Nonlinear Cointegration and Nonlinear Error Correction: Record Counting Cointegration Tests

Alvaro Escribano, Ana E. Sipols and Felipe Aparicio
Department of Economics, Statistics& Statistics
Universidad Carlos III, Universidad Rey Juan Carlos& Universidad Carlos III
anaelizabeth.garcia@urjc.es

Key Words: Unit roots tests; cointegration; nonlinearity; robustness; ranges; jumps; counting statistics; small sample corrections; structural breaks.

ABSTRACT

In this paper we propose a record counting cointegration (RCC) test which is robust to nonlinearities and certain types of structural breaks. The RCC test is based on the synchronicity property of the jumps (or new records) of cointegrated series, counting the number of jumps that simultaneously occur in both series. We obtain the rate of convergence of the RCC statistics under the null and alternative hypothesis. The distribution of RCC under the null of a unit root depends on the short run dependence of the cointegrated series. We propose a small sample correction and show by Monte Carlo simulation techniques their excellent small sample behaviour. Finally we apply our new cointegration test statistic to several financial and macroeconomic time series that have some structural breaks and nonlinearities.

1. INTRODUCTION

Granger (1981) introduced the concept of cointegration and with the contribution of Engle and Granger (1987) and Johansen (1991) this concept has achieved immense popularity among econometricians and applied economists. Only a few recent papers have been dedicated to the simultaneous consideration of nonstationarity and nonlinearity, even though many people agree that those are likely characteristics of many macroeconomic and financial economic relationships. Granger (1995) discussed the concepts of long-range dependence in mean and extended memory which generalize the linear concept of integration, I(1), to a nonlinear framework. On the other hand, there are interesting empirical macroeconomic applications where nonlinearity has been found in a nonstationary context and therefore, there is a need to justify those results econometrically.

Most unit root tests, like Dickey and Fuller (1979) or Phillips and Perron (1988), are not robust to outliers, Franses and Haldrup (1994), nor to structural breaks Perron (1990), nor to nonlinear transformations Granger and Hallman (1991) and Aparicio, Escribano and Garcia (2006b). Therefore, tests for non-cointegration based on the augmented Dickey and Fuller (ADF) test applied to the residuals of the cointegrating relationship, Engle and Granger (1987), have size distortions and losses in power. Aparicio, Escribano and Garcia (2004,2006b) analyzed the asymptotic properties of a new range unit root (RUR) test and provide evidence of their nice behavior by Monte Carlo simulation of nonlinearities and structural breaks and by some empirical applications.

In this paper we analyzed the properties of a record counting cointegration (RCC) test which is robust to monotonic nonlinear transformations and structural changes. This testing procedure works with the ranges, instead of using the actual variables. The range for a given time t is defined as the difference between the cumulative maximum and the cumulative minimum at that time t. The first differences of the ranges are called the jumps and they represent the arrival of a new maximum or minimum (new record). The RCC test analyzed in this paper is based on counting the synchronicity of the jumps (new maximum or minimum) of the cointegrated series. This statistic counts the number of jumps that simultaneously occur in both series. We compare the properties of the well known non-cointegration test, see Engle and Granger (EG) (1987), with those of RCC. We show by Monte Carlo simulations how EG dramatically fails in the presence of nonlinearities and structural breaks. These well known results are general and affect most of the available tests of cointegration or non-cointegration, like Johansen (1991), etc.

As empirical applications we have considered the prices of gold and silver, analyzed by Escribano and Granger (1998) and the UK money demand from 1878 to 2000, analyzed by Escribano (2004). Those data sets were selected because there was evidence that the series were cointegrated only after explicit treatments of nonlinearities and/or structural breaks. However, in this paper we find evidence of cointegration without any previous treatment (pre-filtering) of those problems.

2. ANALYSIS BASED ON RECORD COUNTING COINTEGRATION (RCC)

In this section we introduce nonparametric methods to analyze cointegration that do not impose restrictions on the functional form relating the variables. Some of those procedures are robust to certain types of structural breaks (or level shifts) and monotonic nonlinear transformations.

Aparicio (1995), Aparicio and Granger (1995) and Aparicio, Escribano and Garcia (2006a) propose an alternative nonparametric methodology to test for unit roots. The basic idea behind this approach is to *count recursively the number of new records* (maximum or minimum) in a time series. We expect nonstationary cointegrated series to have many more new co-records (or synchronous new records) than noncointegrated series or stationary series. To be more precise we need to introduce some concepts and definitions that are useful in the rest of the paper.

Definition 1. *Given a time series* x_t *we define the sequence of extremes of* x_t *as the sequence of* $x_{1,t} = \min\{x_1, ..., x_t\}$ *and* $x_{t,t} = \max\{x_1, ..., x_t\}$, *when* $t = 1, 2, \ldots, n$.

Definition 2. *Let* n *be the sample size, the sequence of ranges of* x_t *is defined as:*

$$R_t^{(x)} = x_{t,t} - x_{1,t} \qquad for\ t = 1, ..., n \tag{1}$$

Definition 3. *Given the sequence of ranges we can define the sequence of jumps (new records) as the sequence of first differences of ranges:*

$$\Delta R_t^{(x)} = R_t^{(x)} - R_{t-1}^{(x)} \tag{2}$$

The series of new records $\Delta R_t^{(x)}$ is positive or zero. When there is a new record (maximum or minimum) in $\{x_1, ..., x_t\}$, the first difference of the ranges will be positive at time t,

$\Delta R_t^{(x)} > 0$. A statistics record was proposed by Aparicio, Escribano and Garcia (2006b) for robust unit root testing. When the original series is stationary, with finite variance, the series of first differences of the ranges are positive at the beginning of the sample and the rest is almost a sequence of zeros. On the contrary when the series are nonstationary $I(1)$ new records appear with positive probability as the sample increases. The key idea relied on the different vanishing rates of the long-run frequency of a new record, $n^{-1}\sum_{t=1}^{n}\mathbf{1}(\Delta R_t^{(x)} > 0)$, for an $I(1)$ and an $I(0)$ time series, in such a way that the normalized long-run frequency of records

$$J_x^{(n)} = n^{-1/2}\sum_{t=1}^{n}\mathbf{1}(\Delta R_t^{(x)} > 0) \tag{3}$$

converged in probability to zero under the alternative of stationary, and to a non degenerate positive random variable under the null hypothesis of a unit root. A well-known result from *extremal theory* is that the statistical properties of records from *iid* sequences of random variables are shared by a wide class of dependent stationary time series (see for instance, Lindgren and Rootzén (1987) and Leadbetter and Rootzén (1988). This prompts the question of whether short run dependencies and cross-dependencies may have an impact or not on a record-based test for cointegration. Then, one important alternative model to consider in this context is the case of dependent random walks.

For plots of the sequence of ranges $R_t^{(y)}$ y $R_t^{(x)}$ see Aparicio, Escribano and Garcia (2006a). The following models are useful to motivate the new approach we are proposing. with parameter values equal to $\alpha = 0.5$, $b = 0.6$, where $e_{t,0}$, $e_{t,1}$, $e_{t,2}$ are standard normal distributions and mutually independent, Nid(0,1).

Model 1 (Linear cointegration): $x_t = w_t + e_{t,1}$, $y_t = \alpha w_t + e_{t,2}$ where $w_t = w_{t-1} + e_{t,0}$.

Model 2 (Nonlinear cointegration): $x_t = w_t + e_{t,1}$, $y_t = \alpha w_t^2 + e_{t,2}$ where $w_t = w_{t-1} + e_{t,0}$.

Model 3 (Independent random walks): $x_t = x_{t-1} + e_{t,1}$, $y_t = y_{t-1} + e_{t,2}$.

Model 4 (Random walks with short run dependence ($a \neq 0$)): $\Delta x_t = e_{t,1}$, $\Delta y_t = a\Delta x_t + e_{t,2}$.

Figure 1 shows the cross plots of the sequences of new records or first difference of ranges. From the cross plot of Figure 1 it is clear that dependent but not cointegrated random walks (d) have several new records synchronized indicating some small number of co-records. This property will likely reduce the power of the RCC test if the short run correlation is high relative to cointegration relationship. The idea that cointegrated series imply arrivals of highly synchronized new records , suggest the following nonparametric test statistic that we called *the record counting cointegration test,* (RCC)

$$RCC_{x,y}^{(n)} = \frac{\sum_{t=1}^{n}\mathbf{1}(\Delta R_t^{(x)} > 0)\mathbf{1}(\Delta R_t^{(y)} > 0)}{\log(n)} \tag{4}$$

where the product of the two indicator functions $\mathbf{1}(\Delta R_t^{(x)} > 0)\mathbf{1}(\Delta R_t^{(y)} > 0)$ count the number of arrivals of new records that are coincident or synchronized *(co-records).* That is,

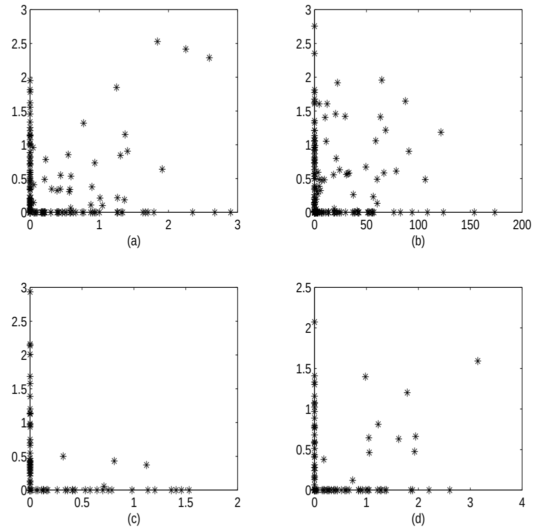

Figure 1: Cross-plots of $\Delta R_t^{(y)}$ and $\Delta R_t^{(x)}$ for pairs of series that are: (a) linearly cointegrated (model 1), (b) nonlinearly cointegrated (model 2), (c) independent random walks (model 3), and (d) random walks with short run dependence (model 4)

the RCC counts how many of the total arrivals of the jumps coincide, relative to the $log(n)$. Therefore, the convergence rates of the test statistic RC is $log(n)$.

Theorem 1 *Let the processes* $x_t = \sum_{i=1}^{t} e_{i,1}$ *and* $y_t = \sum_{i=1}^{t} e_{i,2}$ *for* $t = 1, 2, ..., \infty$*,where* $e_{i,1}$ *and* $e_{i,2}$ *are independent continuous zero-mean iid sequences with finite variances and symmetric pdf. Let* $RCC_{x,y}^{(n)}$ *be the number of joint records of* x_t *and* y_t *in a sample of size* n (4), *then:*

$$\log(n)^{-1} RCC_{x,y}^{(n)} \to 1 \tag{5}$$

$$E\left\{RCC_{x,y}^{(n)}\right\} = O(\log n) \tag{6}$$

$$Var\left\{RCC_{x,y}^{(n)}\right\} = O(\log n) \tag{7}$$

Proof. See Appendix A1. ■

Consistency

If x_t and y_t are cointegrated then for some $a \neq 0$ there exists an $I(0)$ sequence, η_t such that $y_t = ax_t + e_{t,2}$. Since for large t, x_t will dominate $e_{t,2}$, we can write:

$$RCC_{x,y}^{(n)} = \frac{\sum_{t=1}^{n} \mathbf{1}(\Delta R_t^{(x)} > 0)\mathbf{1}(\Delta R_t^{(y)} > 0)}{\log(n)} \simeq \frac{\sum_{t=1}^{n} \mathbf{1}(\Delta R_t^{(x)} > 0)}{\log(n)} \tag{8}$$

But we know from Aparicio, Escribano and Garcia (2006b) that $\sum_{t=1}^{n} \mathbf{1}(\Delta R_t^{(x)} > 0) = O(n^{1/2})$. Thus under the alternative hypothesis of cointegration, the test statistic will satisfy:

$$(\log n)^{-1} RCC_{x,y}^{(n)} \to \infty \tag{9}$$

Invariance against Monotonic Nonlinearities

Monotonic transformations preserve the ordering of the observations in any time series, and thus the *inter-record times*. As a consequence, if we let $f(.)$ and $g(.)$ be a monotonic nonlinear transformations, we must have:

$$RCC_{f(x),g(y)}^{(n)} = RCC_{x,y}^{(n)} \tag{10}$$

More generally, let x_t and y_t be $I(1)$ time series variables, and let $x'_t = f(y_t) + \varepsilon_t$, $y'_t = g(x_t) + \eta_t$, where $\{\varepsilon_t\}_{t\geq 1}, \{\eta_t\}_{t\geq 1}$ are independent *iid* sequences with zero mean and finite variances. Since for large values of t the nonlinear transformations, $f(y_t)$ y $g(x_t)$ dominate ε_t and η_t respectively, the records of x'_t (y'_t) will occur at the same time as x_t (y_t). As a consequence, the count of new records before and after the transformation should be the same. That is,

$$RCC_{(x',y')}^{(n)} = \sum_{t=1}^{n} \mathbf{1}(\Delta R_t^{(x')} > 0)\mathbf{1}(\Delta R_t^{(y')} > 0) \tag{11}$$

$$\simeq RCC_{(x,y)} = \sum_{t=1}^{n} \mathbf{1}(\Delta R_t^{(x)} > 0)\mathbf{1}(\Delta R_t^{(y)} > 0) \tag{12}$$

In finite samples, the actual size will oscillate around the nominal size depending on the type of transformations. For example, certain kinds of transformations can emphasize the $I(1)$ part over the $I(0)$ part. This feature may lead, in finite samples, to size fluctuations around the nominal one.

Small Sample Performance of the RCC Test: Monte Carlo Simulations

Let the data generation process (DGP) be the following *non-cointegrated random walks* but with short run dependence ($a \neq 0$), $H_0 : \Delta x_t = e_{t,1}$, $\Delta y_t = a\Delta x_t + e_{t,2}$

In Table I we estimate the quintals of the empirical distribution of the RCC test statistic, under H_0, for different parameter values of the short run dependence ($a = 0, 0.5, 1, 1.5$), different sample sizes $n = 100, 250, 500, 1000$ and different significance values v. The estimated quintals, based on 10,000 Monte Carlo simulations, are shown in Table I. We observe that for most sample sizes (n) the quintals increase with the short run dependence (a) in small sample sizes and therefore the empirical distribution of RCC is shifted to the right. This simulation results call for the need of a small sample correction of the RCC statistic to make it useful, since in empirical application we do not know the true value of the parameter a. One possibility is to prefilter the series (use of instrumental variables, etc.) but this could be complicated if the dependence is nonlinear. A better alternative is to divide the RCC by the number of new records that are synchronized between the first differences of the series. We call this nonparametric statistic the RCC corrected dependence:

$$RCC_{CD} = \frac{\sum_{t=1}^{n} \mathbf{1}(\Delta R_t^{(x)} > 0)\mathbf{1}(\Delta R_t^{(y)} > 0)}{\log(n) \sum_{t=1}^{n} \mathbf{1}(\Delta R_t^{(\Delta x)} > 0)\mathbf{1}(\Delta R_t^{(\Delta y)} > 0)} \tag{13}$$

Table I: Quintals of the empirical distribution of RCC for dependent, but not cointegrated, random walks, for different values of a and different sample sizes n

		a=0					a=1	
v/n	100	250	500	1000	100	250	500	1000
0.01	0.21	0.18	0.16	0.18	0.43	0.54	0.64	0.71
0.025	0.21	0.18	0.32	0.33	0.65	0.72	0.96	1.12
0.05	0.43	0.36	0.32	0.34	0.86	0.90	1.12	1.31
0.10	0.53	0.54	0.48	0.52	1.86	2.08	2.28	2.42
0.90	2.30	2.08	2.28	2.31	3.17	3.53	4.05	4.51
0.95	3.11	3.16	3.11	3.12	3.38	3.89	4.37	4.71

In Table II we estimate the quintals of the empirical distribution of the RCC_{CD} test statistic, under H_0, for different parameter values of the short run dependence ($a = 0, 1$), different sample sizes $n = 100, 250, 500, 1000$ and different significance values v based on 10,000 Monte Carlo simulations. The asymptotic distribution of the RCC is a good approximation to the distribution of the RCC in sample sizes of moderate size. We can observe the invariance that was obtained with the small sample correction of the RCC statistic for different values of a and different sample sizes n. Figure 5 and 6 show the nonparametric estimates of the density function of RCC and RCC_{CD} test statistics under the null hypothesis of independent random walks ($a = 0$), with $Nid(0, 1)$ errors and $a = 1$, 1,000 replications were done using the Epanechnikov kernel.

Table II: Quintals of the empirical distribution of RCC_{CD} for dependent, but not cointegrated, random walks, for different values of a and different sample sizes n

RCC_{CD}			a=0				a=1	
v/n	100	250	500	1000	100	250	500	1000
0.01	0.21	0.231	0.23	0.24	0.22	0.24	0.242	0.24
0.025	0.21	0.232	0.24	0.25	0.23	0.232	0.25	0.26
0.05	0.45	0.38	0.35	0.378	0.41	0.33	0.332	0.376
0.10	0.56	0.58	0.59	0.60	0.53	0.58	0.59	0.60
0.90	2.3	2.24	2.28	2.3	2.4	2.51	2.49	2.52
0.95	3.32	3.27	3.42	3.42	3.53	3.54	3.57	3.6

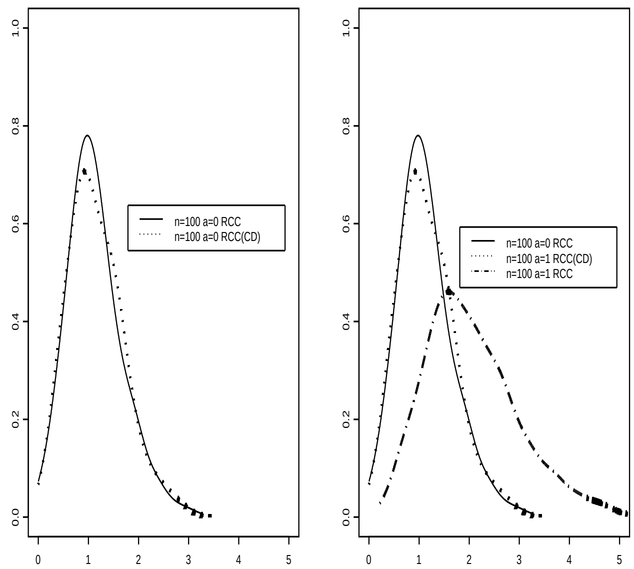

Figure 2: Nonparametric kernel estimates of the density function of the Tests Statistic RCC_{CD} and RCC, under the null hypothesis of no cointegration, for $n = 100$ and for $a = 0, 1$

3. SIZE ADJUSTED POWER OF THE RCC TESTS: MONTE CARLO EVIDENCE

We will analyze the power of the RCC tests using the 5% right tail critical value. The DGP under the alternative hypothesis of cointegration is generated by a bivariate vector error correction model with weakly exogenous variables for the cointegrating parameter vector.

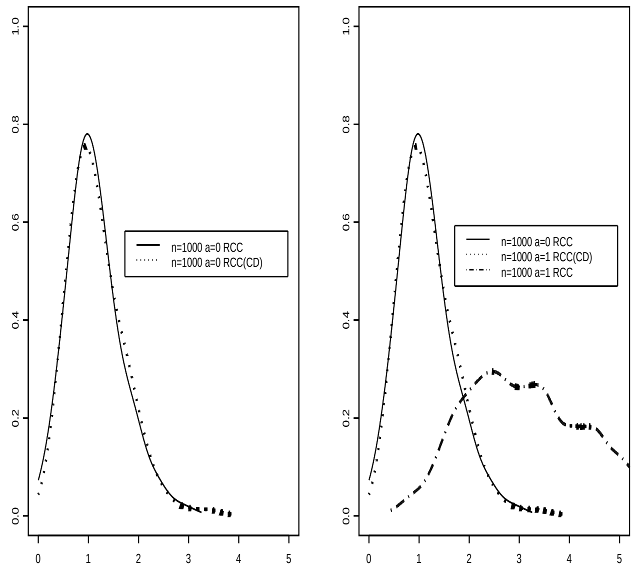

Figure 3: Nonparametric kernel estimates of the density function of the RCC_{CD} and RCC,0ests Statistic under the null hypothesis of no cointegration, for
$n = 1,000$ and for $a = 0, 1$

Consider the following restricted VAR model, for the (y_t, x_t) vector which is generated by,

$$\begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \Delta y_t \\ \Delta x_t \end{bmatrix} = \begin{bmatrix} c \\ 0 \end{bmatrix} + \begin{bmatrix} b \\ 0 \end{bmatrix} (1, -\alpha) \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} w_{1t} \\ w_{2t} \end{bmatrix} \tag{14}$$

where $w_t \sim Nid\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right)$

Linear Cointegration

H_1**:** The alternative hypothesis is a standard linear error correction model (ECM)

DGP**:** $\Delta y_t = c + a\Delta x_t + b(y_{t-1} - \alpha x_{t-1}) + w_{1t} \qquad \Delta x_t = w_{2t}$

where $\alpha = 1$ and $c = 0$.

We will study the power of different non-cointegration tests: the RCC, and the Engle-Granger (EG)(1987), for different parameter values of the parameter (b), say $b = 0$ (non-cointegration), $b = -0.01$, -0.1, -0.25, -0.75, (cointegration) and different short run dependence given by $a = 0, 1$. This DGP follows the parameterization used by Kremers et al. (1992), and Arranz and Escribano(2001). The results of EG are in parenthesis in Table III.

Table III: Power of the RCC and EG Tests for different parameter values of b, a and n

b (EG)		-0.01			-0.1	
$a \backslash n$	100	250	500	100	250	500
0	0.4 (0.047)	0.4 (0.0920)	0.5 (0.1590)	0.5 (0.38)	0.6 (0.99)	0.7 (1)
1	0.32 (0.09)	0.5 (0.12)	0.51 (0.21)	0.44 (0.49)	0.7 (0.99)	0.84 (1)
$a \backslash b$		-0.25			-0.75	
0	0.6 (0.99)	0.8 (1)	0.81 (1)	0.7 (1)	0.9 (1)	0.91 (1)
1	0.54 (0.99)	0.7 (1)	0.75 (1)	0.6 (1)	0.6 (1)	0.7 (1)

Since the EG test is based on estimating the true error correction model (DGP), the EG test for non-cointegration is more powerful than the RCC. However, the RCC is more powerful for small values of b, slow error correction EC adjustment, (see for example $a = 0$, $b = -0.01$ in Table III). Notice that the power of the RCC decreases the larger is a.

Nonlinear cointegration and nonlinear error correction

Escribano (1986, 2004) analyzes ECM models in nonlinear contexts where the cointegration relationship can be linear or nonlinear and the equilibrium correction term can also be linear or nonlinear. He proposes alternative representation theorems for nonlinear error correction (NEC) models based on the concepts of I(1) and I(0) introduced by Escribano (1986, 1987) and Escribano and Mira (2002). Escribano (1986, 2004) proposes a methodology to implement parametric and nonparametric error correction models. Using the data bases of Friedman and Schwartz (1982) and Ericsson, Hendry and Prestwich (1998), extended until the year 2000, he implemented this methodology to estimate a nonlinear money demand for the U.K. from 1878 to 2000. Within the class of parametric models he discusses cubic polynomial (and rational polynomial) error correction models, see also Hendry and Ericsson (1991), Ericsson, Hendry and Prestwich, (1998). Nonlinearities can eliminate most of the power of usual noncointegration test as will be seen in the following Monte Carlo simulations.

Power of RCC against a Nonlinear cointegrating relationship

Consider now that the DGP under the alternative hypothesis is given by the following linear error correction model (ECM) with a nonlinear cointegration relationship.

H_1: The alternative hypothesis is a linear ECM but with *nonlinear cointegration*

DGP: $\Delta y_t = c + a\Delta x_t + b(y_{t-1} - g(x_{t-1}, \alpha)) + w_{1t}$

$\Delta x_t = w_{2t}$

where w_{1t} and w_{2t} are $Nid(0,1)$ and mutually independent errors with $\alpha = 1$.

We analyze now the power of the RCC test for different values of a and we compare the results with the EG test (results in the Table are in parenthesis). Let the nonlinear cointegration relationship be given by the following *polynomial cointegration* term $g(z_{t-1}, \alpha) = z_{t-1}^j$. Based on 10,000 replications of the Monte Carlo experiment we compute the empirical power of the tests for parameter values $a = 0, 1$ and values of $b = -0.01$, -0.1, -0.25, -0.75.

Table IV: Size adjusted power of the RCC test and the EG test (in parenthesis) at 5% significance level based on H_1 with g=z $^{j}_{t-1}$ j=2,3 and 4, for different sample sizes n, different values of b and fixing the parameter a at a=0

	$j\backslash b$		-0.01			-0.25	
		100	250	500	100	250	500
$a = 0$	2	0.71 (0.02)	0.86 (0.002)	0.94 (0.001)	0.88 (0.09)	0.98 (0.03)	1 (0.01)
	3	0.88 (0.06)	1 (0.02)	1 (0.003)	0.9 (0.07)	1 (0.03)	1 (0.009)
	4	0.92 (0.04)	1 (0.01)	1 (0.001)	0.93 (0.08)	1 (0.03)	1 (0.01)
	$j\backslash b$		-0.1			-0.75	
		100	250	500	100	250	500
$a = 0$	2	0.88 (0.1)	0.96 (0.02)	1 (0.01)	0.9 (0.1)	0.95 (0.07)	0.98 (0.07)
	3	0.94 (0.1)	1 (0.03)	1 (0.01)	0.9 (0.17)	1 (0.14)	1 (0.16)
	4	0.93 (0.1)	1 (0.11)	1 (0.10)	1 (0.23)	1 (0.234)	1 (0.275)

Table V: Size adjusted power of the RCC test and EG test (in parenthesis) at 5% significance level based on H_1 with g=z^{j}_{t-1} j=2,3 and 4, for different sample sizes n, different values of b and fixing the parameter a at a=1

	$j\backslash b$		-0.01			-0.25	
		100	250	500	100	250	500
$a = 1$	2	0.35 (0.008)	0.5 (0.003)	0.65 (0)	0.45 (0.057)	0.6 (0.02)	0.7 (0.005)
	3	0.6 (0.05)	0.7 (0.01)	0.87 (0.002)	0.6 (0.09)	0.8 (0.03)	0.9 (0.007)
	4	0.55 (0.04)	0.7 (0.01)	0.8 (0.002)	0.5 (0.09)	0.6 (0.02)	0.7 (0.005)
	$j\backslash b$		-0.1			-0.75	
		100	250	500	100	250	500
$a = 1$	2	0.5 (0.072)	0.6 (0.047)	0.8 (0.02)	0.3 (0.09)	0.4 (0.07)	0.6 (0.057)
	3	0.6 (0.09)	0.8 (0.05)	0.9 (0.02)	0.45 (0.13)	0.64 (0.15)	0.8 (0.13)
	4	0.55 (0.09)	0.7 (0.13)	0.8 (0.09)	0.3 (0.24)	0.45 (0.23)	0.6 (0.27)

As we can see from Table V, when the ECM is linear but with a polynomial cointegration function *RCC* test is much more powerful than the *EG* in all the cases. In particular, we observed that the highest power is obtained for $j = 3$ followed by $j = 4$ and $j = 2$ respectively. The intuition given by Escribano (2004) is the following; cubic polynomials are very flexible and can approximate different level shifts. Furthermore, the error correction term in this case can be equilibrium correcting (stable nonlinear adjustment) and therefore not a deviating error adjustment term as can happen with the quadratic polynomial. Notice that, as expected, the empirical power of the Tests decrease with the parameter a even for large negative values of b. Let the nonlinear cointegrating function be $g(z_{t-1}, \alpha) = \exp(z_{t-1}/100)$. The size adjusted power of the *RCC* test and the *EG* test (in parenthesis) for parameter values $a = 0, 1$ and $b = -0.01, -0.05, -0.1, -0.25, -0.5, -0.75$, are given in Table VI.

Table VI: Size adjusted power of the RCC test and EG test (in parenthesis) at 5% significance level based on H_1 with g=exp(z_{t-1}/100) for different sample sizes n and different values of parameters a and b

		$a = 0$			$a = 1$	
$b \backslash n$	100	250	500	100	250	500
-0.01	0.4 (0.06)	0.48 (0.05)	0.5 (0.05)	0.4 (0.07)	0.45 (0.07)	0.46 (0.08)
-0.05	0.43 (0.05)	0.52 (0.05)	0.57 (0.05)	0.43 (0.05)	0.46 (0.05)	0.5 (0.05)
-0.1	0.5 (0.05)	0.61 (0.05)	0.7 (0.05)	0.43 (0.06)	0.5 (0.06)	0.6 (0.06)
-0.25	0.6 (0.05)	0.8 (0.06)	0.82 (0.09)	0.5 (0.07)	0.54 (0.06)	0.7 (0.08)
-0.5	0.65 (0.07))	0.85 (0.09)	0.9 (0.1)	0.52 (0.08)	0.57 (0.13)	0.76 (0.13)
-0.75	0.7 (0.1)	0.9 (0.11)	1 (0.15)	0.6 (0.14)	0.65 (0.17)	0.8 (0.19)

The power of the EG test is very low since this linear procedure misspecifies the estimation of the cointegrating vector by assuming that it is linear. When there is no short run dependence ($a = 0$) the power of the RCC is very high and it is reduced when the parameter a is large. This is due to the fact that if the short run dependence is high, the nonlinearity also strongly affects the short run dependence of the cointegrating errors and therefore the power of those tests, that do not take this information into account is reduced. Similar results are obtained when the nonlinear cointegrating function is $g(z_{t-1}, \alpha) = \log(z_{t-1} + 100)$, see Garcia (2003). Polynomial transformations, cubic or rational, are flexible functional forms to study general nonlinear error correction adjustments in asymmetric contexts. Escribano (2004) justified those functional forms based on *Pade's approximations.* In what follows we analyze the power of the non cointegration test RCC and EG for nonlinear error correction models.

H_1: Linear cointegration with a nonlinear error correction (NEC)
DGP:

$$\Delta y_t = c + a\Delta x_t + f(y_{t-1} - \alpha x_{t-1}) + w_{1t} \tag{15}$$
$$\Delta x_t = w_{2t} \tag{16}$$

with w_{1t} , w_{2t} $Nid(0, 1)$ errors, mutually independent. Consider the following value of the linear cointegrating parameter $\alpha = 1$, with different nonlinear adjustments, $f(.)$. Let the cointegrating error be $U_{t-1} = y_{t-1} - \alpha x_{t-1}$ and let the nonlinear error correction adjustment be the following polynomial case:

$$f(U_{t-1}) = (U_{t-1} - 0.2)\, U_{t-1}^2 \tag{17}$$

The functional form (17), is justified by the Representation Theorem of NEC models of Escribano (2004) where the stability condition on the nonlinear adjustment is satisfied $(-2 < \frac{df(U_{t-1})}{dx_{t-1}} < 0)$.

Table VII: Size adjusted power of the RCC and the EG (in parenthesis) for different values of a small an n small of model (17)

a\\n	100	250	500
0	0.7 (0.04)	0.8 (0.03)	0.9 (0.02)
0.5	0.6 (0.02)	0.7 (0.03)	0.8 (0.07)
1	0.2 (0.08)	0.43 (0.02)	0.5 (0.01)

It is clear that the power of the RCC is much higher than the power of the EG and that the power of RCC decreases with the short run dependence of both series measures by parameter a. Following Escribano (2004), we combine now both types of linearities.

H_1: Nonlinear cointegration in a nonlinear error correction (NEC) model

DGP**:**

$$\Delta y_t = c + a\Delta x_t + f(y_{t-1} - g(x_{t-1}, \alpha)) + w_{1t} \quad (18)$$

$$\Delta x_t = w_{2t} \quad (19)$$

where w_{1t} , w_{2t} are $Nid(0,1)$, mutually independent and where $\alpha = 1$. We can consider two types of nonlinearities in the cointegration relationship.

Let $U_{t-1} = y_{t-1} - \log(x_{t-1})$ and substitute in (17) . The size adjusted power of the RCC and EG tests , are given in Table VIII if U_{t-1} is included in (17) . Similar power results are obtained for $U_{t-1} = \log(y_{t-1}) - \alpha x_{t-1}$.

Table VIII: Size adjusted power of the RCC and the EG (in parenthesis) for different values of a and n when U_{t-1} is included in (17)

a\\n	100	250	500
0	0.72 (0.02)	0.85 (0.032)	0.93 (0.04)
0.5	0.6 (0.02)	0.7 (0.04)	0.8 (0.02)
1	0.35 (0.09)	0.42 (0.021)	0.45 (0.012)

Linear cointegration with structural changes in the cointegrating vector

There is a body of literature on the effects of cointegration testing in the presence of structural changes. In what follows we want to simulate different cases based on the DGP of Arranz and Escribano (2001) to evaluate the power of the RCC in this context when the break point is in the middle of the sample.

H_1: Linear error correction and cointegration in the presence of structural change in the cointegrating vector

DGP:

$$\Delta x_t = w_{1t} \quad (20)$$

$$\Delta y_t = c + a\Delta x_t + b[y_{t-1} - (c_1 D_{1t-1} x_{t-1} + \alpha x_{t-1})] + w_{2t} \quad (21)$$

where w_{1t} , w_{2t} are $Nid(0,1)$ errors, and mutually independent, where c_1 measures the change in the cointegrating vector. We will consider the following values, $\alpha = 1$, $c_1 = 2$. The structural break is created by the artificial dummy variable D_{1t}, defined by,

$$D_{1t} = \begin{cases} 1, & t \geq \frac{n}{2} \\ 0, & \text{otherwise} \end{cases} \tag{22}$$

Based on 10,000 replications of the Monte Carlo simulations, we obtained the size adjusted power of the tests RCC and EG , see Table IX.

Table IX: Size adjusted power of the RCC and EG (in parenthesis) for different values of a,b, n and c $_1$ =2

$c_1 = 2$		$b = -0.01$			$b = -0.025$	
$a \backslash n$	100	250	500	100	250	500
0	0.5 (0.2)	0.65 (0.1)	0.66 (0.2)	0.65 (0.2)	0.8 (0.24)	0.85 (0.23)
1	0.41 (0.05)	0.5 (0.03)	0.64 (0.02)	0.54 (0.01)	0.61 (0.01)	0.8 (0.007)
$a \backslash n$		$b = -0.5$			$b = -0.75$	
0	0.81 (0.23)	1 (0.25)	1 (0.25)	0.8 (0.25)	1 (0.26)	1 (0.29)
1	0.62 (0.03)	0.8 (0.02)	0.9 (0.01)	0.61 (0.06)	0.73 (0.06)	0.91 (0.07)

Table IX show that the RCC is more powerful that the EG test to reject the null hypothesis of non cointegration in the presence of a structural break in the cointegrating vector. The power of the RCC is very good for moderate short run dependence ($a = 1$) and when the parameter (b) of the error correction adjustment is not slow. Those power results of the RCC test statistic are very promising however, we need to find a small sample correction of the RCC test statistic to make it invariant to the short run dependence indicated by the parameter a. As was mentioned in the previous section, the solution to this problem is in the test statistic that we called RCC_{CD} (RCC corrected for dependence). The idea of the correction is to count the synchronicity of the jumps, or new records, of both series (y and x) but relative to the synchronicity of the jumps in the first differences of the series (Δy and Δx). The RCC with the correction for dependence is, RCC_{CD} mentioned before. We showed in Table II that the critical values of the RCC_{CD} test statistic are now independent of the short run correlation measured by the parameter a. The question now is to evaluate the impact of this small sample correction on the power of the test. However, since the empirical distribution of the RCC_{CD} is very similar to the RCC with $a = 0$, and since for $a = 0$ the RCC test was *most powerful* we expect to also get important power improvements when $a \neq 0$ by using RCC_{CD} instead of RCC.

4. EMPIRICAL APPLICATIONS

Analysis of gold and silver prices

Gold and silver have been actively traded for thousands of years and remain important and closely observed markets. Here, following Escribano and Granger (1998), monthly prices are analyzed from the end of 1971 until 1996. We are interested in testing the existence of contemporaneous relationships between the prices of the two commodities in log levels. These prices are determined in clearly speculative markets and therefore their behavior should be captured by unit-root time series models. The unit root is supported by the Dickey Fuller (DF) type tests using from 1 to 6 lags of the dependent variable and including constant and constant and trend variables in the regression equation. There is, however, a feature of this data that makes it particularly interesting, which is the widely known and well-documented *bubble* in silver prices from roughly June 1979 to March 1980. The objective now is to investigate the effect of the bubble on testing for the existence of a long-run relationship (cointegration). The results in Escribano and Granger (1998) support the use of the intercept-dummies (level shifts in the intercept, introduced to explain the bubble period and their impact on the post-bubble period) greatly strengthens the evidence of cointegration.

When testing the null hypothesis using the EG test on the residuals with one lag we get a t-ratio $t_0 = -2.23$ and therefore we cannot reject H_0, at the 5% significance since the critical value is -3.7. However, if we consider the two level shifts that occur in the cointegration relationship we get $t_0 = -4.54$, which rejects the null of no cointegration[1]. Testing for cointegration between the log prices of gold and silver with the RCC test statistics is interesting since those procedures do not require prior estimation of the cointegrating relationship and could indicate departures from linearity.

On the available data sample (of length $n = 224$), the values obtained for the RCC and for the RCC_{CD} (with x_t, y_t representing now the logarithms of the gold and the silver price series) were 4.8045 (a=0, 5% C.V.=3.12) and 3.56 (5% C.V.=3.3), respectively. These findings support the rejection by H_0 of independent random walks in favor of the cointegration alternative hypothesis (between the log prices) respectively at the 5% significance level.

Analysis for the UK money demand (1878-2000)

Escribano (1986, 2004) found a cointegration relationship between the log of velocity of circulation of money (V) and short run interest rates in nominal terms (RNA), from annual observations from 1878 to 2000. Therefore, this cointegration relationship is nonlinear, and he showed that if we transform V and RNA in logs, they are still cointegrated and the cointegration relationship is linear.

However, when we test the null hypothesis of no cointegration between log(V) and RNA based on the residuals of the EG test statistic we obtain a value of -1.7 and therefore we cannot reject the null hypothesis 5%. This result contradicts the ECM test done in the nonlinear error correction model. Applying our RCC test statistic , we found $RCC = 3.5$, which is larger than the 5% critical value, for $n = 130$ and $a = 0$ (3.11). That is, with the RCC we reject the null of no cointegration if there is no short run contemporaneous correlation. The same result is obtained for certain values of $a \neq 0$ but not for all.For example, we also reject for $a = 0.5$, but not for $a = 1$ and 1.5. Therefore, it is important to apply the RRC_{CD} test statistic,which is independent of a, in order to take a final decision.

4. SUMMARY

[1]There is a body of literature on the effects of structural breaks on unit root and cointegration tests. See, for example, the references in Arranz and Escribano (2001).

In this paper we proposed a non-cointegration test statistic (RCC) based on the first differences of the ranges of the series. This model free cointegration testing device could also be used in the finite samples. The key idea is to count the synchronicity of the new records and it is therefore called record counting cointegration (RCC). The series will not be cointegrated if there is a lack of synchronicity (up to a constant delay) between the sequences of jumps or between the new records of the series. If the two series have a common stochastic component (common trend), their ranges will tend to jump together indicating the synchronicity of the new records. The RCC test statistic clearly outperforms the traditional DF unit root test on the cointegrating residuals (EG) as well as nonparametric tests in a similar context previously proposed by Aparicio, Escribano and Garcia (2006a). The small sample properties of RCC nonparametric test are analyzed by Monte Carlo simulations and with some empirical examples. In particular, we test the null hypothesis of non-cointegration against the alternative hypothesis of a nonlinear error correction (NEC) and/or nonlinear cointegration. Furthermore, we are be able to show that the RCC test for non-cointegration is robust to nonlinearities and structural breaks. One important advantage of this RCC approach is that it does not require previous estimation of the unknown (maybe nonlinear) cointegrating relationship. Our Monte Carlo simulation analysis suggests that the proposed methodology is robust to different departures from the classical linear cointegration context, like nonlinearities in the cointegrating relationship or in the error correction term, or to structrual changes in the cointegrating relationship. The performance evaluation of the RCC in terms of size and power, is compared to the EG cointegration test. However, the RCC is sensitive to the short run dependence between the series. Therefore we suggest a small sample correction for dependence, the RCC_{CD}, which works very well even in very small samples (good size and power properties). Finally our RCC test statistic is applied to different data sets to show it usefulness in identifying cointegration in the presence of important nonlinearities and structural changes.

Appendix A1

Lemma 1. Let $x_t = \sum_{i=1}^{t} e_{i,1}$ where $e_{i,1}$ are continuous i.i.d. random variables with bounded and symmetric pdf, zero means, and finite variances. Suppose that x_0 has also a bounded *pdf* and finite variance. And let $J_{(x)}^{(n)} = n^{-1/2} \sum_{t=1}^{n} \mathbf{1}(\Delta R_t^{(x)} > 0)$. Then we have $\sum_{t=1}^{n} \mathbf{1}(\Delta R_t^{(x)} > 0) = O(n^{-1/2})$

Proof. See Aparicio, Escribano and Garcia (2006b). ■

PROOF OF THEOREM 1. Since x_t is a random walk we have from Lemma 1 :

$$\sum_{t=1}^{n} \mathbf{1}(\Delta R_t^{(x)} > 0) = O(n^{-1/2})$$

$$\Longrightarrow \quad P(\Delta R_t^{(x)} > 0) = O(t^{-1/2})$$

$$\Longrightarrow \quad \left[P(\Delta R_t^{(x)} > 0)\right]^2 = O(t^{-1})$$

$$\Longrightarrow \quad \sum_{t=1}^{n} \left[P(\Delta R_t^{(x)} > 0)\right]^2 = O(\log n)$$

since from Euler‘s formula we can write $\sum_{t=1}^{n} t^{-1} = \log n + \gamma + \frac{1}{2n} + \frac{1}{12n^2} + O(n^{-4})$
$\gamma = 0.57721566$ (Euler's constant)
Now if x_t and y_t are independent we have

$$\begin{aligned} E\left\{RCC_{x,y}^{(n)}\right\} &= \sum_{t=1}^{n} P(\Delta R_t^{(x)} > 0) P(\Delta R_t^{(y)} > 0) \\ &= \sum_{t=1}^{n} \left[P(\Delta R_t^{(x)} > 0)\right]^2 = O(\log n) \end{aligned}$$

Therefore, under H_0, we can write for some positive constant μ:
$\mathrm{RCC}_{x,y}^{(n)} = \mu \log n + \delta_n V$
where V denotes a non-degenerate random variable with unit variance and δ_n defines the asymptotic order for the standard deviation of $RCC_{x,y}^{(n)}$. Our next objective is to determine δ_n. To do this, first note that $\mathrm{E}\left(RCC_{x,y}^{(n)} - \mu \log n\right)^2 = E\left(\delta_n V\right)^2 = \delta_n^2 E\left(V\right)^2$

$$\begin{aligned} E\left\{\left[RCC_{x,y}^{(n)}\right]^2\right\} &= E\left\{\sum_{t=1}^{n}\sum_{t'=1}^{n} \mathbf{1}(\Delta R_t^{(x)} > 0)\mathbf{1}(\Delta R_t^{(y)} > 0)\mathbf{1}(\Delta R_{t'}^{(x)} > 0)\mathbf{1}(\Delta R_{t'}^{(y)} > 0)\right\} \\ &= \sum_{t=1}^{n} \left[P(\Delta R_t^{(x)} > 0)\right]^2 + 2\sum_{t=1}^{n}\sum_{t'=t+1}^{n} \left[P(\Delta R_t^{(x)} \Delta R_{t'}^{(x)} > 0)\right]^2 \\ &= \sum_{t=1}^{n} \left[P(\Delta R_t^{(x)} > 0)\right]^2 + W_{x,y}^{(n)} \end{aligned}$$

where we let

$$\begin{aligned} W_{x,y}^{(n)} &= 2\sum_{t=1}^{n}\sum_{t'=t+1}^{n} \left[P(\Delta R_t^{(x)} \Delta R_{t'}^{(x)} > 0)\right]^2 \\ &= 2\sum_{t=1}^{n}\sum_{t'=t+1}^{n} \left[P(\Delta R_{t'}^{(x)} > 0 \mid \Delta R_t^{(x)} > 0)\right]^2 \left[P(\Delta R_t^{(x)} > 0)\right]^2 \\ &= 2\sum_{t=1}^{n-1} \left[P(\Delta R_t^{(x)} > 0)\right]^2 \sum_{t'=t+1}^{n} \left[P(\Delta R_{t'-t}^{(x)} > 0)\right]^2 \end{aligned}$$

Now observing that

$$\begin{aligned} & \sum_{t=1}^{n-1} \left[P(\Delta R_t^{(x)} > 0)\right]^2 \sum_{t'=t+1}^{n} \left[P(\Delta R_{t'-t}^{(x)} > 0)\right]^2 \\ &= \mu^2 \sum_{t=1}^{n-1} t^{-1}(\log n - \log t) \\ &= \mu^2 \left(\log n\right)^2 - \mu^2 \sum_{t=1}^{n-1} t^{-1} \log t \end{aligned}$$

$$
\begin{aligned}
&= \mu^2 (\log n)^2 - \mu^2 \left\{ \sum_{t=1}^{n} \left(\frac{t}{n}\right)^{-1} \left(\log \frac{t}{n} + \log n\right) \frac{1}{n} \right\} \\
&= \mu^2 (\log n)^2 - \mu^2 (\log n)^2 - \mu^2 \left\{ \sum_{t=1}^{n} \left(\frac{t}{n}\right)^{-1} \left(\log \frac{t}{n}\right) \frac{1}{n} \right\} \\
&\simeq -\mu^2 \int_{1/n}^{1} \frac{\log x}{x} dx \\
&= \frac{1}{2} \mu^2 (\log n)^2
\end{aligned}
$$

it follows:

$$
\begin{aligned}
Var\left\{RCC_{x,y}^{(n)}\right\} &\simeq \mu \log n + \mu^2 (\log n)^2 - \left[E\left\{RCC_{x,y}^{(n)}\right\}\right]^2 \\
&= \mu \log n
\end{aligned}
$$

This entails that

$\delta_n = O((\log n)^{1/2})$ and $(\log n)^{-1} RCC_{x,y}^{(n)} \to 1$

BIBLIOGRAPHY

(1) Aparicio, F.M. *Nonlinear Modelling and Analysis under Long-Range Dependence with an Application to Positive Time Series*, PhD thesis. Swiss Federal Institute of Technology, Dpt. of Electrical Enginnering. Laussane (Suiza), **1995**.

(2) Aparicio, F.M.; Escribano, A.; Garcia, A. *A Range Unit Root Test*, Working Paper of the Department of Statistics of Carlos III University, **2004**.

(3) Aparicio, F.M.; Escribano, A. and Garcia, A. *Range Unit Root (RUR) Tests: Robust against Nonlinearities, Error Distributions, Structural Breaks and Outliers*, Journal of Time series Analysis, **2006a**, Forthcomming.

(4) Aparicio, F.M.; Escribano, A.; Garcia, A. *Synchronicity Between Financial Time Series: An Exploratory Analysis*, Progress in Financial Markets Research, Nova Publishers (forthcoming)**2006b**.

(5) Aparicio, F.M.; Granger, C.W.J. *Nonlinear Cointegration and some New Tests for Comovements Modelling*, Working Paper 95-15, Depart. of Economics of the University of California at San Diego, **1995**.

(6) Arranz, M.A.; Escribano, A. *Cointegration Testing under Structural Breaks: A Robust Extended Error Correction Model*, Oxford Bulletin of Economic and Statistics, **2001**, *62*, 23-52.

(7) Dickey, A.; Fuller, W. *Distribution of the Estimators for AR Time Series with a Unit Root* , Journal of the American Statistical Association, **1979**, *74*, 427-431.

(8) Engle, R.; Granger, C.W.J. *Cointegration and Error Correction: Representation, Estimation and Testing* , Econometrica, **1987**, *55*, 251-276.

(9) Ericsson, N.R.; Hendry, D.F.; Prestwich, K.M. *Friedman and Schwartz (1982) revisited: Assesing annual and phase average models of money demand in the United Kingdom*, Empirical Economics, **1998**, *23*.

(10) Escribano, A. *Cointegration, Time Co-trends and Error Correction Systems: An Alternative Approach*, C.O.R.E. Discussion Paper, 8715, University of Louvain, **1986**.
(11) Escribano, A. *Error Correction Systems: Nonlinear Adjustments to Linear Long-run Relationships*, C.O.R.E. Discussion Paper, 8730, University of Louvain, **1987**.
(12) Escribano, A. *Nonlinear Error Correction: The Case of Money Demand in the U.K. (1878-2000)*, Macroeconomic Dynamics, **2004**, *8*, 76-116.
(13) Escribano, A.; Granger, C.W.J. *Investigating the Relationship between Gold and Silver Prices*, Journal of Forecasting, **1998**, *17*, 81-107. (28)
(14) Escribano, A.; Mira, S. *Nonlinear Error Correction Models*, Journal of Time Series Analysis, **2002**, *23*, 509-522.
(15) Franses, P.H.; Haldrup, N. *The Effects of Additive Outliers on Tests for Unit Roots and Cointegration*, Journal of Business and Economic Statistics, **1994**, *12*, 471-478.
(16) Friedman, M.; Schwartz, A. *Monetary Trends in the USA and the UK: Their Relation to Income, Prices, and Interest Rates, 1867-1975*, University of Chicago Press, Chicago. **1982**.
(17) Garcia, A. *Contrastes Basados en Recorridos para Raices Unitarias y Cointegracion*, Ph.D. Dissertation. Universidad Carlos III de Madrid, **2004**.
(18) Granger, C.W.J. *Some properties of Time Series Data and their use in Econometric Model Specification*, Journal of Econometrics, **1981**, *16*, 121-130.
(19) Granger, C.W.J. *Modelling Nonlinear Relationships between Extended Memory Variables*, Econometrica **1995**, *63 (2)*, 265-279.
(20) Granger, C.W.J.; Hallman, J.J. *Long-Memory Series with Attractors*, Oxford Bulletin of Economics and Statistics, **1991**, *53*, 11-26.
(21) Hendry, D.F.; Ericsson, N.R. *An Econometric Analysis of UK Money Demand in Monetary Trends in the USA and UK by Milton Friedman and Anna J. Schwartz*, American Economic Review, **1991**, *8*.
(22) Johansen, S. *Estimation and hypothesis testing of cointegration vectors in Gaussian vector autorregresive models* , Econometrica **1991**, *59*, 1551-1580.
(23) Kremers, J.J.; Ericsson, N.R.; Dolado, J.J. *The Power of Cointegration Tests*, Economics and Statistics, **1992**, *54*, 325-348.
(24) Leadbetter, M.; Rootz, H. *Extremal theory for stochastic processes*, Annals of Probability, **1987**, *16*, 431-478.
(25) Lindgren, G.; Rootzen, H. *Extremes values: theory and thechnical applications*, Scandinavian Journal of Statistics, **1987**, *14*, 241-279.
(26) Perron, P. *Testing for a Unit Root in a Time Series with a Changing Mean*, Journal of Business and Economic Statistics, **1990**, *70*, 153-162.
(27) Phillips, P.C.B; Perron, P. *Testing for a Unit Root in Time Series Regression* , Biometrika, **1988**, *75*, 335-346.