\Delta\hat{z}_E$ on x_{-1}. Note, however, that there is a simple transformation of (10) which facilitates the computation of the EDF test. Since the projection matrix M eliminates both x_{-1} and Δx, $\hat{\beta}_E$ in (9) is also identical to

$$\hat{\beta}_E = (\hat{z}_{E-1}'\,M\,\hat{z}_{E-1})^{-1}\,\hat{z}_{E-1}'\,M\,\Delta\,\hat{z}_E \qquad (11)$$

which substituted in (10), together with the expression for $\hat{\psi}_i$(i=1, 2, 3) implies that

$$T\,\hat{\beta}_{EDF} = \left(\frac{\hat{z}'_{E-1}\,\hat{z}_{E-1}}{T^2}\right)^{-1} \frac{\hat{z}'_{E-1}\,\Delta\hat{z}}{T} \qquad (12)$$

Then, since $\hat{\alpha}_E$ is a consistent estimator of α, the following theory holds.

Proposition 3 Under the null hypothesis of no cointegration ($\beta = 0$)

$$T\,\hat{\beta}_{EDF} \rightarrow \int B_\epsilon d\, B_\epsilon / \int B_\epsilon^2$$

and

$$t_{EDF} \Rightarrow (\int B_\epsilon^2)^{-1/2} \int B_\epsilon\, d\, B_\epsilon$$

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

$$t_{EDF} = t_E\,(1-\hat{\psi}_1'\,\hat{\psi}_3)^{\frac{1}{2}} + \hat{\psi}_2'\,t(\hat{\psi}_1)$$

$$= (\hat{\sigma}_\epsilon^{-2}\,T^{-2}\,\hat{z}'_{E-1}\,\hat{z}_{E-1})^{1/2}\,T\,\hat{\beta}_{EDF}$$

where $t(\hat{\psi}_1)$ is the vector of t-ratios on the parameter estimates $\hat{\psi}_1$ multiplied by the standard deviation of the residuals in the regresion of x_{-1} on $\hat{z}_{E-1}$, which is a non-parametric estimate of the distribution

$$(\int B_\epsilon^2)^{-1/2} \int B_u\, B_\epsilon$$

Thus, the EDF test statistics have the same property as the C-O test statistics although they differ in that the common factor restriction is not imposed in the regression model on which the test is based.

All in all, we have seen that the test statistics based on Hansen's approach and the EDF procedure achieve the Dickey-Fuller distribution under the null hypothesis, while the ECM procedure depends on the number of regressors. This simple DGP also makes the point that the EDF test is the only test-statistic which does not impose the common-factor restriction and still remains dimension invariant. In the next section the relative power behaviour of the three test statistics under the alternative hypothesis is examined.

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

The alternative hypothesis is that of cointegration which is, for (1)-(2), $-1<\beta<0$. Because the error-correction term in (1) is stationary under this hypothesis, distributional results from conventional central limit theorems, instead of functional central limit theorems, do apply for fixed alternatives. Although under a suitable sequence of local alternatives, the non-conventional asymptotic theory developed by Phillips (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, let us discuss briefly the fixed alternative case to give the intuition behind the results obtained under near-no cointegration.

The basic result is that the ECM test tends to have larger power than the C-O test when $\alpha \neq \lambda$ and Σ_u is large relative to σ_ϵ. Moreover, the EDF test tends to have less power than the C-O test except in some cases under the previous circumstances. The intuition behind these results is as follows. The ECM regression conditions on Δx_t, x_{t-1} and y_{t-1}, whereas the C-O regression conditions on the three sets of variables subject to restrictions.

That loses potentially valuable information. Consider again the alternative representations of (1):

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

As an extreme example, let $\varepsilon_t \approx 0$ but $\alpha \neq \lambda$ and Σ_u 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

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

the estimates of λ and β in the Hansen's procedure will be much more imprecise, affecting the power of the test based upon $\hat{\beta}_{NH}$ in the second stage.

With respect to the EDF test, it is clear that it has no power against a fixed alternative of cointegration, since it tests for the stationarity of $(y_t - \hat{\alpha}_E' x_t)$ which is different from $(y_t - \lambda' x_t)$ given that $\hat{\alpha}_E$ tends to α and not to λ under the alternative. However, for a local alternative with moderate sample size, when the common factor restriction is grossly violated, it may have slightly better power properties than the C-O test, given that it does not impose such a restriction.

To formalise the previous intuition, we start with the case of local alternatives distribution theory discussed in Phillips (1988). 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 to obtain the distribution under fixed alternatives. A similar analysis has been applied by Kremers et al. (1992) for the case where the potential cointegrating vector is assumed to be known and, therefore, it needs not be estimated. For this purpose, we make use of the DGP (1)- (2), where now

$$\beta = 1-\exp(c/T) \simeq -c/T \qquad (13)$$

In (13), c is a fixed scalar. We call time series that are generated by (1)-(2), with ß as in (13), near-no cointegrated, 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_o:ß= 0 which applies when c=0. When c>0, (13) represents a local alternative to H_o, 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)-(13), is directly measurable in terms of the non-centrality parameter c.

To proceed in the analysis of local power, use is made of the following diffussion process

$$K(r) = \int_0^r \exp\left[c(r-s)\right] dB(s) \equiv B(r)+c\int_0^r \exp\left[c(r-s)\right] B ds \qquad (14)$$

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

Using (14) it is possible to show that,

<u>Proposition 4</u> Under the alternative hypothesis of near-no cointegration (c>0):

$$T\hat{\beta}_{NH} \Rightarrow \left[\sigma_e^2 \int K_e^2 - 2\sigma_e (\alpha-\lambda)' \Sigma_u^{1/2} \int B_u K_e + (\alpha-\lambda)' \Sigma_u^{1/2} \int B_u B_u' \Sigma_u^{1/2} (\alpha-\lambda)\right]^{-1}$$

$$\left[-c \left(\sigma_e^2 \int K_e^2 - \sigma_e (\alpha-\lambda)' \Sigma_u^{1/2} \int B_u K_e\right) + \sigma_e \sigma_\epsilon \int K_e dB_\epsilon - \sigma_\epsilon (\alpha-\lambda)' \Sigma_u^{1/2} \int B_u dB_\epsilon\right]$$

$$T\hat{\beta}_E \Rightarrow -c + \left[\sigma_e^2 \left(\int K_e^2 - \int K_e' B_u \Sigma_u^{1/2} (\Sigma_u^{1/2} \int B_u B_u' \Sigma_u^{1/2})^{-1} \Sigma_u^{1/2} \int B_u' K_e\right)\right]^{-1}$$

$$\sigma_e \sigma_\varepsilon \left[\int K_e \, d B_\varepsilon - \int K_e' B_u \Sigma_u^{1/2} (\Sigma_u^{1/2} \int B_u B_u' \Sigma_u^{1/2})^{-1} \Sigma_u^{1/2} \int B_u' \, d B_\varepsilon\right]$$

and

$$T \hat{\beta}_{EDF} \equiv T \hat{\beta}_{NH}$$

whereas in the case of the t-ratios,

$$t_{NH} \Rightarrow \left\{\sigma_\varepsilon^{-2} \left[\sigma_e^2 \int K_e^2 - 2\sigma_e (\alpha-\lambda)' \Sigma_u^{-1/2} \int B_u K_e + (\alpha-\lambda)' \Sigma_u^{1/2} \int B_u B_u' \Sigma_u^{1/2} (\alpha-\lambda)\right]\right\}^{1/2} T\hat{\beta}_{NH}$$

$$t_E \Rightarrow \left\{\sigma_\varepsilon^{-2} \sigma_e^2 \left[\int K_e^2 - \int K_e B_u' \Sigma_u^{1/2} (\Sigma_u^{1/2} \int B_u B_u' \Sigma_u^{1/2})^{-1} \Sigma_u^{1/2} \int B_u' K_e\right]\right\}^{1/2} T\hat{\beta}_E$$

and

$$t_{EDF} \equiv t_{NH}$$

Note that when c=0, the non-centrality parameters of the three statistics are zero, K≡B and the distributions under the null are recovered, i.e. power equals size. Note also that the C-O and the EDF test-statistics have the same asymptotic distribution under a local alternative. The reason is that in both cases $\hat{\lambda}_{NH}$ and $\hat{a}_E$ converge to the true vector of short-run coefficients α at rate $O_p(T^{-1/2})$ in (1). However, as it will be seen below, under a fixed alternative, as $c\uparrow\infty$ and $T\uparrow\infty$, $\hat{\lambda}_{NH}$ will converge to the cointegrating slope at rate $O_p(T^{-1})$ whereas α_E will not converge to λ, implying zero power for the EDF test statistic.

Although the comparison of the asymptotic distributions under the alternative local hypothesis is cumbersome, given the difficulty of the expressions, we expect, that the power of the C-O and EDF tests will be higher than the power of the ECM test under these 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:

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

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) variance of $(\alpha-\lambda)\Delta x_t$ relative to ϵ_t. This ratio will play a prominent role in the analysis.

<u>Proposition 5</u>. When k=1, under the alternative hypothesis of near-no cointegration (c > 0):

$$T\,\hat{\beta}_{NH} \Rightarrow \{-c(\int K_u - B_u)^{-2}\int(K_u^2 - K_u B_u) + O_p(q^{-1})\}$$

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

and $T\,\hat{\beta}_{EDF}$ tends to the same limit as $T\,\hat{\beta}_{NH}$,

whereas in the case of the t-ratios,

$$t_{NH} \rightarrow [\int (K_u-B_u)^2]^{-\frac{1}{2}} [-cq \int K_u(K_u-B_u) + \int (K_u-B_u)\, d B_\varepsilon] + O_p(q^{-1})$$

$$t_E \rightarrow [\int K_u^2-(\int K_uB_u)/(\int B_u^2)]^{-\frac{1}{2}}[-cq+\int K_u\, d B_\varepsilon-(\int K_uB_u)(\int B_u d B_\varepsilon)/\int B_u^2]+O_p(q^{-1})$$

and t_{EDF} tends to the same limit as t_{NH}

Various interesting properties arise from Proposition 5. 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-ratio test statistics.

a) Coefficient Test Statistics

First, as $q \uparrow \infty$, i.e. $\alpha \neq \lambda$ and the variance of Δx_t is large relative to ϵ_t, the non-centrality parameter in the C-O and EDF tests has a stochastic slope given by

$$\int K_u (K_u-B_u) \; / \int (K_u-B_u)^2 \qquad (15)$$

whereas the ECM test has a slope equal to unity, i.e. there is a degenerate distribution centred on (-c). When c is 'large', the variance of the denominator in (15) 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 ß and q are 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 observe very low power of the test based on the ECM coefficient, although higher than that pertaining to the other two tests as q gets larger. This problem does not arise with the t-ratio version of the ECM test, as will be seen below, so that, in this sense, the tests based upon the t-ratios are preferable to those based directly on the scaled coefficients.

b) <u>t-ratio Test Statistics</u>

The discussion of the different tests is similar to the preceding one, except that 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 cases of the C-O and EDF statistics, using similar arguments to those employed in the discussion of (15), the distribution is again centred around zero, irrespectively of the value of c, thus their power will be lower than that of the ECM test, tending towards zero even when the limiting distribution under the local alternative tends to be less skewed to the left than that under the null hypothesis.

For the case of the fixed alternative ($c\uparrow\infty$ and $T\uparrow\infty$), ($y-\lambda x$) is stationary and the non-centrality (NC) parameters of the t-ratio versions of the tests are as follows.

Proposition 6. Under a fixed alternative hypothesis, the non-centrality parameters of the t-tests are

$$NC_{NB} = (1 - (1 + \beta)^2)^{-1/2}\ T^{1/2}\ \beta$$

$$NC_E = (1 + q^2)^{1/2}\ NC_{NB}$$

$$NC_{EDF} = 0$$

As we mentioned above, the values of these non-centralities help explain the better performance of the ECM test relative to the C-O test as long as $q>0$. Similarly, the EDF test has zero power since it tests for cointegration using the wrong cointegrating vector, i.e. $(1:\alpha)$ instead $(1:\lambda)$.

The next section of the paper addresses how well the asymptotic theory fits the finite sample behaviour. In it a small Monte-Carlo experiment is performed.

5. FINITE SAMPLE EVIDENCE

To examine the size and power of the C-O, ECM and EDF 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. 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 pilot study we choose

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

$\alpha = (0.1, 0.9)$

$ß$ = (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, 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 $ß_H$ in the initial iteration. Following the notation in (6), denote by $\hat{ß}_H$ the autoregressive coefficient obtained from (4). Then define

$$\hat{ß}_H^+ = \hat{ß}_H + a/T$$

where a>0 is fixed constant which Hansen suggests to select equal to 10.

To obtain $\hat{ß}_{NH}$, we start by using a quasi-difference of the data as in (6), obtaining an estimator of $\hat{ß}_H$. Eight iterations were performed on this procedure and, at the final stage, a/T was substracted from $\hat{ß}_H^+$, 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 upon the autoregressive model (4), using the OLS residuals from (3). This test suffers both from the 'curse of dimensionality' and the 'common factor restriction' problem (see Phillips and Ouliaris (1990) and Kremers et al. (1992)), so that it is useful to see how it performs relatively to the other three tests discussed in this paper.

Under the null of no cointegration ($\beta = 0$) Figs. 1 and 2 show the densities of the various tests while 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 both the C-O and EDF tests should correspond under the null, the first row in Table 1 offers the cumulative distribution of the DF test as taken from Tables 8-5-1 and 8-5-2 in 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 Propositions 1 to 3. Given this degree of invariance, the reported figures correspond to the averages of the critical values across the chosen range of values for s. It can be observed that the empirical distributions of the C-O and EDF tests are quite 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, yet in both cases the deviation seems to be small. It is interesting, however, to notice that the size of the EDF test-statistic seems to be closer to the theoretical size, even when the C-O test has been adjusted by the ad-hoc correction factor (10/T) recommended by Hansen (1990). Obviously, the empirical distribution of the ECM differs from the unit root distribution, as expected from Propositions 4 and 6. 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. The results point out in the same direction as before. The t-tests seem to be more immune to the "curse of dimensionality" than the coefficient tests and the EDF test seems to be more inmune than the C-O test. Again, as expected, the ECM and EG tests suffer from a shift to the left in the 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 other two tests, reflecting the "curse of dimensionality". However, as q increases, either because α becomes different from λ or s rises, the power of the ECM test overcomes the power of the other two tests. Second, the EDF test seems to be moderately more powerful than the adjusted C-O test for large values of q. Third, 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 which illustrates the "cost of simplicity" 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, whereas the EDF test rejects 8% of the time and the adjusted 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 and EDF tests. 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.

6. GENERALIZATIONS

The common factor "problem" of the C-O statistic remains when (1) includes additional lags. Furthermore, the use of augmented versions of the DF test, such as the ADF statistic, or the non-parametric corrections suggested by Phillips (1987a), on the C-O residuals, as in (6), do not resolve the problem. Since the argument is similar to that given by Kremers et al. (1992) for the case in which the cointegrating vector is assumed to be known and hence not estimated, this section will briefly review their reasoning and refer the reader to their work for the specific details.

First, consider the natural generalization of (1) for which lagged as well as current values of Δy_t and Δx_t may appear. Letting $\gamma(L)$ and $\alpha(L)$ be suitable scalar and vector polynomials in the lag operator L, the generalization of (1) is

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

Substracting $\gamma(L)\ \lambda'\ \Delta\ x_t$ from both sides, and letting now $z_t = y_t - \lambda'\ x_t$, we obtain

$$\gamma(L)\ \Delta\ z_t = \text{ß}\ z_{t-1} + e_t$$

where

$$e_t = [\alpha(L)' - \gamma(L)\ \lambda']\ \Delta\ x_t + \varepsilon_t \qquad (16)$$

which is a simple generalization of (4).

The C-O statistic is based upon (4), having estimated λ by iterated or non-linear C-O procedures, and so it imposes the common factor restriction

$$\alpha\ (L) = \gamma\ (L)\ \lambda$$

which, if invalid, may imply a loss of information and so a loss of power relative to the test based upon the ECM test.

With respect to the implementation of the EDF in this more general case, it is technically very similar to the reparameterisation which gives rise to (12), in this case reformulated as

$$\Delta y_t - \hat{\alpha}_E(L)'\Delta x_t - \hat{\gamma}_E^*(L)'\Delta y_{t-1} = \hat{\text{ß}}_{EDF}\left[y_{t-1} - \hat{\alpha}_E(L)'x_{t-1} - \hat{\gamma}_E^*(L)y_{t-2}\right] + \epsilon_t \qquad (17)$$

where $\hat{\gamma}_E^*\ (L) = 1 - \hat{\gamma}_E\ (L)$

Thus, the EDF test is based upon the transformation z_{Et} which now comprises the more general short-run dynamic structure. It is tedious, but straightforward, to show that both the coefficient and the t-ratio version of the C-O and EDF tests have the DF distributions as their limiting distributions.

Second, it is important to note that similar considerations apply when the assumption that u_t is white noise in the DGP (1)-(2), is replaced by u_t being I(0), not necessarily serially uncorrelated but still long-run independent of ε_t. In that case x_t would only be weakly exogenous instead of being strongly exogenous, as assumed above for simplicity, but OLS on (1) would still be asymptotically efficient (see in Johansen (1992) and the subsequent discussion in Dolado (1992) and Hendry (1992)). Following this weaker assumption, B_u will still be an independent Brownian motion of B_ε, with long-run variance given by:

$$\omega_u^2 = \sigma_u^2 + 2\,\phi_u$$

where

$$\phi_u = T^{-1} \sum_{1}^{\infty} u_{t-1}\, u_t$$

in the case where k=1.

In this case, the limiting distributions contained in Proposition 4 remain similar, except that the 'signal-to-noise' ratio q has to be replaced by its natural generalization

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

given that the ignored information on Δx_t is now I(0) instead of merely white noise. Again, as in the serially uncorrelated case, as q_o gets 'large' the 'curse of dimensionality' affecting the ECM test tends to be overcome by the 'cost of simplicity', implying that the ECM test tends to have better power properties than the other two tests.

Third, system analysis of cointegration based upon the C-O methods faces similar problems (see Kremers et al.), impliying that, when the common factor restriction is violated, it is preferable to use the likelihood-based method approach developed by Johansen (1988, 1991).

Finally, if desired, the data may be demeaned, or demeaned and detrended, before applying the various tests for cointegration. It is not difficult to see that the appropriate distributions for the C-O and EDF tests are given by the unit root tests in Fuller (1976) for models with intercepts and trends. The critical values for the ECM test in its t-ratio version, which are only available for k=1 (see Banerjee et al. (1993)), are extended 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).

7. 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. It is well known that tests for cointegration have low power against reasonable alternatives in typical sample sizes, the reasons being that the distributional theory for those tests depends upon the dimensionality of the system (the number of variables), which is due to the fact that all unit roots in the system are estimated. This does not make effective use of the structure of the alternative hypothesis, since typically the alternative of interest is a low dimensional cointegrating space. Thus, those tests suffer from what is known as the "curse of dimensionality", which is expected to reduce power. To overcome this difficulty, Hansen (1990) has proposed and alternative procedure based upon the second-stage residuals of the cointegrating relationships estimated by Cochrane-Orcutt. Under the null hypothesis of no cointegration the limiting distribution of the autoregressive coefficient in the autoregression model for those residuals is identical to the Dickey-Fuller distribution for a univariate series. Thus, this test-statistic is dimension invariant. In this paper we extend the argument of Kremers et al. (1992) and claim that, in spite of the advantages of this test, it suffers from the problem of possibly imposing invalid common factor restrictions, a problem which may reduce its power relative to other tests which are not dimension invariant. We derive conditions under which this second problem, denoted as the "cost of simplicity", may then reduce the power of Hansen's test relative to other tests, such as the test based on the estimated coefficient of the error correction term in the ECM representation of the model. Moreover, as a byproduct of the analysis, we also show that, when the variability of the regressors is large relative to the variability of the error term in the potential cointegrating equation and the common factor restriction is invalid, the t-ratio version of the ECM has much better power properties than the coefficient version of that test. Finally, it is shown that the only available test which is dimension invariant but does not impose the common factor restriction, is powerless against the alternative hypothesis of cointegration.

APPENDIX

The analysis contained in this appendix draws on a number of results in Phillips (1987b, 1988) (see Banerjee et al. (1992) for an exposition).

Under the null hypothesis of no-cointegration, the DGP (H_o) is given by

$$\Delta y_t = \alpha' \Delta x_t + \epsilon_t \; ; \qquad \epsilon_t \sim \text{nid}(0, \sigma_\epsilon^2)$$

$$\Delta x_t = u_t \; ; \qquad u_t \sim \text{NI}\,(0, \Sigma_u)$$

and the following asymptotic results (R1) are used

$$T^{-2} S'_{\epsilon-1} S_{\epsilon-1} \Rightarrow \sigma_\epsilon^2 \int B_\epsilon^2 \; ; \qquad T^{-1} S_{u-1} \epsilon \Rightarrow \sigma_\epsilon \Sigma_u^{1/2} \int B_u \, d\, B_\epsilon$$

$$T^{-2} S_{u-1} S'_{u-1} \Rightarrow \Sigma_u \int B_u B_u' \; ; \qquad T^{-1} S'_{\epsilon-1} \epsilon \Rightarrow \sigma_\epsilon^2 \int B_\epsilon \, d\, B_\epsilon$$

$$T^{-2} S_{u-1} S_{\epsilon-1} \Rightarrow \sigma_\epsilon \Sigma_u^{1/2} \int B_u B_\epsilon \qquad .$$

where $S_\epsilon (S_u)$ represent the accumulated sum of ϵ (u) from 1 to T.

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

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

$$\Delta x_t = u_t$$

with $\beta = -c/T$; $z = y - \lambda' x$, and $e = (\alpha - \lambda)' u + \epsilon$

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

$$T^{-2} z'z \Rightarrow \sigma_e^2 \int K_e^2 \; ; \qquad T^{-2} x_{-1} z_{-1} \Rightarrow \sigma_e \Sigma_u^{1/2} \int B_u K_e$$

$$T^{-1} z_{-1}' \epsilon \Rightarrow \sigma_e \sigma_\epsilon \int K_e \, d B_\epsilon \; ; \qquad .$$

and

$$\sigma_e^2 \int K_e^2 = (\alpha-\lambda)' \Sigma_u^{1/2} \int B_u B_u' \Sigma_u^{1/2} (\alpha-\lambda) + 2(\alpha-\lambda)' \sigma_\epsilon \Sigma_u^{1/2} \int B_u B_\epsilon + \sigma_\epsilon^2 \int B_\epsilon^2$$

<u>Proof of Proposition 1</u>

The criterion function to minimise in the C-O procedure is given by

$$\min \sum_{2}^{T} \left[\Delta y_t - \lambda' \Delta x_t - ß \, (y_{t-1} - \lambda' x_{t-1})\right]^2$$

From the f.o.c. it follows that

$$\hat{\lambda}_{NH} = (\Delta x^{*\prime} \Delta x^*)^{-1} (\Delta x^{*\prime} \; \Delta y^*) \qquad \text{(A.1)}$$

and

$$\hat{ß}_{NH} = (\hat{z}_{NH\,-1}' \; \hat{z}_{NH\,-1})^{-1} \; \hat{z}_{NH\,-1}' \; (\Delta y - \hat{\lambda}_{NH}' \; \Delta x) \qquad \text{(A.2)}$$

where

$$\Delta x^* = \Delta x - \hat{ß}_{NH} \, x_{-1}$$

$$\Delta y^* = \Delta y - \hat{ß}_{NH} \, y_{-1}$$

and

$$\hat{z}_{NHt} = y_t - \hat{\lambda}_{NH}' \; x_t = y_t - \alpha' x_t - (\hat{\lambda}_{NH} - \alpha)' \; x_t$$

Because β=0 in the DGP(H_o), $\hat{\beta}_{NH} \to \beta$, thus asymptotically

$$\Delta x^* \to \Delta x = u \quad \text{and} \quad \Delta y^* \to \Delta y = \alpha' u + \epsilon . \quad \text{Hence}$$

$$T^{1/2} (\hat{\lambda}_{NH} - \alpha) \Rightarrow N (0 , \sigma_\epsilon^2 \Sigma_u^{-1})$$

and since $\hat{\lambda}_{NH} \to \alpha$, $\hat{z}_{NH} \to S_\epsilon$ and $\Delta y - \hat{\lambda}'_{NH} \Delta x \to \epsilon$

Therefore, using the results (R1) for DGP (H_o)

$$T \hat{\beta}_{NH} \Rightarrow (\int B_\epsilon^2)^{-1} \int B_\epsilon \, d B_\epsilon$$

Finally, given that $\hat{\sigma}_\epsilon^{\,2} \to \sigma_\epsilon^{\,2}$, the proof for the t-ratio follows trivially

<u>Proof of Proposition 2</u>

Given that $\hat{\beta}_E$ is given in (9) and that the projection matrix **M** eliminates Δx and x_{-1}, it follows that

$$T\hat{\beta}_E = \left[T^{-2} (S'_{\epsilon-1} S_{\epsilon-1} - (S'_{\epsilon-1} S_{u-1})(S_{u-1} S'_{u-1})^{-1} (S'_{u-1} S_{\epsilon-1}))\right]^{-1}$$
$$T^{-1} \left[S'_{\epsilon-1} \epsilon - (S'_{\epsilon-1} S_{u-1})(S_{u-1} S'_{u-1})^{-1} (S'_{u-1} \epsilon)\right] \qquad \text{(A3)}$$

Therefore, using the results (R1) for DGP(H_o), the distribution of $T\hat{\beta}_E$ follows. From that distribution, given that $\hat{\sigma}_\epsilon^{\,2} \to \sigma_\epsilon^{\,2}$, the distribution of the t-ratio easily follows.

Proof of Proposition 3

Solving out for $(T^{-2} S'_{\epsilon-1} S_{\epsilon-1})^{-1} T^{-1} S'_{\epsilon-1} \epsilon$ in (A.3), the limiting distribution of $T\, T\hat{\beta}_{EDF}$ is obtained. Given that $\hat{\sigma}_{\epsilon}^{2} \to \sigma_{\epsilon}^{2}$, the distribution of the t-ratio is easily obtained.

Proof of Proposition 4

Note that z_t in DGP(H_{1a}) can be written as

$$z = y_t - \lambda' x_t = \hat{z}_{NHt} + (\hat{\lambda}_{NH} - \lambda)' x_t = \hat{z}_{NHt} + (\alpha - \lambda)' x_t + op\ (1) \qquad (A.4)$$

given that $\hat{\lambda}_{NH} \to \alpha$ at rate $Op(T^{-1/2})$ as in Proposition (1).

Similarly

$$z = \hat{z}_E + (\hat{\alpha}_E - \lambda)' x = \hat{z}_E + (\alpha - \lambda)\ x + op\ (1) \qquad (A.5)$$

given that $\hat{\alpha}_E \to \alpha$ at rate $Op(T^{-1/2})$.

Then, the different estimators of ß are given by

$$T\ \hat{\beta}_{NH} = (T^{-2}\ \hat{z}'_{NH-1}\ \hat{z}_{NH-1})^{-1}\ T^{-1}\ \hat{z}'_{NH-1}\ (\Delta y - \hat{\lambda}'_{NH}\ \Delta x) \qquad (A.6)$$

$$T\ \hat{\beta}_E = (T^{-2}\ \hat{z}'_{E-1}\ M\ \ z_{E-1})^{-1}\ T^{-1}\ \hat{z}'_{E-1}\ M\ \ \Delta\ \hat{z}_E \qquad (A.7)$$

and

$$T\ \hat{\beta}_{EDF} = (T^{-2}\ \hat{z}'_{E-1}\ \hat{z}_{E-1})^{-1}\ T^{-1}\ \hat{z}'_{E-1}\ \Delta\ \hat{z}_E \qquad (A.8)$$

where

$$M = I - x_{-1}\ (x_{-1}\ x'_{-1})\ x'_{-1}$$

Then, substituting the results (R2) for DGP (H_{1a}) and (A.4) - (A.5) into (A.6) - (A.8) and taking into account that $ß=-c/T$, the limiting distributions are obtained. The proofs for the t-ratio follow similarly.

Proof of Proposition 5

Inmediate from Proposition 4 and the definition of $q=(\alpha-\lambda)\sigma_u/\sigma_\epsilon$, given the results (R2) for DGP (H_{1a}).

Proof of Proposition 6

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

$$\Delta y_t = \alpha \Delta x_t + ß (y_{t-1} - \lambda x_{t-1}) + \epsilon_t$$

$$\Delta x_t = u_t$$

Define $z_t=y_t-\lambda x_t$ which is governed by the following AR(1) process

$$z_t = \frac{(\alpha - \lambda) u_t + \epsilon_t}{1 - (1 + ß) L}$$

where L is the lag operator and $V(z)=\sigma_e^2[1-(1+ß)^2]^{-1}$

Then

$$T^{1/2} (\hat{ß}_{NH} - ß) = (T^{-1} z_{-1}' z_{-1})^{-1} T^{-1/2} z_{-1}' \Delta z + op(1) \Rightarrow N[0, 1 - (1+ß)^2]$$

Hence

$$t_{NH} = [1 - (1+ß)^2]^{-1/2} T \hat{ß}_{NH} + op(1) \Rightarrow N\left[[1 - (1+ß)^2]^{-1/2} T^{1/2} ß, 1\right]$$

Similarly

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

Thus

$$t_E \rightarrow N\left[[1 - (1+\beta)^2]^{-1/2}\ (1+q^2)^{1/2}\ T^{1/2}\ \beta\ ,\ 1\right]$$

Finally $T^{-1}\ \hat{z}'_{E-1}\ \Delta\ \hat{z}_E$ is Op(1) since $\hat{\lambda}_E$ is not a consistent estimator of α under H_a. Therefore the non-centrality parameter of t_{EDF} is zero.

REFERENCES

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, forthcoming

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.

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 Cointaining 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.

Dolado, J. (1992) "A Note on Weak Exogeneity in VAR Cointegrated Systems" Economics Letters, 38, 139-143.

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.

Hansen, B.E. (1990) "A Powerful, Simple Test for Cointegration using Cochrane-Orcutt" University of Rochester. Working Paper No. 230 (paper presented at the 6 th. World Meeting of the Econometric Society, Barcelona, 1990).

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

Hendry, D.F. (1992) 'On the Interactions of Unit Roots and Exogeneity' Nuffield College, Oxford (mimeo).

S. Johansen (1988) "Statistical Analysis of Cointegration Vectors" Journal of Economic Dynamics and Control, 12, 132-154.

S. Johansen (1991) "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models" Econometrica, 59, 1551-1580.

Johansen, S. (1992) "Cointegration in Partial Systems and the Efficiency of Single Equation Analysis" Journal of Econometrics 52, 389-402.

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.

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. Ouliaris (1990) "Asymptotic Properties of Residual Based Tests for Cointegration" Econometrica 58, 165-193.

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)
EDF	- 8.08 (-1.96)	- 5.76 (-1.62)	0.97 (0.92)	1.36 (1.33)
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); EDF: Corrected ECM test (with the Dickey Fuller distribution); ECM: ECM coefficient test; EG: Engle and Granger test (computed from the OLS residuals of the static regression).

Table 2

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

Test	Size 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)
(2)	- 9.76 (-2.15)	- 6.94 (-1.77)	0.98 (0.86)	1.34 (1.28)
(3)	- 9.50 (-2.19)	- 6.17 (-1.81)	0.99 (0.90)	1.36 (1.31)
(4)	- 9.73 (-2.28)	- 6.47 (-1.91)	0.98 (0.88)	1.37 (1.32)
(5)	-10.37 (-2.38)	- 7.22 (-1.98)	0.94 (0.88)	1.36 (1.34)
EDF (k=1)	- 7.92 (-1.94)	- 5.65 (-1.62)	0.97 (0.92)	1.36 (1.33)
(2)	- 8.12 (-1.97)	- 5.72 (-1.61)	0.97 (0.90)	1.34 (1.32)
(3)	- 8.19 (-1.95)	- 5.75 (-1.62)	0.97 (0.90)	1.35 (1.32)
(4)	- 8.11 (-1.94)	- 5.80 (-1.61)	0.94 (0.85)	0.91 (1.23)
(5)	- 8.53 (-2.02)	- 5.98 (-1.66)	0.96 (0.90)	0.96 (1.34)
ECM (k=1)	-12.38 (-2.60)	- 9.66 (-2.27)	0.66 (0.44)	1.21 (0.88)
(2)	-16.38 (-3.03)	-13.14 (-2.68)	0.09 (0.04)	0.93 (0.51)
(3)	-19.72 (-3.36)	-16.24 (-3.01)	-0.80 (-0.31)	0.33 (0.14)
(4)	-22.99 (-3.63)	-19.28 (-3.26)	-1.95 (-0.59)	-0.50 (-0.17)
(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)
(2)	-21.07 (-3.37)	-17.74 (-3.04)	-2.83 (-1.03)	-1.58 (-0.68)
(3)	-26.31 (-3.80)	-22.72 (-3.49)	-5.29 (-1.49)	-3.76 (-1.21)
(4)	-31.04 (-4.19)	-27.10 (-3.87)	-7.86 (-1.87)	-6.25 (-1.59)
(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 exogenous regressors.

Table 3

Size Adjusted Powers of 5% Tests
(percentages)

	β=-0.05								
Test		s=0.05		s=1.00		s=5.00		s=20.00	
	C-O	30	(30)	8	(7)	0	(0)	0	(0)
α=0.1	EDF	30	(31)	8	(8)	0	(0)	0	(0)
	ECM	22	(18)	14	(23)	0	(88)	0	(100)
	EG	14	(15)	11	(11)	5	(5)	4	(4)
	C-O	30	(30)	28	(28)	16	(16)	1	(1)
α=0.9	EDF	30	(30)	29	(29)	16	(17)	2	(2)
	ECM	21	(17)	21	(17)	18	(19)	5	(48)
	EG	14	(14)	13	(14)	12	(12)	7	(7)
	β=-0.10								
	C-O	69	(68)	8	(8)	0	(0)	0	(0)
α=0.1	EDF	68	(68)	9	(9)	0	(0)	0	(0)
	ECM	53	(54)	44	(67)	8	(100)	0	(100)
	EG	36	(36)	30	(30)	18	(18)	17	(17)
	C-O	70	(70)	67	(67)	27	(26)	1	(1)
α=0.9	EDF	70	(70)	65	(65)	26	(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 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
	25	-4.53	-3.64	-3.24	-2.60
(2)	50	-4.29	-3.57	-3.20	-2.63
	100	-4.22	-3.56	-3.22	-2.67
	500	-4.11	-3.50	-3.19	-2.66
	25	-4.92	-3.91	-3.46	-2.76
(3)	50	-4.59	-3.82	-3.45	-2.84
	100	-4.49	-3.82	-3.47	-2.90
	500	-4.37	-3.77	-3.45	-2.90
	25	-5.27	-4.18	-3.68	-2.90
(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
	25	-5.53	-4.36	-3.82	-2.99
(5)	50	-5.04	-4.23	-3.82	-3.18
	100	-4.92	-4.20	-3.85	-3.28
	500	-4.81	-4.19	-3.86	-3.32
		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
	25	-5.12	-4.18	-3.72	-3.04
(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
	25	-5.42	-4.39	-3.89	-3.16
(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
	25	-5.79	-4.56	-4.04	-3.26
(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
	25	-6.18	-4.76	-4.16	-3.31
(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

Figure 1

t–Test Histograms (2 variables)

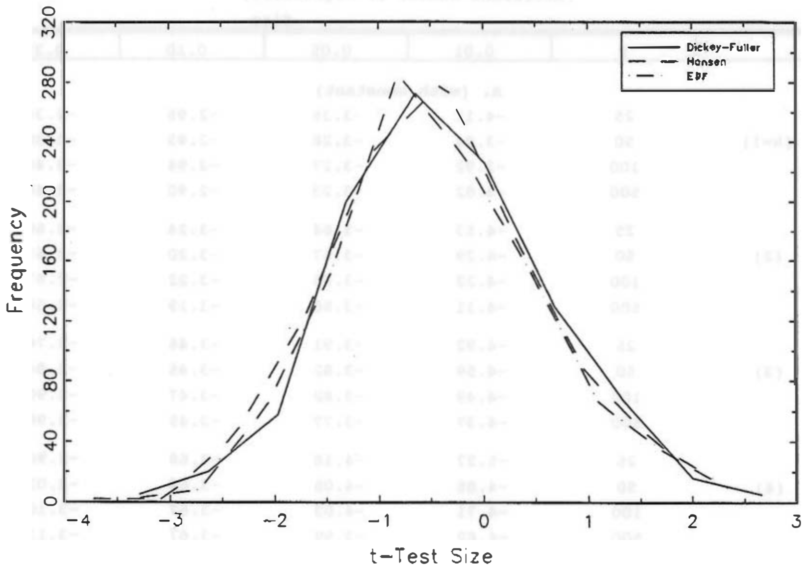
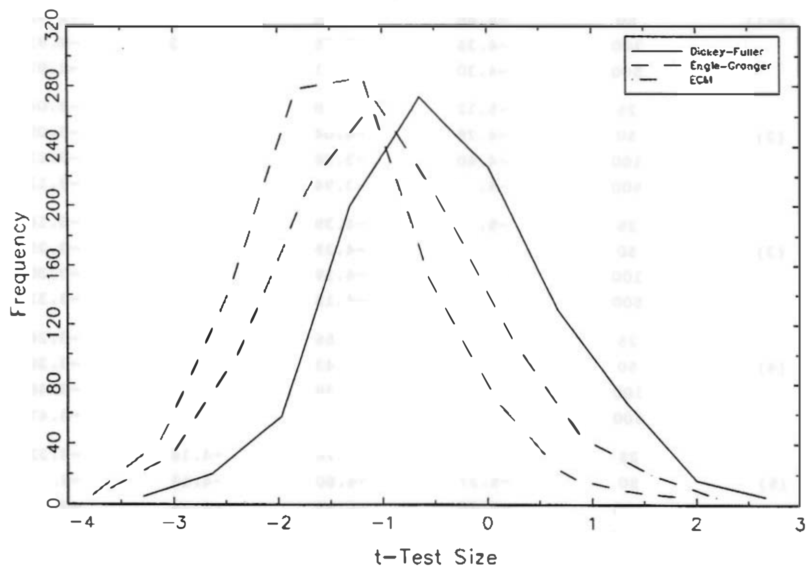


Figure 2

t–Test Histograms (2 variables)



DOCUMENTOS DE TRABAJO (1)

9101 **Javier Vallés:** Estimation of a growth model with adjustment costs in presence of unobservable shocks.

9102 **Javier Vallés:** Aggregate investment in a growth model with adjustment costs.

9103 **Juan J. Dolado:** Asymptotic distribution theory for econometric estimation with integrated processes: a guide.

9104 **José Luis Escrivá y José Luis Malo de Molina:** La instrumentación de la política monetaria española en el marco de la integración europea. (Publicada una edición en inglés con el mismo número.)

9105 **Isabel Argimón y Jesús Briones:** Un modelo de simulación de la carga de la deuda del Estado.

9106 **Juan Ayuso:** Los efectos de la entrada de la peseta en el SME sobre la volatilidad de las variables financieras españolas. (Publicada una edición en inglés con el mismo número.)

9107 **Juan J. Dolado y José Luis Escrivá:** La demanda de dinero en España: definiciones amplias de liquidez. (Publicada una edición en inglés con el mismo número.)

9108 **Fernando C. Ballabriga:** Instrumentación de la metodología VAR.

9109 **Soledad Núñez:** Los mercados derivados de la deuda pública en España: marco institucional y funcionamiento.

9110 **Isabel Argimón y José Mª Roldán:** Ahorro, inversión y movilidad internacional del capital en los países de la CE. (Publicada una edición en inglés con el mismo número.)

9111 **José Luis Escrivá y Román Santos:** Un estudio del cambio de régimen en la variable instrumental del control monetario en España. (Publicada una edición en inglés con el mismo número.)

9112 **Carlos Chuliá:** El crédito interempresarial. Una manifestación de la desintermediación financiera.

9113 **Ignacio Hernando y Javier Vallés:** Inversión y restricciones financieras: evidencia en las empresas manufactureras españolas.

9114 **Miguel Sebastián:** Un análisis estructural de las exportaciones e importaciones españolas: evaluación del período 1989-91 y perspectivas a medio plazo.

9115 **Pedro Martínez Méndez:** Intereses y resultados en pesetas constantes.

9116 **Ana R. de Lamo y Juan J. Dolado:** Un modelo del mercado de trabajo y la restricción de oferta en la economía española.

9117 **Juan Luis Vega:** Tests de raíces unitarias: aplicación a series de la economía española y al análisis de la velocidad de circulación del dinero (1964-1990).

9118 **Javier Jareño y Juan Carlos Delrieu:** La circulación fiduciaria en España: distorsiones en su evolución.

9119 **Juan Ayuso Huertas:** Intervenciones esterilizadas en el mercado de la peseta: 1978-1991.

9120 **Juan Ayuso, Juan J. Dolado y Simón Sosvilla-Rivero:** Eficiencia en el mercado a plazo de la peseta.

9121 **José M. González-Páramo, José M. Roldán y Miguel Sebastián:** Issues on Fiscal Policy in Spain.

9201 **Pedro Martínez Méndez:** Tipos de interés, impuestos e inflación.

9202 **Víctor García-Vaquero:** Los fondos de inversión en España.

9203 **César Alonso y Samuel Bentolila:** La relación entre la inversión y la «Q de Tobin» en las empresas industriales españolas. (Publicada una edición en inglés con el mismo número.)

9204 **Cristina Mazón:** Márgenes de beneficio, eficiencia y poder de mercado en las empresas españolas.

9205 **Cristina Mazón:** El margen precio-coste marginal en la encuesta industrial: 1978-1988.

9206 **Fernando Restoy:** Intertemporal substitution, risk aversion and short term interest rates.

9207 **Fernando Restoy:** Optimal portfolio policies under time-dependent returns.

9208 **Fernando Restoy and Georg Michael Rockinger:** Investment incentives in endogenously growing economies.

9209 **José M. González-Páramo, José M. Roldán y Miguel Sebastián:** Cuestiones sobre política fiscal en España.

9210 **Angel Serrat Tubert:** Riesgo, especulación y cobertura en un mercado de futuros dinámico.

9211 **Soledad Núñez Ramos:** Fras, futuros y opciones sobre el MIBOR.
9212 **Federico J. Sáez:** El funcionamiento del mercado de deuda pública anotada en España.
9213 **Javier Santillán:** La idoneidad y asignación del ahorro mundial.
9214 **María de los Llanos Matea:** Contrastes de raíces unitarias para series mensuales. Una aplicación al IPC.
9215 **Isabel Argimón, José Manuel González-Páramo y José María Roldán:** Ahorro, riqueza y tipos de interés en España.
9216 **Javier Azcárate Aguilar-Amat:** La supervisión de los conglomerados financieros.
9217 **Olympia Bover:** Un modelo empírico de la evolución de los precios de la vivienda en España (1976-1991). (Publicada una edición en inglés con el mismo número.)
9218 **Jeroen J. M. Kremers, Neil R. Ericsson and Juan J. Dolado:** The power of cointegration tests.
9219 **Luis Julián Alvarez, Juan Carlos Delrieu y Javier Jareño:** Tratamiento de predicciones conflictivas: empleo eficiente de información extramuestral. (Publicada una edición en inglés con el mismo número.)
9220 **Fernando C. Ballabriga y Miguel Sebastián:** Déficit público y tipos de interés en la economía española: ¿existe evidencia de causalidad?
9221 **Fernando Restoy:** Tipos de interés y disciplina fiscal en uniones monetarias. (Publicada una edición en inglés con el mismo número.)
9222 **Manuel Arellano:** Introducción al análisis econométrico con datos de panel.
9223 **Angel Serrat:** Diferenciales de tipos de interés onshore/offshore y operaciones swap.
9224 **Angel Serrat:** Credibilidad y arbitraje de la peseta en el SME.
9225 **Juan Ayuso y Fernando Restoy:** Eficiencia y primas de riesgo en los mercados de cambio.
9226 **Luis J. Alvarez, Juan C. Delrieu y Antoni Espasa:** Aproximación lineal por tramos a comportamientos no lineales: estimación de señales de nivel y crecimiento.
9227 **Ignacio Hernando y Javier Vallés:** Productividad, estructura de mercado y situación financiera.
9228 **Angel Estrada García:** Una función de consumo de bienes duraderos.
9229 **Juan J. Dolado and Samuel Bentolila:** Who are the insiders? Wage setting in spanish manufacturing firms.
9301 **Emiliano González Mota:** Políticas de estabilización y límites a la autonomía fiscal en un área monetaria y económica común.
9302 **Anindya Banerjee, Juan J. Dolado and Ricardo Mestre:** On some simple tests for cointegration: the cost of simplicity.

(1) Los Documentos de Trabajo anteriores a 1991 figuran en el catálogo de publicaciones del Banco de España.

Información: Banco de España
Sección de Publicaciones. Negociado de Distribución y Gestión
Teléfono: 338 51 80
Alcalá, 50. 28014 Madrid