Petr Mariel

A comparison of cointegration tests

Applications of Mathematics, Vol. 41 (1996), No. 6, 411–431

Persistent URL: http://dml.cz/dmlcz/134335

A COMPARISON OF COINTEGRATION TESTS

PETR MARIEL, Bilbao

(Received June 1, 1995)

Summary. In this paper some of the cointegration tests applied to a single equation are compared. Many of the existent cointegration tests are simply extensions of the unit root tests applied to the residuals of the cointegrating regression and the habitual H_0 is no cointegration. However, some non residual-based tests and some tests of the opposite null hypothesis have recently appeared in literature. Monte Carlo simulations have been used for the power comparison of the nine selected tests (ADF, $\hat{Z}_\alpha$, $\hat{Z}_t$, DHS, $J1$, $H1$, $H2$, C, LBI) using several types of data generating processes.

Keywords: Integrated Processes, Monte Carlo Simulation

AMS classification: 62J05, 62E25

1. Introduction

Two or more time series are said to be cointegrated if each of the series taken individually is $I(1)$ while some linear combination of the series is $I(0)$. Cointegration means that many shocks can cause permanent changes in the $I(1)$ series but there is some long-run equilibrium relation tying the series together. The series can then deviate from the equilibrium in the short run but not in the long run.

Cointegration analysis is often used in financial economics. Kasa [9], for example, looks for common stochastic trends among the stock markets of the U.S.A., Japan, England, Germany and Canada. His results indicate the presence of a single common trend driving these countries' stock markets. In a similar study, Arshanapalli and Doukas [1] discover that the U.S.A. stock market has a considerable impact on the French, German and U.K. markets in the post-October 1987 period. Johansen and Juselius [8] look for relations among prices, interest rates and exchange rates for the U.K., and test some hypotheses on cointegration relations among these variables. Haldrup [6] studies single-equation regression models containing both $I(1)$ and $I(2)$

variables. His paper is completed with an empirical application of money demand in the U.K.. Ngama and Sosvilla [12] examine the purchasing power parity theory for both the Spanish Peseta/U.S. Dollar and the Spanish Peseta/Deutsche Mark exchange rates, using both the consumer and the wholesale price indices.

Given that testing the existence of a cointegration relationship in a single equation is equivalent to testing for the unit root in the residuals of the cointegrating regression, many cointegration tests are usually different ways of testing for this unit root, that is to say, they are extensions of unit root tests. This is the first reason for the existence of more tests of the null hypothesis of no cointegration than of the null of cointegration, since the null hypothesis of the majority of the unit root tests is the existence of the unit root at the frequency zero. Another reason is related to the asymptotic theory of the models with $I(1)$ variables. As Phillips and Ouliaris [15] argue, the conventional asymptotic theory fails under the H_0 of cointegration, that is to say, it is not valid for certain statistics (such as the long run variance).

Among the residual based tests of the H_0 of no cointegration the first to be mentioned is the well known Augmented Dickey-Fuller test (see [3]). This test has been and still is used in many applications for its simple and intuitive procedure. Other tests belonging to this class are the very well known $\hat{Z}_\alpha$ and $\hat{Z}_t$ [15] and the less known Durbin-Hausman test of Choi [2]. Among the non residual based tests of the H_0 of no cointegration, the test t_{ECM} of Kremers, Ericsson and Dolado [10] could be mentioned. It is simply the t statistic of the correction error term in the ECM model. Another statistic is the $J2$, proposed by Park [13], which is based on the signification of some additional regressors in the cointegrating regression.

The class of tests of the H_0 of cointegration has not been studied as much as the previous one and up to now there have been relatively few proposals of this type of tests. Among the oldest, the $J1$ statistic of Park [13] stands out. Hansen [7] suggests to test the cointegration using some of the statistics proposed for parameter instability against several alternatives in the context of cointegrated regression models. Other non residual based tests of the H_0 of cointegration are the tests $H1$ and $H2$ proposed by Fernández-Macho ([4], [5]) which are based on the Durbin-Hausman principle. From the residual based tests of the H_0 of cointegration, the tests C of Shin [17] and LBI of Leybourne and McCabe [11] should be mentioned.

2. Cointegration tests

2.1. Preliminaries

Let an m-vector time series $\{\mathbf{z}_t\}_0^\infty$ be generated by $\mathbf{z}_t = \mathbf{z}_{t-1} + \mathbf{w}_t$. We assume that $\{\mathbf{w}_t\}$ is a stationary and ergodic random sequence with zero mean and finite variance. The initial variable $\mathbf{z}_0$ is assumed to be any random variable. We also assume that the partial sum process of $\{\mathbf{w}_t\}$ satisfies the multivariate invariance principle (see [14]). That is, for $r \in [0,1]$,

$$X_T(r) = T^{-1/2} \sum_{1}^{[Tr]} \mathbf{w}_t \Rightarrow B(r),$$

where $B(r)$ is the m-vector Brownian motion with a covariance matrix

$$\Omega = \begin{bmatrix} \underset{(1 \times 1)}{\omega_{11}} & \underset{(1 \times (m-1))}{\omega_{21}'} \\ \underset{((m-1) \times 1)}{\omega_{21}} & \underset{((m-1) \times (m-1))}{\mathbf{\Omega}_{22}} \end{bmatrix}.$$

We consider the linear cointegrating regression, partitioning $\mathbf{z}_t = (y_t, \mathbf{x}_t')'$:

$$y_t = \hat{\gamma}'\mathbf{x}_t + \hat{u}_t. \tag{1}$$

The majority of the cointegration tests are based on this regression for which we assume that $\{\mathbf{x}_t\}$ is not cointegrated. Obviously $\{\hat{u}_t\}$ is $I(0)$ when $\{y_t\}$ and $\{\mathbf{x}_t\}$ are cointegrated and is $I(1)$ when they are not. The residual based tests are simply different ways of testing the cointegrating regression residuals for a unit root.

2.2. H_0: No cointegration

2.2.1. Augmented Dickey-Fuller test (ADF)

This is a well known unit root test applied to the residuals of the cointegrating regression $\{\hat{u}_t\}$. The ADF statistics is the t-ratio of the coefficient α_* in the regression:

$$\Delta\hat{u}_t = \hat{\alpha_*}\hat{u}_{t-1} + \sum_{i=1}^{p} \hat{\varphi}_i \Delta\hat{u}_{t-1} + \hat{v}_t^* \tag{2}$$

$$ADF = t_{\hat{\alpha_*}}.$$

Under H_0: no cointegration

The asymptotic distribution of the ADF statistic is, as in the case of the following tests, a stochastic integral of the continuous stochastic processes which are continuous functionals of the standard m-dimensional Brownian motion:

$$t_{\hat{\alpha_*}} \Rightarrow \int_0^1 R\,\mathrm{d}S,$$

provided the order of the autoregression in the ADF is such that $p \to \infty$ as $T \to \infty$, $p = o\ (T^{1/3})$ and

$$R(r) = Q(r)\Big/\left(\int_0^1 Q^2\right)^{1/2},$$

$$S(r) = Q(r)/(\kappa'\kappa)^{1/2},$$

$$Q(r) = W_1(r) - \int_0^1 W_1 W_2'\left(\int_0^1 W_2 W_2'\right)^{-1} W_2(r),$$

$$\kappa' = \left(1, -\int_0^1 W_1 W_2'\left(\int_0^1 W_2 W_2'\right)^{-1}\right),$$

$$W(r) = \begin{bmatrix} \underset{(1)}{W_1} \\ \underset{(m-1)}{W_2} \end{bmatrix}.$$

$W(r)$ denotes the m-vector standard Brownian motion which is written as W to achieve notational economy as $\int_0^1 W_1 W_2'$ is used for $\int_0^1 W_1(r) W_2'(r)\, dr$. The asymptotic distribution depends only on the number of regressors in (1).

Under H_a: cointegration

$$ADF = O_p\ (T^{1/2}). \tag{3}$$

The critical values and more details about the asymptotic distribution of the ADF statistic can be found in [15].

2.2.2. The Phillips' $\hat{Z}_\alpha$ and $\hat{Z}_t$ test

After the estimation of the regression $\hat{u}_t = \hat{\alpha}\hat{u}_{t-1} + \hat{v}_t$, the $\hat{Z}_\alpha$ and $\hat{Z}_t$ statistics are defined as

$$\hat{Z}_\alpha = T(\hat{\alpha} - 1) - (1/2)(s_{T\ell}^2 - s_v^2)\left(T^{-2}\sum_2^T \hat{u}_{t-1}^2\right)^{-1},$$

$$\hat{Z}_t = \left(\sum_2^T \hat{u}_{t-1}^2\right)^{1/2} (\hat{\alpha} - 1)/s_{T\ell} - (1/2)(s_{T\ell}^2 - s_v^2)\left[s_{T\ell}\left(T^{-2}\sum_2^T \hat{u}_{t-1}^2\right)^{1/2}\right]^{-1},$$

where

$$s_v^2 = T^{-1}\sum_1^T \hat{v}_t^2,$$

$$s_{T\ell}^2 = T^{-1}\sum_1^T \hat{v}_t^2 + 2T^{-1}\sum_{s=1}^{\ell} w_{s\ell} \sum_{t=s+1}^{T} \hat{v}_t \hat{v}_{t-s}$$

for a certain window such as $w_{s\ell} = 1 - s/(\ell + 1)$.

The first terms of the statistics $\hat{Z}_\alpha$ and $\hat{Z}_t$ are based on the conventional statistics for the H_0: $\alpha = 1$ in the relation $\hat{u}_t = \hat{\alpha}\hat{u}_{t-1} + \hat{v}_t$. The difference between them is that $\hat{Z}_\alpha$ is based only on the estimation of α but $\hat{Z}_t$ includes the standard error of regression. The second terms correct for the possible heteroskedasticity and autocorrelation of v_t. This correction is small when the estimations of the long run and short run variances do not differ too much; then the value of $(s^2_{T\ell} - s^2_v)$ is small while it would be large in the opposite case.

Under H_0: no cointegration

$$\hat{Z}_\alpha \Rightarrow \int_0^1 R\,\mathrm{d}R$$
$$\hat{Z}_t \Rightarrow \int_0^1 R\,\mathrm{d}S,$$

where $R(r)$ and $S(r)$ have been defined in the previous section.

These distributions depend only on the number of regressors in (1). The statistics $\hat{Z}_t$ and ADF have the same asymptotic distribution (this result depends crucially on the condition $p \to \infty$ in the regression (2)).

Under H_a: cointegration

$$\hat{Z}_\alpha = O_p(T) \tag{4}$$
$$\hat{Z}_t = O_p(T^{1/2}). \tag{5}$$

The statistic $\hat{Z}_\alpha$ diverges faster as $T \to \infty$ under the alternative than the statistics $\hat{Z}_t$ and ADF. The results (5) and (3) are valid only if $f_{qq}(0) > 0$, where $f_{qq}(\lambda)$ is the spectral density of the stationary process $q_t = h'\mathbf{z}_t$ ($h'h = 1$, that is to say, h is the cointegration vector). If $f_{qq}(0) = 0$ then $\hat{Z}_t = O_p(T)$. The critical values of the statistics $\hat{Z}_\alpha$ and $\hat{Z}_t$ can be found in [15].

Phillips and Ouliaris [15] recommend the statistic $\hat{Z}_\alpha$, because this test is likely to have higher power than $\hat{Z}_t$ and ADF in samples of moderate size. See [15] for further details.

2.2.3. Durbin-Hausman test

The null hypothesis (no cointegration) for this test is $H_0 : \alpha = 1$ in the regression $\hat{u}_t = \hat{\alpha}\hat{u}_{t-1} + \hat{v}_t$. The idea of this test is based on the Durbin-Hausman statistic composed of two estimators which are both consistent under the null but have different probability limits under the alternative hypothesis.

Two estimators which fulfil these conditions are the pseudo IV estimator

$$\hat{\alpha}_{IV} = \sum_{t=1}^{T} \hat{u}_t^2 \Big/ \sum_{t=1}^{T} \hat{u}_t \hat{u}_{t-1}$$

and the OLS estimator

$$\hat{\alpha}_{OLS} = \sum_{t=1}^{T} \hat{u}_t \hat{u}_{t-1} \Big/ \sum_{t=1}^{T} \hat{u}_{t-1}^2 .$$

The Durbin-Hausman statistic proposed by Choi [2] is

$$DHS = (\hat{\alpha}_{IV} - \hat{\alpha}_{OLS})^2 \Big/ \left\{ \hat{\delta}\hat{s}^2 \left(\sum_{t=1}^{T} \hat{u}_{t-1}^2 \right)^{-1} \right\},$$

where

$$\begin{aligned} \hat{s}^2 &= T^{-1} \sum_{t=1}^{T} (\hat{u}_t - \hat{\alpha}_{OLS} \hat{u}_{t-1})^2, \\ \hat{\delta} &= \hat{s}^2 / \hat{\sigma}^2, \\ \hat{\sigma}^2 &= T^{-1} \sum_{t=1}^{T} (\hat{u}_t - \hat{\alpha}_{OLS} \hat{u}_{t-1})^2 \\ &\quad + 2T^{-1} \sum_{s=1}^{\ell} w_{s\ell} \sum_{t=s+1}^{T} (\hat{u}_t - \hat{\alpha}_{OLS} \hat{u}_{t-1})(\hat{u}_{t-s} - \hat{\alpha}_{OLS} \hat{u}_{t-s-1}). \end{aligned}$$

$w_{s\ell}$ is a selected window as in the case of the previous statistics $\hat{Z}_\alpha$ and $\hat{Z}_t$. The term $\hat{\delta}$ appears in the statistic to correct possible serial correlation and guarantees that the asymptotic distribution is free of nuisance parameters.

Under H_0: no cointegration

$$DHS \Rightarrow (1 + f_{21}' F_{22}^{-2} f_{21}) \Big/ \int_0^1 Q^2,$$

where $Q(r)$ is defined in the previous section and

$$\int_0^1 WW' = \begin{bmatrix} f_{11} & f_{21}' \\ f_{21} & F_{22} \end{bmatrix}.$$

Under H_a: cointegration

If we assume, as in the previous section, that $f_{qq}(0) > 0$, then

$$DHS = O_p(T).$$

The statistic diverges in probability at the same rate as $\hat{Z}_\alpha$ which indicates good power of the test in finite-sample applications. The critical values for the DHS statistic can be found in [2].

2.3. H_0: Cointegration

2.3.1. Test for cointegration by variable addition

Park [13] proposes two alternative statistics: the first for the null hypothesis of cointegration and the other for the null hypothesis of no cointegration. Both of them are based on the same idea. To the cointegrating regression we add some superfluous regressors and when the cointegration is present (the regression errors are stationary) the test is able to detect the insignificance of the added variables. When there is no cointegrating relationship among the variables (the regression errors are $I(1)$) the regression is spurious and the added variables are likely to be significant.

We estimate the relation

$$y_t^* = \hat{\alpha}_*' \mathbf{x}_t^* + \hat{\beta}_1' \mathbf{s}_{1t} + \hat{\beta}_2' \mathbf{s}_{2t} + \hat{e}_t \tag{6}$$

by OLS. $\{\mathbf{s}_{1t}\}$ is a vector of k_2 deterministic trends and $\{\mathbf{s}_{2t}\}$ is a vector of m_2 processes $I(1)$ which are not cointegrated with $\{\mathbf{x}_t\}$. More details about the added regressors can be found in [13]. The variables y_t^* and $\mathbf{x}_t^*$ are transformations of the original variables y_t and $\mathbf{x}_t$, because the asymptotic distribution of the Wald test for $H_0 : \beta_1 = \beta_2 = 0$ in (6) with the original variables y_t and $\mathbf{x}_t$ is not free of nuisance parameters.

A simple version of this transformation for a bivariate case which is used for the power comparison in Section 3 is the following.

We regress

$$y_t = \hat{\alpha}_1 + \hat{\alpha}_2 x_t + \hat{e}_t,$$
$$x_t = \hat{\alpha}_3 + \hat{d}_t,$$

where α_1 and α_3 are constants, $\hat{e}_t$ and $\hat{d}_t$ are the corresponding residuals. Then for $\hat{q}_t = (\hat{e}_t, \Delta\hat{d}_t)'$, we define

$$\begin{aligned}
\hat{\Sigma}_{aa} &= 1/T\sum_{t=1}^{T}\hat{q}_t\hat{q}_t',\\
\hat{\Omega}_{aa} &= \begin{bmatrix}\hat{\omega}_{00} & \hat{\omega}_{01}\\ \hat{\omega}_{10} & \hat{\Omega}_{11}\end{bmatrix}\\
&= 1/T\sum_{t=1}^{T}\hat{q}_t\hat{q}_t' + 1/T\sum_{k=1}^{\ell}w_{s\ell}\sum_{t=k+1}^{T}(\hat{q}_t\hat{q}_{t-k}' + \hat{q}_{t-k}\hat{q}_t'),\\
\hat{\Delta}_{aa} &= \begin{bmatrix}\hat{\delta}_{00} & \hat{\delta}_{01}\\ \hat{\delta}_{10} & \hat{\Delta}_{11}\end{bmatrix},\\
&= 1/T\sum_{t=1}^{T}\hat{q}_t\hat{q}_t' + 1/T\sum_{s=1}^{\ell}w_{s\ell}\sum_{t=s+1}^{T}\hat{q}_t\hat{q}_{t-k}',\\
\hat{\Delta}_{1a} &= (\hat{\delta}_{10}, \hat{\Delta}_{11}),\\
\hat{c}_a &= (0, \hat{\omega}_{01}\hat{\Omega}_{11}^{-1}),
\end{aligned}$$

where $w_{s\ell}$ is some window which fulfils the condition $\sum_{s=1}^{\ell}|w_{s\ell}| = O(T^{\delta})$, $0 \leqslant \delta \leqslant \frac{1}{2}$.

The transformed variables are defined as

$$\begin{aligned}
\mathbf{x}_t^* &= \mathbf{x}_t - \hat{\Delta}_{1a}\hat{\Sigma}_{aa}\hat{q}_t,\\
y_t^* &= y_t - (\hat{\gamma}'\hat{\Delta}_{1a}\hat{\Sigma}_{aa}^{-1} + \hat{c}_a')\hat{q}_t.
\end{aligned}$$

The statistic for the H_0 of cointegration is

$$J_1 = \frac{RSS_x^* - RSS_{xs}^*}{\hat{\omega}_{00.1}},$$

where RSS_{xs}^* and RSS_x^* are the residual sum of squares of the regression (6) with and without the superfluous regressors and $\hat{\omega}_{00.1} = \hat{\omega}_{00} - \hat{\omega}_{01}\hat{\Omega}_{11}\hat{\omega}_{10}$.

Under H_0: cointegration

$$J_1 \overset{a}{\sim} \chi^2_{m_2+k_2}.$$

Under H_a: no cointegration

$$J_1 = O_p(T^{1-\delta}),$$

where δ is set by the definition of the corresponding window $w_{s\ell}$. Under the alternative the statistic diverges in probability depending on the selected window. The smaller δ, the faster the divergence of J_1.

An obvious advantage is that no special tables for this test are required, but the choice of the superfluous regressors and the window $w_{s\ell}$ will probably be crucial for the power of the test.

2.3.2. The Shin's C and Leybourne and McCabe's LBI tests

The C statistic [17] and LBI statistic [11] are based on the structural model of local trend. Consider the following model:

$$\begin{aligned} y_t &= \gamma' \mathbf{x}_t + u_t \\ u_t &= \nu_t + \psi_t \\ \nu_t &= \nu_{t-1} + \varepsilon_t \end{aligned}$$

where ψ_t is a stationary process and ε_t is $iid(0, \sigma_\varepsilon^2)$ and ψ_t and ε_t are independent. The null hypothesis is $H_0 : \sigma_\varepsilon = 0$, that is to say, u_t has no random walk component.

The C and LBI statistics are defined as

$$C = LBI = \frac{\sum_{t=1}^{T} S_t^2}{T^2 s^2(\ell)},$$

where S_t is the partial sum process of the residuals from the cointegrating regression and $s^2(\ell)$ is a consistent semiparametric estimator of the long run variance of the regression error u_t. Shin [17] uses the triangular Barlett window for the estimation of the long run variance in the empirical section of his paper:

$$s_C^2(\ell) = T^{-1} \sum_{t=1}^{T} \hat{u}_t^2 + 2T^{-1} \sum_{s=1}^{\ell} (1 - s(\ell+1)^{-1}) \sum_{t=s+1}^{T} \hat{u}_t \hat{u}_{t-s}, \tag{7}$$

whereas Leybourne and McCabe [11] use the rectangular window:

$$s_{LBI}^2(\ell) = T^{-1} \sum_{t=1}^{T} \hat{u}_t^2 + 2T^{-1} \sum_{s=1}^{\ell} \sum_{t=s+1}^{T} \hat{u}_t \hat{u}_{t-s}. \tag{8}$$

If the variables $\mathbf{x}_t$ are strictly exogenous with respect to u_t, then the test statistics C and LBI have the following limiting distribution.

Under H_0: cointegration

$$LBI = C \Rightarrow \int_0^1 G^2, \tag{9}$$

where

$$G = W_1 - \left(\int_0^r W_2'\right)\left(\int_0^1 W_2 W_2'\right)^{-1}\left(\int_0^1 W_2 dW_1\right).$$

Under H_0: no cointegration

$$C = O_p(T/\ell),$$

where ℓ is the selected lag of the window used for the estimation of $s^2(\ell)$. Then the power of the test depends on ℓ.

The exogeneity assumption is too restrictive in time series modeling. Shin [17] (using the result of Saikkonen [16]) proposes a modification of the cointegrating regression in order to relax this assumption. The proposed modification consists in using the present, past and future values of $\Delta \mathbf{x}_t$ as additional regressors. See [17] for further details. Therefore, we consider the modified least squares regression equation

$$y_t = \hat{\gamma}'\mathbf{x}_t + \sum_{j=-K}^{K} \hat{\pi}_j' \Delta \mathbf{x}_{t-j} + \hat{u}_t^*. \tag{10}$$

The C statistic defined using the new residuals $\hat{u}_t^*$ has the same limiting distribution (9) without the assumption of strict exogeneity.

The statistics C and LBI are very simple to calculate and the critical values can be found in [17]. Of course, the choice of the lag truncation parameter ℓ and the number K for the leads and lags of the additional regressors will be a problem in empirical applications. Shin [17] proposes the use of the AIC or BIC criteria for the selection of K.

2.3.3. Hausman-like tests $H1$ and $H2$

As in the case of Choi's test, the statistics $H1$ and $H2$ (see [4] and [5]) are based on the comparison of two estimators, both of them consistent under the null but one of them inconsistent under the alternative hypothesis.

The OLS estimation of (10) will produce under cointegration an efficient and superconsistent (the estimates converge to its true value at the rate T^{-1} rather than the usual $T^{-1/2}$) estimator of the cointegrating vector $(1, \hat{\gamma}_L{}')'$ whose asymptotic distribution is free of the nuisance parameters π_j (see [16]). The error term u_t^* is asymptotically uncorrelated with the increments of $\mathbf{x}_t$ at all leads and lags. We define

$$z_t = y_t - \sum_{j=-K}^{K} \hat{\pi}_j' \Delta \mathbf{x}_{t-j}.$$

The initial model (1) may thus be rewritten as

$$z_t = \gamma' \mathbf{x}_t + u_t^*.$$

Taking differences we get

$$\Delta z_t = \gamma' \Delta \mathbf{x}_t + \Delta u_t^*,$$

where $\Delta \mathbf{x}_t$ and Δu_t^* are asymptotically uncorrelated and the OLS estimator $\hat{\gamma}_D$ will give us the usual $\sqrt{T}$-consistent estimator under cointegration.

Both estimators $\hat{\gamma}_L$ and $\hat{\gamma}_D$ are used for defining Hausman-like statistics H_1 and H_2 as $\hat{\gamma}_L$ is inconsistent under no cointegration (spurious regression) and $\hat{\gamma}_D$ is an $O_p(T^{-1/2})$ estimator (asymptotically biased since the correlation between $\Delta \mathbf{x}_t$ and Δu_t^* may not be equal to zero in general) under the alternative hypothesis.

The test statistics are

$$H_1 = (\hat{\gamma}_L - \hat{\gamma}_D)'(\hat{V}_D + \hat{V}_L)(\hat{\gamma}_L - \hat{\gamma}_D),$$
$$H_2 = (\hat{\gamma}_L - \hat{\gamma}_D)'\hat{V}_D(\hat{\gamma}_L - \hat{\gamma}_D),$$

where $\hat{V}_D$ and $\hat{V}_L$ are consistent estimates of the covariance matrices of $\hat{\gamma}_L$ and $\hat{\gamma}_D$, respectively. Since V_L is $O_p(T^{-2})$ while V_D is $O_p(T^{-1})$ the two statistics are asymptotically equivalent.

Under H_0: cointegration

$$H1 \overset{a}{\sim} \chi^2_{m-1}$$

$$H2 \overset{a}{\sim} \chi^2_{m-1}.$$

Therefore, the asymptotic distribution of the test statistics $H1$ and $H2$ is a standard χ^2 distribution. The critical values for the sample size of $T = 10, 20, \ldots, 500$ and for 1 to 4 regressors can be found in [5].

Under H_a: no cointegration

$$H1 = O_p(T)$$
$$H2 = O_p(T).$$

3. Comparison of the empirical size and power

In this section the empirical size and power of the nine above presented tests in a bivariate case are compared. For the comparison the following data generating process (DGP) is used:

DGP(1)

$$x_t = x_{t-1} + v_t \qquad v_t = \varphi_1 v_{t-1} + \varepsilon_{1t} + \theta_1 \varepsilon_{1,t-1}$$

$$y_t = x_t + u_t \qquad u_t = \varphi_2 u_{t-1} + \varepsilon_{2t} + \theta_2 \varepsilon_{2,t-1}$$

$$\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix} = N(0, \Sigma) \qquad |\varphi_1|, |\theta_1| < 1.$$

The values of φ_2 and θ_2 determine whether the variables x_t and y_t are cointegrated. If $\theta_2 = 0$, then for the values $|\varphi_2| \in [0, 1)$ the series y_t and x_t are cointegrated and the cointegrating vector is $(1, -1)$. Increasing the value of φ_2 from 0 to 1 we are approaching the case of no cointegration and we can observe changes in

- the power of the test for H_0: no cointegration
- the size of the test for H_0: cointegration

for different values of φ_2. Changes in the value of φ_2 should not have any impact on size or power of the tests. We should observe a drastic change of size (or power) in the extreme case $\varphi_2 = 1$ where the cointegration is replaced by no cointegration.

If we maintain the value of $\varphi_2 = 1$ and decrease the value of θ_2 from 0 to -1, we are moving away from the extreme case of no cointegration ($\varphi_2 = 1$ and $\theta_2 = 0$) and at the same time we are approaching the initial case of cointegration since, for $\varphi_2 = 1$ and $\theta_2 = -1$, the process $u_t = u_{t-1} + \varepsilon_{2t} - \varepsilon_{2,t-1}$ cannot be distinguished from the process $u_t = \varepsilon_{2t}$ (if $u_0 = \varepsilon_0$), which is a $I(0)$ series.

Then, decreasing the values of θ_2 we can observe changes in

- the power of the test for H_0: cointegration
- the size of the test for H_0: no cointegration.

All the previous cases can be summarized in the following scheme.

$$\underbrace{\begin{array}{l}\theta_2 = 0\\ \varphi_2 = 0 \cdots 0.2 \cdots 0.4 \cdots 0.6 \cdots 0.8 \cdots\end{array}}_{\text{cointegration}}\underbrace{\begin{array}{l}\varphi_2 = 1\\ \theta_2 = 0 \cdots -0.2 \cdots -0.4 \cdots -0.6 \cdots -0.8 \cdots\end{array}}_{\text{no cointegration}}\underbrace{\begin{array}{r}\varphi_2 = 1\\ \theta_2 = -1\end{array}}_{\text{cointegration}}$$

For the empirical size and power comparison, the DGP(1) has been used with the following parameter values:

$$\varphi_1 = 0.45, \qquad \theta_1 = 0.35,$$

$$\underbrace{\varphi_2 = 0; 0.2; 0.4; 0.6; 0.8; 1,}_{\theta_2 = 0} \qquad \underbrace{\theta_2 = -0.2; -0.4; -0.6; -0.8; -1,}_{\varphi_2 = 1}$$

$$\Sigma = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}.$$

DGP(1) is complicated enough to be considered a good approach to the large scale of real processes. We have to point out that processes u_t and v_t are serially and contemporaneously correlated. That is to say, the regressor in the cointegrating regression is serially and contemporaneously correlated with the disturbance term.

The results of Monte Carlo simulation for sample size $T = 100$ and 5000 iterations are summed up in tables (1)–(4). For the simulation, the software package RATS version 4.10 was used and the random normal numbers were generated by the function %RAN(x).

The triangular Barlett window $w_{s\ell} = 1 - s/(\ell + 1)$ was used for the estimation of the long run variance in the case of $\hat{Z}_\alpha$, $\hat{Z}_t$, DHS, $J1$ and C statistic. The ℓ value in the tables (1)–(4) represents the lag truncation parameter for the Barlett window. The value of p in the case of ADF statistics is the number of lags $\Delta\hat{u}_t$ in the equation (2). The parameter K which represents the number of leads and lags of $\Delta\mathbf{x}_t$ in the modified cointegrating regression (10) was set equal to 0 because the results with higher values of K were worse for both the C and the Hausman-like statistic.

The following conclusions can be drawn.

- ADF: The power of this test is sufficiently high but it does not reach the power of the $\hat{Z}_\alpha$, $\hat{Z}_t$, or DHS test, especially in cases where $\varphi_2 \in [0.6, 1)$. Another conclusion is that the empirical size depends crucially on the number of lags p. If we add a few lags of $\Delta\hat{u}_t$ the test will suffer from a serious overrejection problem which will make it rather useless in practice. When $p = 0$ then for the values $\varphi_2 = 1$ and $\theta_2 = -0.6$ (under *No cointegration*) this test rejects the right null hypothesis of no cointegration about 42 % of time. We can observe that the overrejection problem can be solved by increasing the value of p, but the cost of this solution is a lower power under the alternative hypothesis. For example for the values $p = 4$, $\varphi_2 = 0.8$ and $\theta_2 = 0$ the test rejects the wrong null hypothesis only in 57 % of the cases.
- $\hat{Z}_\alpha$ and $\hat{Z}_t$: The power of these tests is extraordinary, but it seems to be at the expense of distortion of the size. The number of rejections of the true null hypothesis for the values $\theta_2 = -0.4$ and lower ($\varphi_2 = 1$) moves between 35% and 90%. One advantage we should point out is the independence of the results of the value of ℓ.

When we use these tests we should keep in mind that they sometimes tend to reject the null hypothesis of no cointegration when it is the true hypothesis. Then we tend to find the cointegrating relationship among no cointegrating series.

- *DHS*: The conclusions for the *DHS* statistics could be the same as in the case of the previous tests $\hat{Z}_\alpha$ and $\hat{Z}_t$. The simulation experiments confirm the results of Choi [2], who argues that the *DHS* statistic performs slightly better than $\hat{Z}_\alpha$ and $\hat{Z}_t$ tests when φ_2 is close to 1. Obviously this is a very problematic region where it is very difficult to distinguish between cointegration and near cointegration. The large distortion of the size, however, has to be left for further research.
- *J*1: The power of the *J*1 statistic is very good but the overrejection problem under the null hypothesis seems to be too large. As Park (see [13]) argues, the selection of the additional regressors is crucial for the test performance. In our simulation experiments, 5 additional regressors have been used. Two of them were deterministic trends ($0.2t$ and $0.003t^2$) and the other three regressors were independent random walks:

$$x_{1t} = x_{1,t-1} + \varepsilon_{1t}$$
$$x_{2t} = x_{2,t-1} + \varepsilon_{2t}$$
$$x_{3t} = x_{3,t-1} + \varepsilon_{3t}$$

$$\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix} = N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 9 \end{pmatrix}\right).$$

The results for less additional regressors (one deterministic trend and one random walk) were still more disastrous.
- *C*: As we have argued in Section 2, the performance of this test depends on ℓ (under H_0: no cointegration $C = O_p(T/\ell)$). For small values of ℓ the test has a very high power which does not depend on the value of θ_2, but the overrejection problem under the null hypothesis seems to be too serious. If we increase the value of ℓ, the number of rejections under the null hypothesis gets closer to the theoretical value of 5 % (1 %) but the power of the test (under the alternative) drops quickly. The evident advantage as compared to the previous test is that the power does not depend on the value of θ_2.
- *LBI*: The number of rejections of this test under the null hypothesis is stable enough for the majority of the values of φ_2. The lack of power under the alternative seems to be an important problem of this test. The values around 40 %

(for $\ell = 4$) are not enough for practical purposes, that is to say, we can easily accept the cointegrating relationship among series that are not cointegrated.

- $H1$, $H2$: The tests perform very well under the null hypothesis of cointegration and the number of rejections of the null is very close to the theoretical value (5 % and 1 %) for the majority of the values of φ_2 (< 1). Under the alternative hypothesis the tests enjoy a high power for the values $\theta_2 > -0.6$ ($\varphi_2 = 1$) and when θ_2 gets closer to -1 the tests lose power. But this is the only region where these tests perform worse than any other test (with the empirical size close to the nominal one) of the null hypothesis of cointegration.

Some of these results are summarized in Figure 1. At first sight we can observe that the tests of the null hypothesis of no cointegration have a higher power[1] than the test with the opposite null hypothesis. The overrejection problem under the null hypothesis seems to be the cost paid for this extraordinary power. On the contrary, the tests of the null hypothesis of cointegration perform better under the null hypothesis than under the alternative.

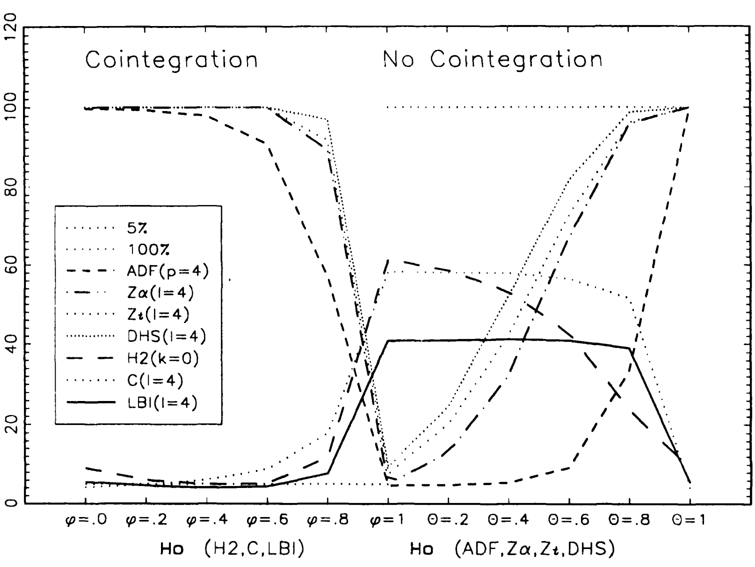

Figure 1. Power and empirical size of cointegration tests

[1] However, the direct comparison would be incorrect

4. Conclusion

Which null hypothesis should we use in our studies? If we base the decision on the results of our Monte Carlo experiment, there is no obvious answer. In general, the tests of the null hypothesis of cointegration perform very well under the null but tend to fail under the alternative. The tests of the null of no cointegration exhibit excellent power under the alternative but suffer from the overrejection problem under the null. A clear conclusion of the Monte Carlo experiment is that in the empirical studies no-cointegration will be a much more difficult relationship to be discovered than cointegration.

Nevertheless, the present Monte Carlo study offers the possibility of combining two cointegration tests with opposite null hypotheses to obtain more precise results. If we accept no cointegration and reject cointegration, then there is strong evidence for no cointegration. Similarly, if we reject no cointegration and accept cointegration, then this is strong evidence for cointegration. These statements assume that both tests have good power and do not suffer the overrejection problem under the null hypothesis. If both null hypotheses are rejected, then a type I error may have occurred in one of the tests. If both null hypotheses are accepted, then one of the tests may suffer lack of power.

Thus, using Monte Carlo simulation for a large scale of DGP we could conclude that the best combination for the testing of cointegration could be $\hat{Z}_{\alpha}(DHS)$–$H2$.

Acknowledgement. I would like to thank Professors J. Fernández-Macho and M. Regúlez for helpful advice and valuable comments on this paper.

		Cointegration					No Cointegration
		$\varphi_2 = 0.0$	$\varphi_2 = 0.2$	$\varphi_2 = 0.4$	$\varphi_2 = 0.6$	$\varphi_2 = 0.8$	$\varphi_2 = 1.0$
H_0: *No Cointegration*							
	$p = 1$	100.00	100.00	100.00	99.92	83.46	5.46
ADF	$p = 4$	99.72	99.36	97.92	90.88	56.84	4.66
	$p = 8$	77.32	72.96	66.28	54.80	31.46	3.82
	ℓ=0	100.00	100.00	100.00	100.00	88.48	3.82
$\hat{Z}_\alpha$	ℓ=4	100.00	100.00	100.00	100.00	89.46	5.04
	ℓ=8	100.00	100.00	100.00	100.00	89.82	4.88
	ℓ=12	100.00	100.00	100.00	100.00	88.48	3.82
	ℓ=0	100.00	100.00	100.00	100.00	91.84	7.62
$\hat{Z}_t$	ℓ=4	100.00	100.00	100.00	100.00	91.78	7.62
	ℓ=8	100.00	100.00	100.00	100.00	91.78	7.86
	ℓ=12	100.00	100.00	100.00	100.00	92.34	8.32
	ℓ=0	100.00	100.00	100.00	100.00	96.84	8.10
DHS	ℓ=4	100.00	100.00	100.00	100.00	96.96	10.34
	ℓ=8	100.00	100.00	100.00	100.00	97.48	10.14
	ℓ=12	100.00	100.00	100.00	100.00	97.22	9.42
H_0: *Cointegration*							
$J1$	ℓ=0	97.48	97.72	97.50	97.24	97.90	99.66
	ℓ=4	99.4	99.78	98.96	95.48	86.84	95.56
$H1$		8.98	5.96	5.06	5.12	11.56	58.28
$H2$		9.02	6.02	4.92	5.18	11.72	61.28
	ℓ=0	4.56	9.38	18.14	32.88	58.60	91.96
C	ℓ=4	4.10	4.74	6.06	8.78	17.74	58.16
	ℓ=8	3.74	4.04	4.46	5.68	9.82	42.72
LBI	ℓ=2	4.54	4.22	4.90	6.82	15.64	57.32
	ℓ=4	5.48	4.62	4.08	4.38	7.72	40.80

Significance level 5 %

Table 1. Power and empirical size

		Cointegration					No Cointegration
		$\varphi_2 = 0.0$	$\varphi_2 = 0.2$	$\varphi_2 = 0.4$	$\varphi_2 = 0.6$	$\varphi_2 = 0.8$	$\varphi_2 = 1.0$
H_0: *No Cointegration*							
	$p = 1$	100.00	100.00	99.98	97.66	46.14	1.40
ADF	$p = 4$	94.76	91.04	81.56	60.72	21.44	0.88
	$p = 8$	39.66	35.92	30.02	21.22	8.92	0.58
	ℓ=0	100.00	100.00	100.00	99.76	49.54	0.74
$\hat{Z}_\alpha$	ℓ=4	100.00	100.00	100.00	99.72	53.42	0.94
	ℓ=8	100.00	100.00	100.00	99.64	50.18	0.76
	ℓ=12	100.00	100.00	100.00	99.76	49.54	0.74
	ℓ=0	100.00	100.00	100.00	99.90	62.84	2.80
$\hat{Z}_t$	ℓ=4	100.00	100.00	100.00	99.90	62.30	2.76
	ℓ=8	100.00	100.00	100.00	99.90	63.04	2.80
	ℓ=12	100.00	100.00	100.00	99.94	65.06	2.98
	ℓ=0	100.00	100.00	100.00	100.00	80.88	2.78
DHS	ℓ=4	100.00	100.00	100.00	100.00	82.02	3.44
	ℓ=8	100.00	100.00	100.00	100.00	81.92	3.40
	ℓ=12	100.00	100.00	100.00	100.00	78.98	2.94
H_0: *Cointegration*							
$J1$	ℓ=0	94.68	95.00	94.50	94.08	95.48	99.52
	ℓ=4	99.84	99.52	98.12	92.16	76.78	90.96
$H1$		1.96	1.04	0.86	1.28	3.76	47.80
$H2$		1.96	1.06	0.90	1.20	3.98	51.62
	ℓ=0	0.58	2.48	6.74	15.84	38.16	81.74
C	ℓ=4	0.32	0.50	0.74	1.76	5.54	39.80
	ℓ=8	0.14	0.20	0.26	0.40	1.66	20.44
LBI	ℓ=2	0.62	0.50	0.50	1.02	4.50	38.80
	ℓ=4	0.80	0.48	0.36	0.28	1.06	18.94

Significance level 1 %

Table 2. Power and empirical size

		Cointegration					No Cointegration
		$\theta_2 = 0.0$	$\theta_2 = -0.2$	$\theta_2 = -0.4$	$\theta_2 = -0.6$	$\theta_2 = -0.8$	$\theta_2 = -1.0$
H_0: *No Cointegration*							
	$p = 0$	5.46	6.76	15.54	41.98	85.80	100.00
ADF	$p = 4$	4.66	4.70	5.34	9.20	33.50	99.72
	$p = 8$	3.82	3.92	3.78	4.62	10.14	77.48
	$\ell=0$	3.82	14.86	41.06	75.82	97.98	100.00
$\hat{Z}_\alpha$	$\ell=4$	5.04	12.98	32.38	66.96	95.86	100.00
	$\ell=8$	4.88	15.60	38.82	73.32	97.22	100.00
	$\ell=12$	3.82	14.86	41.06	75.82	97.98	100.00
	$\ell=0$	7.62	19.94	41.14	71.88	95.78	100.00
$\hat{Z}_t$	$\ell=4$	7.62	20.24	42.34	72.86	96.18	100.00
	$\ell=8$	7.86	19.96	40.36	71.18	95.84	100.00
	$\ell=12$	8.32	20.50	41.80	72.68	96.22	100.00
	$\ell=0$	8.10	23.62	52.26	82.50	98.84	100.00
DHS	$\ell=4$	10.34	21.00	42.96	74.56	97.16	100.00
	$\ell=8$	10.14	23.26	48.12	78.76	98.18	100.00
	$\ell=12$	9.42	24.68	52.10	81.84	98.72	100.00
H_0: *Cointegration*							
$J1$	$\ell=0$	99.66	99.80	99.80	99.76	99.34	97.34
	$\ell=4$	95.56	96.02	97.02	97.86	99.14	99.90
$H1$		58.28	56.26	51.16	40.88	23.04	8.98
$H2$		61.28	58.50	53.22	42.72	23.64	9.00
	$\ell=0$	91.96	91.26	89.66	85.46	70.88	4.32
C	$\ell=4$	58.16	57.90	57.78	56.38	51.36	3.82
	$\ell=8$	42.72	43.40	43.60	43.10	41.04	3.76
LBI	$\ell=2$	57.32	57.16	56.84	55.76	50.88	4.54
	$\ell=4$	40.80	40.82	41.14	40.82	38.88	5.48

Significance level 5 %

Table 3. Power and empirical size

		Cointegration					No Cointegration
		$\theta_2 = 0.0$	$\theta_2 = -0.2$	$\theta_2 = -0.4$	$\theta_2 = -0.6$	$\theta_2 = -0.8$	$\theta_2 = -1.0$
H_0: *No Cointegration*							
	$p = 0$	1.40	1.90	5.16	24.24	74.06	100.00
ADF	$p = 4$	0.88	1.08	1.16	2.24	14.56	94.78
	$p = 8$	0.58	0.56	0.58	0.82	2.40	39.76
	ℓ=0	0.74	4.78	22.90	60.98	94.90	100.00
$\hat{Z}_\alpha$	ℓ=4	0.94	4.40	18.84	54.68	92.78	100.00
	ℓ=8	0.76	5.70	24.90	62.78	95.22	100.00
	ℓ=12	0.74	4.78	22.90	60.98	94.90	100.00
	ℓ=0	2.80	9.10	25.92	58.28	92.02	100.00
$\hat{Z}_t$	ℓ=4	2.76	9.30	27.32	59.96	92.46	100.00
	ℓ=8	2.80	9.22	26.38	59.42	92.80	100.00
	ℓ=12	2.98	9.48	27.56	61.98	93.98	100.00
	ℓ=0	2.78	12.04	36.76	73.14	97.38	100.00
DHS	ℓ=4	3.44	10.76	30.18	65.60	95.68	100.00
	ℓ=8	3.40	12.78	35.50	71.60	98.18	100.00
	ℓ=12	2.94	13.96	39.82	74.98	97.80	100.00
H_0: *Cointegration*							
$J1$	ℓ=0	99.52	99.40	99.66	99.46	98.64	94.92
	ℓ=4	90.96	92.14	93.72	95.58	98.28	99.76
$H1$		47.80	44.60	38.30	27.48	11.58	1.96
$H2$		51.62	48.40	41.56	29.68	12.22	2.04
	ℓ=0	81.74	81.04	78.94	73.40	57.08	0.64
C	ℓ=4	39.80	40.28	40.04	38.86	33.48	0.32
	ℓ=8	20.44	20.86	20.84	20.56	18.28	0.12
LBI	ℓ=2	38.88	38.80	38.56	37.54	31.98	0.62
	ℓ=4	18.94	19.00	19.06	19.34	17.96	0.08

Significance level 1 %

Table 4. Power and empirical size

References

[1] *B. Arshanapalli, J. Doukas*: International stock market linkages: Evidence from the pre- and post-october 1987 period. Journal of Banking and Finance *17* (1993), 193–208.

[2] *I. Choi*: Durbin-Hausman tests for cointegration. Journal of Economic Dynamics and Control *18* (1992), 467–480.

[3] *R.F. Engle, C.W.J. Granger*: Co-integration and error correction: Representation, estimation and testing. Econometrica *55* (1987), 251–276.

[4] *F.J. Fernández-Macho*: Hausman-like and variance-ratio test statistics for the null of cointegration. Working paper D.T. 94/15, Dpto. de Econometría y Estadística, Universidad del País Vasco, Bilbao. 1994.

[5] *F.J. Fernández-Macho, P. Mariel*: Testing the null of cointegration: Hausman-like test for regressions with a unit root. Working paper D.T. 94/18, Dpto. de Econometría y Estadística, Universidad del País Vasco, Bilbao, 1994.

[6] *N. Haldrup*: The asymptotics of single-equation cointegration regressions with I(1) and I(2) variables. Journal of Econometrics *63* (1994), 153–181.

[7] *B.E. Hansen*: Tests for parameter instability in regressions with I(1) processes. Journal of Business and Economic Statistics *10* (1992), 321–335.

[8] *S. Johansen, K. Juselius*: Testing structural hypotheses in a multivariate cointegration analysis of the PPP and the UIP for UK. Journal of Econometrics *53* (1992), 211–244.

[9] *K.Kasa*: Common stochastic trends in international stock markets. Journal of Monetary Economics *29* (1992), 95–124.

[10] *J.J.M. Kremers, N.R. Ericsson, J.J. Dolado*: The power of cointegration tests. Oxford Bulletin of Economics and Statistics *54* (1992), 325–348.

[11] *S.J. Leybourne, B.P.M. McCabe*: A simple test for cointegration. Oxford Bulletin of Economics and Statistics *56* (1994), 97–103.

[12] *Y. Ngama, S. Sosvilla-Rivero*: An empirical examination of absolute purchasing power parity: Spain 1977–1988. Working paper 91-08, University of Birmingham, 1991.

[13] *J.Y. Park*: Testing for unit roots and cointegration by variable addition. In: Advances in Econometrics: Co-integration, Spurious Regressions and Unit Roots. T.B. Fomby and G.F. Rhodes, Greenwich, JAI Press, 1990, pp. 107–133.

[14] *P.C.B. Phillips, S.N. Durlauf*: Multiple time series regression with integrated processes. Review of Economic Studies *53* (1986), 473–496.

[15] *P.C.B. Phillips, S. Ouliaris*: Asymptotic properties of residual based tests for cointegration. Econometrica *58* (1990), 165–193.

[16] *P. Saikkonen*: Asymptotically efficient estimation of cointegration regressions. Econometric Theory *7* (1991), 1–21.

[17] *Y. Shin*: A residual-based test of the null of cointegration against the alternative of no cointegration. Econometric theory *10* (1994), 91–115.

Author's address: *Petr Mariel*, Dpto. de Econometría y Estadística, Fac. de Ciencias Económicas (UPV/EHU), Lehendakari Agirre 83, E48015 Bilbao, Spain, e-mail (internet): `etpmaxxp@bs.ehu.es`.