A CHARACTERIZATION of COINTEGRATING RELATIONSHIPS using INDUCED-ORDER STATISTICS

*Felipe M. Aparicio and Alvaro Escribano**

Universidad Carlos III de Madrid

Department of Statistics and Econometrics

C/Madrid, 126-128

28903 Getafe (Madrid)

Spain

Fax: +34 1 624 98 49

Phone: +34 1 624 98 71

E-mail: aparicio@est-econ.uc3m.es

*This research, initiated at the Department of Economics (Aparicio and Granger, 1995[2]) and the Institute for Nonlinear Science of the University of California at San Diego, was partially supported by the Spanish DGICYT under Grant PB95-0298. We are grateful to Ignacio Pena and to Z. Ding for kindly providing the foreign exchange-rate and the stock return series, respectively.

Abstract

In this paper we explore the usefulness of *induced-order statistics* in the characterization of integrated series and of cointegrating relationships. We propose some cointegration testing devices which do not require prior estimation of the cointegration parameter, and therefore lead to null distributions which are free from nuisance parameters. To test the null of non-cointegration, we proposed using the parametric Dickey-Fuller test statistic on a pair of series obtained from the original ones. These series must be cointegrated with cointegration parameter equal to unity whenever the original series are cointegrated. Thus the null distribution of this test is exactly the Dickey-Fuller one, since it does not depend on the estimated regression parameter. We also proposed a pair of non-parametric test statistics for testing the null hypothesis of stationarity of the errors in the regression of two $I(1)$ time series. These tests are powerful against the wide alternative of non-cointegration or of pair of (possibly co-moving) $I(0)$ series. Therefore they do not require prior testing for unit roots in the series.

KEY WORDS: *Integrated series, cointegration, Dickey-Fuller test, induced-order statistics, robustness.*

1 Introduction

Processes which exhibit common trends or similar long waves in their sample paths are often called *cointegrated*. The concept of cointegration originated in macroeconomics and finance (c.f. Granger, 1983 [10]; Granger and Engle, 1987 [8]), where in some cases the theory suggests the presence of economic or social forces preventing two or more series to drift too far apart from each other. Take for example, those series as income and expenditure, the prices of a particular good in different markets, the interest rates in different parts of a country, etc. Cointegration relationships may also appear in other contexts, such as in the outputs of different sensing or processing devices having a limited storage capacity (or memory) and driven by a common persistent input flow (c.f. Aparicio, 1995 [1]).

Underlying the idea of cointegration is that of an *equilibrium* relationship (i.e. one that holds on the average) between two cointegrated variables, x_t, y_t. A strict equilibrium exists when for some $\alpha \neq 0$, one has $y_t = \alpha x_t$. This unrealistic situation is replaced, in practice, by that of (linear) cointegration, in which the equilibrium error $z_t = y_t - \alpha x_t$ is different from zero but fluctuates around this value much more frequently than the individual series, and the size of the fluctuations are much smaller.

Standard tests for cointegration, aimed at testing the null hypothesis of non-cointegration, can be decomposed into two stages:

- A test for long-memory in the variables, say $x_t \sim I(d_x), y_t \sim I(d_y)$, and estimation of the long-memory parameters d_x, d_y. Then a test of significance for the stochastic difference $\hat{d}_x - \hat{d}_y$. If it is too large as compared to both d_x and d_y, then the variables cannot be cointegrated. Otherwise, we assume $d_x = d_y$ and go to next step.

- A test for long-memory in the cointegrating residuals $\xi_t = y_t - \hat{\alpha} x_t$, and estimation of its long-memory parameter, d_ξ. Then a test of significance for the stochastic difference $\hat{d}_x - \hat{d}_\xi$. Large positive values of this difference as compared to d, can be taken as evidence of the existence of a cointegrating relationship between x_t and y_t.

A most investigated case corresponds to when the long-memory features in the variables are only due to unit roots (i.e. $d_x = d_y = 1$). This simplifies the procedure since no estimation of long-memory parameters is required in this case. However, a *test for unit roots* is needed to confirm this hypothesis.

In a one-sided test for unit roots, the null hypothesis of a unit root in a series x_t, H_0 : $(1-B)x_t = \xi_t$, is tested against the alternative H_1 : $(1-\rho B)(x_t - \mu) = \xi_t$ with $|\rho| < 1$, where ξ_t denotes a sequence of *i.i.d. r.v.*'s. In a more general setting, the alternative hypothesis of random walk with drift, H_3 : $(1-B)x_t = c + \xi_t$, and that of stationary AR(1) with a linear trend, H_4 : $(1-B)(x_t - a - bt) = \xi_t$, may also be tested if H_1 is discarded at a first stage (see Dickey et al., 1986 [6] for details).

The decision between H_0 and H_1 is based on the significance of the estimate of $\rho - 1$ in the regression:

$$\Delta x_t = (\rho - 1)x_{t-1} + \epsilon_t \tag{1}$$

with Δ denoting the first differencing operator. The test statistic $\tau = N(\hat{\rho} - 1)$ is commonly used, with N denoting the sample size. Under H_0, this test stastistic can be written as

$$\tau_0 = N \frac{\sum_{t=1}^{N} x_{t-1}\epsilon_t}{\sum_{t=1}^{N} x_{t-1}^2} \tag{2}$$

The distribution of τ_0 is a non-standard one, and is obtained under the following assumptions (c.f. Phillips, 1987[13]):

ASSUMPTION AS1 :

$E(\epsilon_t) = 0$.

ASSUMPTION AS2 :

$\sup_t E(|\epsilon_t|^\gamma) < \infty$ for some $\gamma > 2$.

ASSUMPTION AS3 :

$0 < \lim_{N\to\infty} E\left(N^{-1}(\sum_{t=1}^{N} \epsilon_t)^2\right) < \infty$.

ASSUMPTION AS4 :

ϵ_t is strong mixing, with mixing coefficients α_i satisfying $\sum_{i=1}^{\infty} \alpha_i^{1-2/\gamma} < \infty$.

Under assumptions AS1 - AS4, and under H_0, one has

$$\tau \Rightarrow \frac{w_1^2 - \sigma_\epsilon^2/\sigma^2}{2\int_0^1 w_s^2 ds} \tag{3}$$

where w_t is the standard Wiener process, and

$$\sigma_\epsilon^2 = \lim_{N\to\infty} N^{-1} \sum_{t=1}^{N} E(\epsilon_t^2) \tag{4}$$

$$\sigma^2 = \lim_{N\to\infty} N^{-1} \left(\sum_{t=1}^{N} \epsilon_t \right)^2 \tag{5}$$

Another commonly used test statistic is the t-ratio of the parameter in the regression (1), call it t. Under assumptions (AS1)-(AS4), and under H_0, one has for this statistic (see Phillips, 1987[13])

$$t \Rightarrow \frac{\sigma(w_1^2 - \sigma_\epsilon^2/\sigma^2)}{2\sigma_\epsilon\left(\int_0^1 w_s^2 ds\right)^{1/2}} \tag{6}$$

Both distributions were tabulated by Dickey and Fuller (1979 [7]). The test of unit roots based on the regression in (1), with either of the previous test statistics, is known as the *Dickey-Fuller* (DF) *test*. Although this test is not robust to serial correlation in the model error structure, it can be "augmented" by including sufficient lagged first differences Δx_{t-i}, $i \geq 1$, in the left-hand side of (1), so as to remove as much as possible of this correlation in ϵ_t. The resulting test is often referred to as the ADF test.

The second part of a standard cointegration test consists in checking for a unit root in the cointegration residuals $\hat{z}_t = y_t - \hat{\alpha} x_t$ (the hypothesis of a unit root is usually referred to as the null of non-cointegration), for which an estimate, $\hat{\alpha}$, of the cointegrating parameter, is required. It follows that the null distributions of the DF test statistics τ and t applied to the cointegration residuals are no longer those given above. These distributions have to be estimated, as it was done by Engle and Yoo (1991 [9]), because of the presence of a nuisance parameter.

One possible way to palliate this problem may be to state the null hypothesis as that of cointegration. In this case, the new null distribution of the test statistics will benefit from the *super-consistency* of the ordinary least-squares (OLS) estimator of this parameter, $\hat{\alpha}_{ols}$ (Stock, 1987 [14]). This property means that $N^{1-\delta}(\hat{\alpha}_{ols} - \alpha)$ for some very small δ ($< 1/2$), rather than $N^{1/2}(\hat{\alpha}_{ols} - \alpha)$, converges to a non-degenerate distribution. The tendency of the OLS estimator of α to converge at a faster rate than usual is intuitively explained by the fact that any perturbation to the true parameter, α, in the equation

$$y_t = \alpha x_t + z_t \tag{7}$$

will be severely penalized by the OLS estimation procedure, which searches for an optimal match between both sides of (7) at the second-order. Since we are dealing with integrated variables, the slightest mismatch in this parameter would lead to an imbalance of major properties between both sides of the regression equation.

Although most studies on cointegration rest on the assumption of a linear relationship between the variables, the possibility that these variables depend on each other through a jointly nonlinear Data Generating Mechanism (DGM), and which standard cointegration tests may fail to capture, has open the way to new research trends. Since the concept of cointegration is inherently nonlinear, early attempts focussed on extending standard definitions, and on understanding how the standard tests were affected by the presence of neglected nonlinearity. Hallman (1990 [12]) and Granger and Hallman (1991 [11]) proposed the concept of a cointegrating *nonlinear attractor* for a pair of univariate integrated series $x_t, y_t \sim I(d)$, by requiring the existence of nonlinear measurable functions $f(.), g(.)$ such that $f(x_t)$ and $g(y_t)$ are both $I(d)$, $d > 0$, and $s_t = f(x_t) - g(y_t)$ is $\sim I(d')$, with $d' < d$.

Assuming that f and g can be expanded as Taylor series up to some order $p \geq 2$ around the origin, we may write $s_t = c_0 + c_1 z_t + HOT(x_t, y_t)$, where $z_t = y_t - \alpha x_t$, and with $HOT(.,.)$ denoting higher-order terms. It follows that the linear approximation, z_t, to the true cointegration errors differs from the latter by some higher-order terms which express that the *strengh of attraction* onto the cointegration line $y_t = \alpha x_t$ varies with the levels of the series, x_t, y_t.

Also here, the case where $d = 1$, $d' = 0$ and the cointegration residuals have finite variance is most important in practice, since it allows a straighforward interpretation in terms of equilibrium concepts. Figure 1 illustrates the case of a cointegrating nonlinear attractor obtained by simulating a nonlinearly related pair of random walks with *i.i.d.* Gaussian errors.

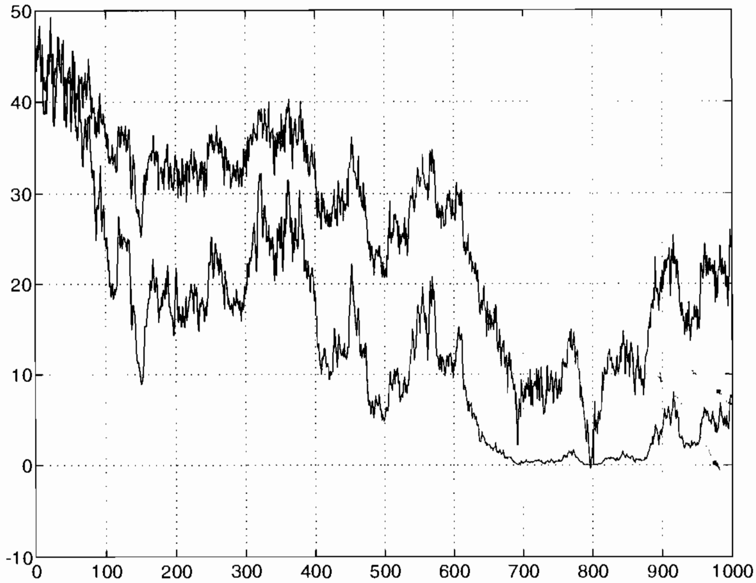

Figure 1: A cointegrating nonlinear attractor obtained with a factorial model. The upper series was generated as $x_t = w_t + \epsilon_{t,x}$, while the lower one corresponds to $y_t = g(w_t) + \epsilon_{t,y}$, where $g(.)$ represents a third-order polynomial of its argument random walk variable w_t, and $\epsilon_{t,x}, \epsilon_{t,y}$ are independent *i.i.d.* Gaussian sequences of random variables.

Hallman (1990[12]) also noted that DF tests were affected by the presence of nonlinearity in the series, and proposed running an ADF test on the *ranks* (RADF test), as a way to robustize this test against non-normality, and particularly against monotonic nonlinearity in their relationship (the rank of x_i in the sample of x_t, of size n, is defined as $r_{i,n}^{(x)} = \sum_{j=1}^{n} \mathbf{1}(x_i \geq x_j)$ -see for example David, 1981 [5]). However, a distribution for ranks is discrete, and thus different from a distribution of the levels. In spite that an approximation could always be found, Breitung and Gourieroux (1997 [4]) reported that any approximation to the asymptotic null rank distribution must differ considerably from the distributions tabulated by Dickey and Fuller (1979 [7]).

More recently, Aparicio and Granger (1994 [2]) proposed using *induced-order statistics* for characterizing cointegrating relationships. This new methodology presents some advantages. First, it allows us to free from the nuisance cointegration parameter the null distributions of the DF test statistics. Second, it allows us to construct a fully non-parametric cointegration testing procedure, which do not impose severe constraints on the time series models, such as the presence of unit roots. Third, it inherits the mild robustness of the ranks against monotonic nonlinearity in the relationship between the variables, circumventing the problems arising from a a null rank distribution. And forth, it is apparently robust to serial cointegration in the cointegration errors.

In this paper, we develop this methodology and provide simulation results supporting it.

2 Cointegrating relationships and induced-order statistics

For a sample of size n, $x_1, \cdots, x_n$, the order statistics of x_t are given by the sequence $x_{1,n} \leq \cdots \leq x_{n,n}$ obtained after a permutation, of the indexes $\{1, \cdots, n\}$ such that $x_{i,n} \leq x_{i+j,n}$, $\forall\, j > 0$. Related to order statistics are the so called *induced order statistics* (Bhattacharya, 1984[3]). The induced Y-order statistics based on the ordering of x_t are defined as: $\hat{y}_{i,n} = y_j$ if $x_{i,n} = x_j$. Obviously, $\hat{y}_{i,n} \neq y_{i,n}$, in general. It is known that the ranks are invariant with respect to order preserving transformations, such as monotonic nonlinear functions, and this property was used by Hallman (1990[12]) to increase the robustness of the DF test against monotonic nonlinear departures from the linear cointegration assumption. However, it was questionable whether the classical unit-root regression theory still applied when the variables had discrete probability distributions, as it is the case for rank variables. Breitung and Gourieroux (1997 [4]) reported some problems when applying the DF test to the ranks, and showed that the asymptotic null distribution is different in this case.

Induced-order statistics inherit the robustness properties of the DF test applied to the ranks against monotonic nonlinearities in the relationship, but have the advantage that the asymptotics of a test statistic constructed from the induced-order sample are easier to handle.

In the following we show how to obtain the order statistics of one series, y_t, induced by the orderings of the other, x_t, by using standard linear algebra, and define our test statistics. Let $\left\{\mathbf{P}_x^{(n)}, \mathbf{P}_y^{(n)}\right\}$ be sequences of stochastic permutation matrices, defined as

$$\begin{aligned} \mathbf{P}_x^{(n)} X &= X_{(o)} \\ \mathbf{P}_y^{(n)} Y &= Y_{(o)} \end{aligned} \tag{8}$$

where $Z = (z_1, \cdots, z_n)'$ and $Z_{(o)} = (z_{1,n}, \cdots, z_{n,n})'$, $Z = X, Y$. The vector of induced Y-order statistics (induced by the ordering of X), $\hat{Y} = (\hat{y}_{1,n}, \cdots, \hat{y}_{n,n})'$ is obtained as

$$\hat{Y} = \mathbf{P}_x^{(n)} Y \tag{9}$$

with $\tilde{Y} = (\tilde{y}_1, \cdots, \tilde{y}_n)'$. Now since $\mathbf{P}_y^{(n)}$ is invertible [1], we can form the sequence

$$\tilde{Y} = \left(\mathbf{P}_y^{(n)}\right)^{-1} \mathbf{P}_x^{(n)} Y, \tag{10}$$

In order to propose test statistics with no nuisance parameters in their null distributions, we consider now three candidate test statistics, which are computed directly on the series levels and not on the estimated residuals, thus leading to null distributions free of the nuisance cointegration parameter. Firstly, we consider the DF test statistic, τ, applied to the pair of univariate series $(y_t, \tilde{y}_t)$. Secondly, we propose two non-parametric statistics for comparing the orderings of the two time series y_t and $\hat{y}_t$. These are the *Kolmogorov-Smirnov* (K-S) (hereafter $K1$), and the χ^2 statistics (hereafter $K2$), and we use them as follows. Let $\hat{F}_Y^{(n)}(y)$ be the empirical distribution function obtained from a sample of length n of y_t. That is, $\hat{F}_Y^{(n)}(y) = n^{-1} \sum_{t=1}^{n} \mathbf{1}(y_t \leq y)$, where $\mathbf{1}(.)$ denotes the indicator function. Then one can define

$$K1 = \max_{j=1,n} |\hat{F}_Y^{(n)}(\hat{y}_{l_s(j),n}) - \hat{F}_Y^{(n)}(y_{j,n})| \tag{11}$$

[1] Note that $\mathbf{P}^{-1} = \mathbf{P}'$ for a permutation matrix, where $\mathbf{P}'$ denotes the transpose of $\mathbf{P}$.

and

$$K2 = n^{-1} \sum_{j=1}^{n} \frac{\left(\hat{F}_Y^{(n)}(\hat{y}_{l_s(j),n}) - \hat{F}_Y^{(n)}(y_{j,n})\right)^2}{\hat{F}_Y^{(n)}(y_{j,n})}, \tag{12}$$

where s takes the value of the sign of the estimated cointegration parameter, and $l_s(j)$ is equal to j or to $n-j$, depending on whether $s=1$ or $s=-1$ (respectively).
Large values of $K1$ or $K2$ suggest that the hypothesis of cointegration should be rejected.

Figure 2 shows the discrepancy between the sequences $\hat{F}_Y^{(n)}(y_{i,n})$ and $\hat{F}_Y^{(n)}(\hat{y}_{l_s(i),n})$. The K-S and χ^2 statistics, $K1$ and $K2$, respectively, are just a measure of the variability of the sequence $\hat{F}_Y^{(n)}(\hat{y}_{l_s(i),n})$ around the diagonal line, represented by the sequence $\hat{F}_Y^{(n)}(y_{i,n})$. In this figure, it is apparent the comparatively high variability of the $\hat{F}_Y^{(n)}(\hat{y}_{l_s(i),n})$ when the simulated unit-root time series have independent DGM's.

Table 1 shows the mean values and standard deviations (in brackets) of both statistics, $K1$ and $K2$, obtained from 100 replications of pairs of series that are cointegrated (a), of series that are related by a cointegrating quadratic attractor (b), and finally, of independent random walks. The sample size was $n = 1000$. It suggests that the alternative of non-cointegration could be easily isolated from the other two, because of the relatively large values adopted by the test statistics (especially $K2$).

Test statistic	*cointegration*	*quadratic attractor*	*independent random walks*
$K1$	0.259 (0.119)	0.156 (0.074)	0.773 (0.134)
$K2$	0.021 (0.04)	0.008 (0.025)	0.893 (0.743)

Table 1: Mean and standard deviations values of the Kolmogorov-Smirnov-type (K_1) and the χ^2-type (K_2) test statistics estimated from pairs of: (a) cointegrated series, (b) series related through a quadratic attractor, and (c) independent random walks.

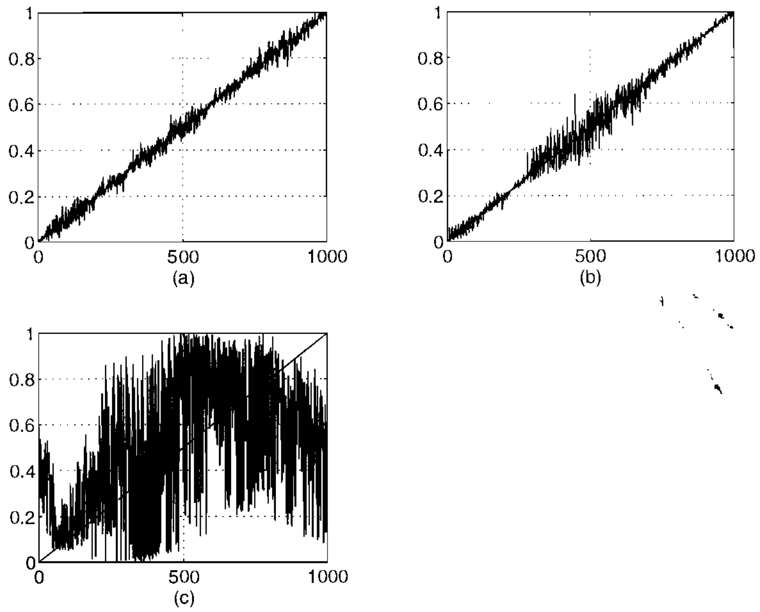

Figure 2: Illustration of the proposed non-parametric cointegration measures for pairs of: (a) cointegrated series, (b) series related through a quadratic atractor, and (c) independent random walks. In the plots, the straight diagonal line corresponds to the sequence $\hat{F}_Y^{(n)}(y_{i,n})$, while the superimposed trending noise corresponds to $\hat{F}_Y^{(n)}(\hat{y}_{l_s(i),n})$, where s stands for the sign of the cointegration parameter. The maximum absolute deviation or the standard deviation of this noise gives the proposed K-S or the χ^2 test statistics, respectively.

3 Some inference results

In relation with our discussion in the preceeding section, we can state the following theorem:

THEOREM TH2 :

A necessary and sufficient condition for two I(1) series x_t, y_t to be cointegrated (CI) is that the sequences $\tilde{y}_t, y_t$ are cointegrated with cointegration parameter equal to 1, that is, that $(\tilde{y}_t, y_t) \sim CI(1)$.

To prove this theorem, we invoked the following lemmas.

LEMMA LE4 :

Two I(1) time series x_t, y_t are cointegrated if and only if $x_{t,n}, y_{t,n}$ are cointegrated monotonically trending sequences with the same cointegration parameter.

PROOF OF LEMMA LE4 :

Let n be the sample size, and suppose that x_t, y_t are I(1) cointegrated series with cointegration parameter α, but $x_{t,n}, y_{t,n}$ are not. Let $Z_{(o)} = Y_{(o)} - \Sigma X_{(o)}$, where Σ is an n-dimensional diagonal matrix with diagonal elements equal to α. Then $n^{-3/2}\mathbf{1}'Z_{(o)} \Rightarrow \eta = \int_0^1 w_t dt$ ($\mathbf{1}'$ denotes an n-dimensional vector of ones), where w_t is the standard Wiener process. By continuity of η, $\forall\, \delta \in (0,1)\ \exists \xi$ such that $P(|\eta| > \xi) > 1 - \delta$, and thereby $P(n^{-1/2}|\mathbf{1}'Z_{(o)}| > b_n) \to 1 - \delta$, with $b_n = o(n)$. But $\mathbf{1}'Z_{(o)} = \mathbf{1}'Z$, with Z denoting the n-dimensional vector of cointegration residuals $z_t = y_t - \alpha x_t$. Thus the result contradicts the fact that z_t are ergodic. With the reverse sense of the proof, we may proceed similarly. $\square$.

PROOF OF THEOREM TH2 :

Let $\mathbf{\Sigma}$ be a diagonal matrix with nonzero elements equal to α, the cointegration parameter of the relation between x_t and y_t. Then one can write $\mathbf{P}_x^{(n)}Y = \mathbf{P}_x^{(n)}\mathbf{\Sigma}X + \mathbf{P}_x^{(n)}Z = \mathbf{\Sigma}\mathbf{P}_x^{(n)}X + \mathbf{P}_x^{(n)}Z$, with $Z = (z_1, \cdots, z_n)'$ and $z_t = y_t - \alpha x_t$. Some little algebra leads to $\mathbf{P}_y^{(n)}(\mathbf{P}_y^{(n)\prime}\mathbf{P}_x^{(n)} - \mathbf{I}^{(n)})Y = \mathbf{\Sigma}\mathbf{P}_x^{(n)}X - \mathbf{P}_y^{(n)}Y + \mathbf{P}_x^{(n)}Z = W$, where $W = (w_1, \cdots, w_n)'$. Now, from lemma LE4 it follows that w_t is I(0) and so will be the sequence v_t defined by

$V = \mathbf{P}_y^{(n)} W$. Therefore $\tilde{y}_t, y_t$ are cointegrated with cointegration parameter equal to 1. $\square$

REMARKS :

If x_t, y_t are cointegrated, so will be $y_t, \tilde{y}_t$, with a cointegration parameter equal to unity. As a consequence, the DF test statistic for testing cointegration has exactly the DF distribution under the hypothesis of non-cointegration (the cointegration parameter is no longer a nuisance as it need not be estimated).

Noting that $\hat{y}_{t,n} = y_{t,n} + e_{t,n}$, it can be seen that a test for cointegration amounts to minimizing near the spectral origin the stochastic matrix distance (using an appropriate norm)

$$||(\mathbf{P}_y^{(n)})^{-1}\mathbf{P}_x^{(n)} - \mathbf{I}^{(n)}|| \tag{13}$$

that arises in the regression equation

$$\left((\mathbf{P}_y^{(n)})^{-1}\mathbf{P}_x^{(n)} - \mathbf{I}^{(n)}\right) Y = U \tag{14}$$

where $\mathbf{I}^{(n)}$ is the $n \times n$ identity matrix, and $U = (u_1, \cdots, u_n)'$. Thus, we may say that cointegration implies a very small low-frequency component of the matrix on the left-hand side of the previous equation. We may think of this matrix as a nonlinear filter. Remark that although it operates linearly, this filter is inherently nonlinear since it depends on the given sample Y, through matrix $\mathbf{P}_y^{(n)}$. If the spectral support of this filter lies entirely in the high-frequencies, then it will preserve only the short-run behaviour of Y, and remove any existing long wave.

PROPOSITION PR2 :

A sufficient condition for the pair of $I(1)$ series x_t, y_t to be cointegrated is that $\mathbf{A}^{(n)} = (\mathbf{P}_y^{(n)})^{-1}\mathbf{P}_x^{(n)} - \mathbf{I}^{(n)}$ be a linear differencing (non-necessarily stationary) filter (i.e. high-pass filter).

PROOF :

By linearity of the filtering equation (14), we may rewrite it in the spectral domain as

$|F_A^{(n)}(\lambda)|^2 S_y^{(n)}(\lambda) = S_u^{(n)}(\lambda)$, with $F_A^{(n)}(\lambda)$ denoting the Fourier transform of the rows of $\mathbf{A}^{(n)}$. Now since $S_u^{(n)}(\lambda) \sim cte > 0$ and $S_u^{(n)}(\lambda) \sim \lambda^{-2}$, as $\lambda \to 0$, the spectral equation is balanced with $F_A^{(n)}(\lambda) \sim \lambda$ as $\lambda \to 0$. $\square$

REMARKS :

Suppose $x_t, y_t \sim I(1)$ series. Under cointegration, $U \sim I(0)$, and thus the variance of u_t as well as the sum of its autocorrelations will be finite and nonzero. This implies that $\mathbf{1}'E(UU')\mathbf{1} = n^2 c$, with $|c| < \infty$. As

$$c = n^{-1}\,\mathbf{1}'E\left[\left((\mathbf{P}_y^{(n)})^{-1}\mathbf{P}_x^{(n)} - \mathbf{I}^{(n)}\right) YY' \left((\mathbf{P}_y^{(n)})^{-1}\mathbf{P}_x^{(n)} - \mathbf{I}^{(n)}\right)'\right]\mathbf{1}, \tag{15}$$

it follows that the row vectors of $(\mathbf{P}_y^{(n)})^{-1}\mathbf{P}_x^{(n)} - \mathbf{I}^{(n)}$ should have high-frequency behaviour in order to offset the slow variations in the elements of $E(YY')$.

It is reasonable to think that induced-order statistics will be useful in testing for any sort of prominent low-frequency comovements, whether the individual series are integrated (even fractionally) or not. However, it is expected that they would fail in capturing relations among $I(0)$ time series, since the high-frequency fluctuations in the latter tend to randomize the relative orderings of the series. We will illustrate this fact with a simulation experiment in the next section.

4 Simulation experiments with the cointegration tests

In this section, we provide simulation evidence that that the two non-parametric test statistics $K1$ and $K2$ are useful for testing the null hypothesis of an equilibrium relationship between two integrated time series (cointegration), or in other words, of the stationarity of the errors $r_t = y_t - \tilde{y}_t$, where $\tilde{y}_t$ is the replica of y_t induced by the orderings of x_t.

For the purpose of comparison, we also consider the parametric DF test statistic applied on the pair of series $(y_t, \tilde{y}_t)$, even though this would be more suitable for testing the null of non-cointegration, since as we argued before, its distribution under this null is exactly the DF one.

Test statistic	*n=100*	*n=500*	*n=1000*
$\tau(y,\tilde{y};x)$	-8.440	-20.604	-29.875
$K1$	0.610	0.396	0.325
$K2$	0.171	0.0369	0.0193

Table 2: Empirical 5% right critical values of the test statistics $\tau(y,\tilde{y};x)$, $K1$ and $K2$, estimated from 1000 replications of linearly cointegrated series with *i.i.d.* Gaussian errors with zero mean and unit variance, and for different sample size n.

First of all, we estimate the 5% right critical values of the empirical distribution under the null hypothesis of cointegration of three alternative statistics, namely $\tau(y,\tilde{y};x)$, $K1$ and $K2$. Table 2 shows these values.

Next we obtained the power of the corresponding tests against the different alternative models given below. For our study, we considered three different sample sizes, $n = 100, 500, 1000$, and 1000 replications from each model.

1. *Model A1*

$$\Delta y_t = \epsilon_t \tag{16}$$

$$\Delta x_t = e_t \tag{17}$$

2. *Model A2*

$$y_t = \epsilon_t \tag{18}$$

$$x_t = e_t \tag{19}$$

3. *Model A3*

$$
\begin{aligned}
y_t &= 2.0x_t + \gamma\epsilon_t && (20)\\
x_t &= \beta x_{t-1} + e_t && (21)
\end{aligned}
$$

4. *Model A4*

$$
\begin{aligned}
y_t &= 2.0x_t + z_t && (22)\\
\Delta x_t &= e_t && (23)\\
z_t &= \beta z_{t-1} + \eta_t && (24)
\end{aligned}
$$

5. *Model A5*

$$
\begin{aligned}
y_t &= \alpha x_t + z_t && (25)\\
\Delta x_t &= e_t && (26)\\
z_t &= \epsilon_t - 0.5\epsilon_{t-1}, && (27)
\end{aligned}
$$

6. *Model A6*

$$
\begin{aligned}
y_t &= g(x_t) + \epsilon_t && (28)\\
\Delta x_t &= e_t && (29)\\
g(x) & \quad a_1 x + a_2 x^2 + a_3 x^3, && (30)
\end{aligned}
$$

and the following cases:

(a) $a_1 = 1.0, a_2 = 0.25, a_3 = 0.1,$

(b) $a_1 = 1.0, a_2 = 0.25, a_3 = 0.0,$

(c) $a_1 = 1.0, a_2 = 0.01, a_3 = 0.0,$

Test statistic	*n=100*	*n=500*	*n=1000*
$\tau(y, \tilde{y}; x)$	0.1995	0.672	0.84
$K1$	0.471	0.8195	0.931
$K2$	0.572	0.924	0.985

Table 3: Powers at the 5% significance level of the test statistics $\tau(y, \tilde{y}; x)$, $K1$ and $K2$ against the alternative of *model A1*, and for different sample size n.

Test statistic	*n=100*	*n=500*	*n=1000*
$\tau(y, \tilde{y}; x)$	0.047	0.030	0.039
$K1$	1.000	1.000	1.000
$K2$	1.000	1.000	1.000

Table 4: Powers at the 5% significance level of the test statistics $\tau(y, \tilde{y}; x)$, $K1$ and $K2$ against the alternative of *model A2*, and for different sample size n.

(d) $a_1 = 1.0, a_2 = 0.001, a_3 = 0.0,$

In the previous DGM's, e_t, η_t and ϵ_t represent three independent sequences of *i.i.d.* Gaussian random variables with zero mean and unit variance. The power results are shown in tables 3 to 8.

REMARKS :

Tables 3 and 4 show that, under independence of x_t and y_t the tests based on $K1$ and $K2$ are powerful (whether the series are $I(1)$ or $I(0)$). However, the one based on $\tau(y, \tilde{y}; x)$ is powerless when the series are $I(0)$. Table 5 shows that both $K1$ and $K2$ can discriminate between comovements in $I(0)$ series and cointegration in samples sizes of $n = 500$ and larger, in contrast to $\tau(y, \tilde{y}; x)$, which again has poor discrimination performance. Moreover, the signal-to-noise ratio has a clear impact on the power of $K1$ and $K2$. When the noise standard deviation is increased by a factor of 1000, both $K1$ and $K2$ have excellent power against comovements in $I(0)$ series, even for samples sizes as small as $n = 100$. Also the power

Test statistic	*n=100*	*n=500*	*n=1000*
$\tau(y, \tilde{y}; x)$	0.033 (0.047)	0.027 (0.032)	0.035
	0.047 (0.047)	0.024	0.046
	0.037 (0.042)	0.032	0.046
$K1$	0.000 (1.000)	0.083 (1.000)	0.721
	0.002 (1.000)	0.808	1.000
	0.006 (1.000)	1.000	1.000
$K2$	0.000 (1.000)	0.070 (1.000)	0.859
	0.000 (1.000)	0.933	1.000
	0.000 (1.000)	1.000	1.000

Table 5: Powers at the 5% significance level of the test statistics $\tau(y, \tilde{y}; x)$, $K1$ and $K2$ against *model A3*, for different sample size n. From top to bottom, the power values given for each statistic correspond to parameter values $\beta = 0.8$, $\beta = 0.5$ and $\beta = 0.0$. For most cases, we took $\gamma = 1.0$, except when the values are given in brackets (corresponding to $\beta = 0.8$), for which we took $\gamma = 1000$, in order to study the incidence of noise on the detection performance.

Test statistic	*n=100*	*n=500*	*n=1000*
$\tau(y,\tilde{y};x)$	0.081	0.094	0.076
	0.095	0.183	0.217
	0.15	0.457	0.637
$K1$	0.004	0.009	0.005
	0.019	0.038	0.04
	0.055	0.28	0.387
$K2$	0.008	0.008	0.005
	0.034	0.081	0.077
	0.124	0.422	0.617

Table 6: Powers at the 5% significance level of the test statistics $\tau(y,\tilde{y};x)$, $K1$ and $K2$ against *model A4*, for different sample size n. From top to bottom, the power values given for each statistic correspond to parameter values $\beta = 0.90$, $\beta = 0.95$ and $\beta = 0.99$.

Test statistic	*n=100*	*n=500*	*n=1000*
$\tau(y,\tilde{y};x)$	0.043	0.047	0.050
$K1$	0.061	0.072	0.073
$K2$	0.063	0.068	0.066

Table 7: Powers at the 5% significance level of the test statistics $\tau(y,\tilde{y};x)$, $K1$ and $K2$ against *model A5*, for different sample size n.

Test statistic	*n=100*	*n=500*	*n=1000*
$\tau(y, \tilde{y}; x)$	0.047	0.071	0.088
	0.391	0.557	0.598
	0.085	0.197	0.258
	0.050	0.041	0.046
$K1$	0.002	0.003	0.004
	0.049	0.140	0.203
	0.205	0.250	0.278
	0.058	0.058	0.049
$K2$	0.002	0.002	0.000
	0.044	0.147	0.213
	0.216	0.290	0.327
	0.058	0.055	0.055

Table 8: Powers at the 5% significance level of the test statistics $\tau(y, \tilde{y}; x)$, $K1$ and $K2$ against *model A6*, for different sample size n.

of these two tests seems to be an increasing function of the parameter β, which provides a measure of *memory* in x_t. We can provide an intuitive explanation for these results. First, remark that when the series are $I(0)$, the statistics $K1$ and $K2$ may not be able to detect any possible comovement, since the high-frequency fluctuactions of $I(0)$ series tend to shadow the trending pattern of the comovement, translating into a misalignement between the rank orderings of the two series. This behaviour of $K1$ and $K2$ differs remarkably from that of $\tau(y, \tilde{y}; x)$. By construction, the latter takes large and negative values both under the null of cointegration, and under the assumption of $I(0)$ series (either comoving or not). Indeed, consider the regression model $y_t = \alpha x_t + z_t$ where $y_t, x_t, z_t \sim I(0)$, and let $\hat{z}_t$ represent the regression residuals. Now, writing $r_t = (\rho - 1)r_{t-1} + \eta_t$, and noting that $r_t \sim I(0)$, one must have $\rho < 1$, and thus $\tau(y, \tilde{y}; x)$ which is proportional to $\hat{\rho} - 1$ must take negative values not close to zero. Therefore $\tau(y, \tilde{y}; x)$ is not suitable for discriminating between $I(0)$ comoving series and cointegrated series.

Tables 6 and 7 demonstrate that the three testing devices for the null of cointegration, are robust against correlation in the cointegration errors. Finally, robustness against mild non-linearities in the relationship is studied in table 8, which shows that the tests $K1$ and $K2$ have little or no power against the forms of nonlinearity considered.

All things considered, we conclude that test statistic $\tau(y, \tilde{y}; x)$, derived from the parametric DF test statistic τ should rather be used for testing the null of non-cointegration, for which it would benefit from the absence of nuisance parameters in its null distribution. Recall that if we had taken the null hypothesis as that of non-cointegration, as it is done usually, then the null distribution of $\tau(y, \tilde{y}; x)$ would be exactly the DF distribution.

5 Unit roots and induced-order statistics

In this section we study the implications of unit roots in time series from the standpoint of induced-order statistics. We also present a unit-root testing device based on the latter.
The key idea is to remark that if $x_t \sim I(1)$ then the pair $(x_t, x_{t-\tau})$ is cointegrated for τ small, and thus, $(x_{t-\tau}, \tilde{x}_{t-\tau}) \sim CI(1)$, for small τ, and where $\tilde{x}_{t-\tau}$ denotes the replica of $x_{t-\tau}$ induced by the orderings of x_t. The heuristics underlying this idea are that, for a $I(1)$ series, a delayed replica of itself must have similar trending patterns if the delay τ is small, whereas if the series is $I(0)$ then any delayed replica will exhibit, in general, very different sample paths (even for a small delay). Thus the regression of x_t on $x_{t-\tau}$ will lead in this latter case to a regression parameter estimate close to zero in a statistical sense.

Thus we may also define an integrated time series x_t as one for which the pair $(\tilde{x}_{t-\tau}, x_{t-\tau}) \sim CI(1)$ for small τ. Since the smaller the delay τ the better our characterization, henceforth, we will assume in our definition $\tau = 1$. Thus we will say that $x_t \sim I(1)$ if and only if $(x_{t-1}, \tilde{x}_{t-1}) \sim CI(1)$, where $\tilde{x}_{t-1}$ is obtained from the following sequence of transformations:

$$X_{(o)} = \mathbf{P}_x^{(n-1)} X \tag{31}$$

$$\hat{X}^{(-1)} = \mathbf{P}_x^{(n-1)} X^{(-1)} \tag{32}$$

$$\tilde{X}^{(-1)} = \mathbf{P}_{x^{(-1)}}^{(n-1)'} \hat{X}^{(-1)}, \tag{33}$$

where $X = (x_2, \cdots, x_n)'$, $X^{(-1)} = (x_1, \cdots, x_{n-1})'$, and $\mathbf{P}_{x^{(-1)}}^{(n-1)}$ is defined by the relation $X_{(o)}^{(-1)} = \mathbf{P}_{x^{(-1)}}^{(n-1)} X^{(-1)}$.
Therefore if $x_t \sim I(1)$ then we have:

$$(I^{(n-1)} - \mathbf{P}_{x^{(-1)}}^{(n-1)'} \mathbf{P}_x^{(n-1)}) X^{(-1)} \sim I(0). \tag{34}$$

In other words, the $(n-1) \times (n-1)$ matrix $\mathbf{A}^{(n-1)} = I^{(n-1)} - \mathbf{P}_{x^{(-1)}}^{(n-1)'} \bar{\mathbf{P}}_x^{(n-1)}$ must have the properties of a differencing (high-pass) filter or linear operator. This result has implications as to the nature of the operations performed by the matrix $\mathbf{B}^{(n-1)} = \mathbf{P}_{x^{(-1)}}^{(n-1)'} \mathbf{P}_x^{(n-1)}$. For example, suppose we have:

$$\mathbf{A}^{(n-1)} = \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 \\ 0 & 1 & -1 & \cdots & 0 \\ \vdots & \cdots & \cdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 & -1 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

then

$$\mathbf{B}^{(n-1)} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \cdots & \cdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 0 & 0, \end{pmatrix}$$

which represents a one-step-forward shifting operator. Now, since $\mathbf{P}_x$ is a permutation matrix, we have $\mathbf{P}'_x\mathbf{P}_x = I$ and $\mathbf{P}_x\mathbf{P}'_x = I$. Now, writing $\mathbf{B} = I + \Delta I$, it is easy to obtain that $\mathbf{P}'_{x(-1)} = \mathbf{B}\mathbf{P}'_x$. That is, $\mathbf{P}_{x(-1)}$ results from shifting $\mathbf{P}_x$ to the left by one element.

Finally, if $x_t \sim I(0)$ then $\mathbf{B} \approx \mathbf{0}$, that is, $\mathbf{P}_{x(-1)}$ and $\mathbf{P}_x$ will be quasi-orthogonal, which means that X and $X^{(-1)}$ have completely different orderings.

6 Simulation experiments with the unit-root tests

Now we show the simulation results obtained in our analysis of unit-roots using induced-order statistics. Our test statistics were applied to the pair of series $(x_{t-1}, \tilde{x}_{t-1})$, where $\tilde{x}_{t-1}$ denotes the induced-order replica of x_{t-1} from the orderings of x_t.

In table 9 below, we show the empirical 5% right critical values for the DF test statistic applied to the pair $(x_{t-1}, \tilde{x}_{t-1})$, that is $\tau(x_{-1}, \tilde{x}_{-1}; x)$, and for our test statistics, $K1$ and $K2$. These critical values were estimated from 1000 simulations of a random walk with *i.i.d.* Gaussian errors with zero mean and unit variance, and for different sample sizes $(n = 100, 500, 1000)$.

Test statistic	*n=100*	*n=500*	*n=1000*
$\tau(x_{-1}, \tilde{x}_{-1}; x)$	-8.4080	-20.7345	-30.0459
$K1$	0.4848	0.3106	0.2442
$K2$	0.0994	0.0196	0.0097

Table 9: Empirical 5% right critical values of the test statistics $\tau(x_{-1}, \tilde{x}_{-1}; x)$, $K1$ and $K2$, estimated from 1000 replications of a random walk with *i.i.d.* Gaussian errors, and for different sample size n.

Now the powers against the alternative hypothesis $H_1 : x_t \sim I(0)$ are shown in table 10 for the following two models:

1. *Model B1*

$$x_t = e_t, \tag{35}$$

2. *Model B2*

$$x_t - 0.8x_{t-1} = e_t, \tag{36}$$

where e_t represents once again a sequence of *i.i.d.* Gaussian random variables, with zero mean and unit variance.

For the same reason as explained in a previous section, $\tau(x_{-1}, \tilde{x}_{-1}; x)$ is not useful in detecting $I(0)$ patterns, since its behaviour is similar for both $I(1)$ and $I(0)$ series (it takes large negative values).

Finally, we computed the power of the tests based on the test statistics $K1$ and $K2$ for the models:

1. *Model B3*

Test statistic	n=100	n=500	n=1000
$\tau(x_{-1}, \tilde{x}_{-1}; x)$	0.018	0.020	0.021
	0.033	0.033	0.034
$K1$	1.000	1.000	1.000
	0.841	1.000	1.000
$K2$	1.000	1.000	1.000
	0.846	1.000	1.000

Table 10: Powers at the 5% significance level of the test statistics $\tau(x_{-1}, \tilde{x}_{-1}; x)$, $K1$ and $K2$ for *model B1* (top) and *model B2* (bottom). These powers were estimated from 1000 replications of the two models and for different sample size n.

$$x_t = g(y_t) + e_t \tag{37}$$

$$\Delta y_t = \xi_t \tag{38}$$

with $g(y) = y + 0.25y^2$.

2. *Model B4*

$$\Delta x_t = v_t \tag{39}$$

$$v_t - 0.5v_{t-1} = e_t, \tag{40}$$

where e_t and ξ_t are independent sequences of *i.i.d.* of Gaussian random variables, with zero mean and unit variance. The power results are given in table 11.

The previous power results show that a unit-root test based on either of the two non-parametric statistics considered, is mildly robust to the quadratic nonlinearity, but its size is severely altered by the presence of serial correlation in the model errors.

Test statistic	*n=100*	*n=500*	*n=1000*
$K1$	0.090	0.088	0.118
	0.607	0.658	0.694
$K2$	0.109 (0.654)	0.192 (0.811)	0.257 (0.851)
	0.654	0.811	0.851

Table 11: Powers at the 5% significance level of the test statistics $K1$ and $K2$ for *model B3* (top) and *model B4* (bottom). These powers were estimated from 1000 replications of the two models and for different sample size n.

7 Conclusion

In this paper, we presented induced-order statistics in the context of characterizing cointegrating relationships and integrated series. Based on this characterization, we proposed some cointegration testing devices which do not require prior estimation of the cointegration parameter, and therefore lead to null distributions which are free from nuisance parameters. We proposed using the parametric DF test statistic on the pair of series $(y_t, \tilde{y}_t)$, where $\tilde{y}_t$ is the induced-ordered replica of y_t following the ranks of x_t, to test the null of non-cointegration, since its distribution under this null is exactly the DF distribution. We also proposed a pair of non-parametric test statistics for testing the null hypothesis of stationarity of the errors in the regression of two $I(1)$ time series. These tests are powerful against the wide alternative of non-cointegration (including possibly comoving $I(0)$ series). Therefore they do not require prior testing for unit roots in the individual series.

References

[1] F.M. Aparicio Acosta. *Nonlinear modelling and analysis under long-range dependence with an application to positive time series.* PhD thesis, Swiss Federal Institute of Technology (EPFL) -Signal Processing Laboratory-, Lausanne, Switzerland, 1995.

[2] F.M. Aparicio Acosta and C.W.J. Granger. Nonlinear cointegration and some new tests for comovements. Working Paper (Depart. of Economics of the University of California at San Diego), March 1995.

[3] P.K. Bhattacharya. Induced order statistics: theory and applications. In P.R. Krishnaiah and P.K. Sen, editors, *Handbook of statistics*, volume 4, pages 383–403. Elsevier Science Publishers, Amsterdam, 1984.

[4] J. Breitung and C. Gourieroux. Rank tests for unit roots. *Journal of econometrics*, 81:7–27, 1997.

[5] H.A. David. *Order statistics.* John Wiley, N.Y., 1981.

[6] D.A. Dickey, W.R. Bell, and R.B. Miller. Unit roots in time series models: tests and implications. *The American statistician*, 40(1):12–26, 1986.

[7] D.A. Dickey and W.A. Fuller. Distribution of the estimators for AR time series with a unit root. *Journal of the American stastistical association*, 74(366):427–431, 1979.

[8] R.F. Engle and C.W.J. Granger. Cointegration and error correction: representation, estimation and testing. *Econometrica*, 55:251–276, 1987.

[9] R.F. Engle and B.S. Yoo. Forecasting and testing in cointegrated systems. In R.F. Engle and C.W.J. Granger, editors, *Long-run economic relationships*, pages 114–129. Oxford Science Publications, 1991.

[10] C.W.J. Granger. Cointegrated variables and error-correcting models. Discussion paper 83-13 (Depart. of Economics of the University of California at San Diego), 1983.

[11] C.W.J. Granger and Hallman J.J. Long-memory series with attractors. *Oxford Bulletin of Economics and Statistics*, 53(1):11–26, 1991.

[12] J.J. Hallman. *Nonlinear integrated series, cointegration, and an application.* PhD thesis, Department of Economics of the University of California at San Diego, La Jolla, USA, 1990.

[13] P.C.B. Phillips. Time series regression with a unit root. *Econometrica*, 55:277–301, 1987.

[14] J.H. Stock. Asymptotic properties of LS estimators of cointegrating vectors. *Econometrica*, 55(5):1035–1056, September 1987.