A model-free CUSUM-type statistic for testing the null of cointegration

Afonso-Rodríguez, Julio A.
Department of Applied Economics and Quantitative Methods
University of La Laguna
Faculty of Business and Economics
Camino La Hornera, s/n. Campus de Guajara, C.P. 38071 La Laguna, Tenerife, Canary Islands
Teléfono: 922317041, e-mail: jafonsor@ull.es

Abstract

We propose a new and simple to compute semiparametric CUSUM-type statistic based on the sequence of centered and squared OLS residuals from the estimation of a single-equation cointegrating regression model as the basis to test the null hypothesis of cointegration against no cointegration. The main novelty of this testing procedure is that, besides very simple corrections for serial correlation and endogeneity of the integrated regressors and the only use of OLS residuals, the non-standard limiting null distribution is invariant to the number and type of components appearing in the estimated regression. We derive such a limiting null distribution, establish its consistency rate under no cointegration and also present some numerical results to illustrate its finite-sample size and power properties.

Keywords and phrases: *Cointegrating regression, OLS estimators and residuals, covariance stationarity, CUSUM-type statistics*

JEL classification: *C12, C22*

A model-free CUSUM-type statistic for testing the null of cointegration

1. Introduction

Since the seminal contributions by Granger (1981) and Engle and Granger (1987), the literature on cointegration analysis has occupied a prominent place in the econometric analysis of multiple nonstationary time series. In macroeconometrics there are many examples where, given the nonstationary behaviour of the series involved, cointegration analysis plays a central role in examining their long-run joint behavior through the specification of a single cointegration relationship based on a regression equation. One of the questions that has received more attention, besides the issues on model specification and estimation, is the development of testing procedures with good size and power properties in finite samples to consistently discriminate between cointegration and no cointegration. There are many available parametric or semiparametric test statistics with such a good properties, but their tabulated limiting null distributions generally depend on the number and nature of the trending components appearing in the specification of the cointegrating regression model.

This paper propose a relatively simple to compute new statistic, that only require the use of the OLS residuals, for testing the null hypothesis of cointegration in a single-equation cointegrating regression model, with an arbitrary number of integrated regressors and a very general form of the deterministic trend component, whose limiting null distribution is invariant to the structure of the components of the estimated model. The proposed testing procedure is robust to endogenous integrated regressors, and through a simulation experiment it can be verify that have good size and power properties.

2. The model, assumptions and some basic results

It is assumed that the set of k+1, $k \geq 1$, observed series Y_t, $\mathbf{X}_{k,t} = (X_{1,t},...,X_{k,t})'$, t = 1, ..., n, are generated by the following unobserved components model

$$\begin{pmatrix} Y_t \\ \mathbf{X}_{k,t} \end{pmatrix} = \begin{pmatrix} d_{0,t} \\ \mathbf{d}_{k,t} \end{pmatrix} + \begin{pmatrix} \eta_{0,t} \\ \boldsymbol{\eta}_{k,t} \end{pmatrix} \tag{2.1}$$

where $d_{0,t}$ and $\mathbf{d}_{k,t} = (d_{1,t},...,d_{k,t})'$ are the deterministic components, and $\boldsymbol{\eta}_t = (\eta_{0,t}, \boldsymbol{\eta}'_{k,t})'$ denotes the stochastic trend component given by $\boldsymbol{\eta}_t = \boldsymbol{\eta}_{t-1} + \boldsymbol{\varepsilon}_t$, with initial value $\boldsymbol{\eta}_0 = o_p(n^{1/2})$, [1] and $\boldsymbol{\varepsilon}_t = (\varepsilon_{0,t}, \boldsymbol{\varepsilon}'_{k,t})'$, a strictly stationary and ergodic zero-mean vector error with a finite long-run covariance matrix, $\boldsymbol{\Gamma} = \boldsymbol{\Gamma}_0 + \boldsymbol{\Gamma}_1 + \boldsymbol{\Gamma}'_1$, where $\boldsymbol{\Gamma}_0 = E[\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}'_t]$, and $\boldsymbol{\Gamma}_1 = \sum_{j=1}^{\infty} E[\boldsymbol{\varepsilon}_{t-j} \boldsymbol{\varepsilon}'_t]$, with $\boldsymbol{\Gamma}$ partitioned as $\boldsymbol{\Gamma} = (\gamma_0^2, \boldsymbol{\gamma}_{k0}, \boldsymbol{\Omega}_{kk})$, $\boldsymbol{\gamma}_{k0} = \boldsymbol{\gamma}'_{0k}$. Then, if there exists a k-vector $\boldsymbol{\beta}_k$ such that

$$u_t = \eta_{0,t} - \boldsymbol{\beta}'_k \boldsymbol{\eta}_{k,t} = (1, -\boldsymbol{\beta}'_k) \begin{pmatrix} \eta_{0,t} \\ \boldsymbol{\eta}_{k,t} \end{pmatrix} = \boldsymbol{\kappa}'_k \boldsymbol{\eta}_t \tag{2.2}$$

is stationary with continuous spectral density, then it is said that $\eta_{0,t}$ and $\boldsymbol{\eta}_{k,t}$ are cointegrated in the sense of Engle and Granger (1987), with cointegrating vector

[1] This general assumption on the initial value includes the case of any random variable with bounded second moments, and also includes the case of a fixed constant value.

$\boldsymbol{\kappa}_k = (1, -\boldsymbol{\beta}_k')'$, and $\boldsymbol{\beta}_k = \boldsymbol{\Omega}_{kk}^{-1}\boldsymbol{\gamma}_{k0}$.[2] To obtain an operative version of (2.2) based on the observed variables in (2.1) we need to introduce a particular but quite general assumption of the structure of the deterministic components in (2.1). Likewise, we need to formulate a convenient set of assumption for the generating mechanism and stochastic properties of the error terms driving the stochastic trend components in (2.1) and the equilibrium error term u_t in (2.2). These assumptions are presented below.

Assumption 2.1. Deterministic components

It is assumed that the deterministic trend components in (2.1), $d_{i,t}$, can be factorized as $d_{0,t} = \boldsymbol{\alpha}_{0,m}'\boldsymbol{\tau}_{m,t}$ and $\mathbf{d}_{k,t} = \mathbf{A}_{k,m}\boldsymbol{\tau}_{m,t} + \mathbf{A}_{k,q}\boldsymbol{\tau}_{q,t}$, with $k > q$, where $\boldsymbol{\tau}_{m,t} = (t^{p_1}, ..., t^{p_m})'$, $\boldsymbol{\tau}_{q,t} = (t^{p_{m+1}}, ..., t^{p_{m+q}})'$, with integer powers $0 \le p_1 < ... < p_{m+q}$, and $q \ge 0$, where, whenever the trend coefficient matrices $\mathbf{A}_{k,m}$ and $\mathbf{A}_{k,q}$ are non-zero matrices, each column of the trend coefficient matrix $\mathbf{A}_{k,m}$ contains a non-zero element, and $\mathbf{A}_{k,q}$ is full rank, i.e. $\text{Rank}(\mathbf{A}_{k,q}) = q < k$.

Assumption 2.2. Multivariate linear process for the error terms

(A) The zero-mean k+1-vector $\boldsymbol{\xi}_{0,t} = (\upsilon_t, \boldsymbol{\varepsilon}_{k,t}')'$ is strictly stationary and ergodic and follows a linear process as $\boldsymbol{\xi}_{0,t} = \mathbf{D}(L)\mathbf{e}_t$, where $\mathbf{e}_t = (e_{0,t}, \mathbf{e}_{k,t}')'$ is a k+1-variate white noise process with zero mean, covariance matrix $\boldsymbol{\Sigma}_e > 0$ and (2+m)*th*-order finite moment, $E[|\mathbf{e}_t|^{2+m}] = E[(\mathbf{e}_t'\mathbf{e}_t)^{2+m}] < \infty$ for some $m \ge 0$. Also, for the infinite order polynomial matrix in the lag operator L, $\mathbf{D}(L) = \sum_{j=0}^{\infty}\mathbf{D}_j L^j = (\mathbf{d}_0(L), \mathbf{D}_k'(L))'$, it is assumed that $\mathbf{D}(1)$ has full rank, with coefficients satisfying the summability condition $\sum_{j=0}^{\infty} j^a \|\mathbf{D}_j\|^2 < \infty$, $a \ge 2$, with $\|\mathbf{D}_j\| = [Tr(\mathbf{D}_j'\mathbf{D}_j)]^{1/2}$.

(B) The regression error term, u_t, is given by $u_t = \alpha u_{t-1} + \upsilon_t$, with $0 \le \alpha \le 1$.

Part (A) of Assumption 2.2, with the use of the well-known BN decomposition $\mathbf{D}(L) = \mathbf{D}(1) - (1-L)\tilde{\mathbf{D}}(L)$, where $\tilde{\mathbf{D}}(L) = \sum_{j=0}^{\infty}\tilde{\mathbf{D}}_j L^j$, $\tilde{\mathbf{D}}_j = \sum_{i=j+1}^{\infty}\mathbf{D}_i$, j = 0, 1, ..., implies the following representation of the scaled partial sum process from $\boldsymbol{\xi}_{0,t}$

$$(1/\sqrt{n})\sum_{t=1}^{[nr]}\boldsymbol{\xi}_{0,t} = \mathbf{D}(1)(1/\sqrt{n})\sum_{t=1}^{[nr]}\mathbf{e}_t + (1/\sqrt{n})(\tilde{\mathbf{e}}_0 - \tilde{\mathbf{e}}_{[nr]})$$

with $\tilde{\mathbf{e}}_t = \tilde{\mathbf{D}}(L)\mathbf{e}_t = O_p(1)$ a well-defined stationary process, so that the last term is asymptotically negligible and $\mathbf{D}(1)n^{-1/2}\sum_{t=1}^{[nr]}\mathbf{e}_t \Rightarrow \mathbf{B}(r) = \mathbf{BM}(\boldsymbol{\Omega}_0)$, 0≤ r ≤1 by the classical Donsker's theorem, with $\mathbf{B}(r) = (B_\upsilon(r), \mathbf{B}_k(r)')'$ a (k+1)-dimensional Brownian motion with covariance matrix $\boldsymbol{\Omega}_0 = \mathbf{D}(1)\Sigma_e\mathbf{D}(1)'$. Once established this result, and taking into account that $\sup_{0\le r\le 1}|(1/\sqrt{n})\sum_{t=1}^{[nr]}(\boldsymbol{\xi}_t - \mathbf{D}(1)\mathbf{e}_t)| \le 2\max_{0\le t\le n}|(1/\sqrt{n})\tilde{\mathbf{e}}_t| = o_p(1)$,

[2] In such a case, as discussed in Phillips (1986) and Phillips and Ouliaris (1990), the stationarity of the cointegrating error sequence u_t given by (2.2) implies that the long-run variance of its first differences, $v_t = \Delta u_t = \boldsymbol{\kappa}_k'\Delta\boldsymbol{\eta}_t = \boldsymbol{\kappa}_k'\boldsymbol{\varepsilon}_t$, is zero, that is $\omega_v^2 = \sum_{j=-\infty}^{\infty}E[v_t v_{t-j}] = 0$. Given that $\det(\boldsymbol{\Gamma}) = \gamma_{0.k}^2\det(\boldsymbol{\Omega}_{kk})$, where $\gamma_{0.k}^2 = \gamma_0^2 - \boldsymbol{\gamma}_{k0}'\boldsymbol{\Omega}_{kk}^{-1}\boldsymbol{\gamma}_{k0} = \gamma_0^2(1-\rho_{k0}^2)$, $\rho_{k0}^2 = \boldsymbol{\gamma}_{k0}'(\gamma_0^2\boldsymbol{\Omega}_{kk})^{-1}\boldsymbol{\gamma}_{k0}$, is the conditional long-run variance of $\varepsilon_{0,t}$ given $\boldsymbol{\varepsilon}_{k,t}$, $\omega_v^2 = \boldsymbol{\kappa}_k'\boldsymbol{\Gamma}\boldsymbol{\kappa}_k$ and it is assumed that $\boldsymbol{\Omega}_{kk} > 0$ which excludes the possibility of integrated cointegrated regressors (subcointegration), with $\boldsymbol{\kappa}_k = (1, -\boldsymbol{\beta}_k')'$ the cointegration vector where $\boldsymbol{\beta}_k = \boldsymbol{\Omega}_{kk}^{-1}\boldsymbol{\gamma}_{k0}$, then we get $\omega_v^2 = \gamma_{0.k}^2$, with $\boldsymbol{\Gamma}$ a singular covariance matrix.

this also ensures that the partial sum process from $\boldsymbol{\xi}_{0,t}$ satisfies a multivariate invariance principle such that $n^{-1/2}\sum_{t=1}^{[nr]}\boldsymbol{\xi}_{0,t} \Rightarrow \mathbf{B}(r)$.[3] On the other hand, from part (B) and for any value $0 \leq \alpha < 1$, we can define the k+1-vector

$$\boldsymbol{\xi}_t = \begin{pmatrix} u_t \\ \boldsymbol{\varepsilon}_{k,t} \end{pmatrix} = \begin{pmatrix} (1-\alpha L)^{-1}\upsilon_t \\ \boldsymbol{\varepsilon}_{k,t} \end{pmatrix} = \mathbf{C}(L)\mathbf{e}_t = \begin{pmatrix} \mathbf{c}'(L)\mathbf{e}_t \\ \mathbf{D}_k(L)\mathbf{e}_t \end{pmatrix}$$

with $\mathbf{c}(L) = (1-\alpha L)^{-1}\mathbf{d}(L)$, that also satisfy the same summability conditions as for the coefficients in $\mathbf{d}(L)$,[4] which implies that the scaled partial sum of $\boldsymbol{\xi}_t$ also weakly converges to a well defined limit as for $\boldsymbol{\xi}_{0,t}$, where now $\mathbf{B}(r) = (B_u(r), \mathbf{B}_k(r)')'$ is a k+1-vector Brownian process with covariance matrix $\boldsymbol{\Omega} = \mathbf{C}(1)\Sigma_e\mathbf{C}(1)'$, with

$$\mathbf{C}(1) = \begin{pmatrix} \mathbf{c}'(1) \\ \mathbf{D}_k(1) \end{pmatrix} = \begin{pmatrix} (1-\alpha)^{-1}\mathbf{d}'(1) \\ \mathbf{D}_k(1) \end{pmatrix}$$

which implies that the long-run variance of u_t and covariance between u_t and $\boldsymbol{\varepsilon}_{k,t}$ are given by $\omega_u^2 = (1-\alpha)^{-2}\omega_\upsilon^2$, and $\boldsymbol{\omega}_{ku} = (1-\alpha)^{-1}\boldsymbol{\omega}_{k\upsilon}$, respectively, with ω_υ^2 and $\boldsymbol{\omega}_{k\upsilon}$ the corresponding elements in $\boldsymbol{\Omega}_0$. Furthermore, under the additional assumption of no cointegrated integrated regressors $\mathbf{X}_{k,t}$, so that $\boldsymbol{\Omega}_{kk} > 0$, the Brownian process $B_u(r)$ admits the decomposition $B_u(r) = B_{u.k}(r) + \boldsymbol{\omega}'_{ku}\boldsymbol{\Omega}_{kk}^{-1}\mathbf{B}_k(r)$, with $B_{u.k}(r) = \omega_{u.k}W_u(r)$, where $\omega_{u.k}^2 = \omega_u^2 - \boldsymbol{\omega}'_{ku}\boldsymbol{\Omega}_{kk}^{-1}\boldsymbol{\omega}_{ku}$ is the conditional long-run variance of u_t given $\boldsymbol{\varepsilon}_{k,t}$, so that $E[\mathbf{B}_k(r)B_{u.k}(r)] = \mathbf{0}_k$ and hence the jointly Gaussian processes $B_{u.k}(r)$ and $\mathbf{B}_k(r)$ are independent.

From the structure of the underlying deterministic components established in Assumption 2.1, and by combining (2.1)-(2.2) we obtain the following estimable version of the single-equation cointegrating regression model

$$Y_t = \boldsymbol{\alpha}'_m\boldsymbol{\tau}_{m,t} + \boldsymbol{\beta}'_k\mathbf{X}_{k,t} + u_t \tag{2.3}$$

which only incorporates the common structure of the deterministic components of each series. Next, defining the vector of trend regressors in (2.3) as $\mathbf{m}_t = (\boldsymbol{\tau}'_{m,t}, \mathbf{X}'_{k,t})'$, we must find a non-singular weighting matrix $\mathbf{W}_n$ such as $\mathbf{m}_t = \mathbf{W}_n\mathbf{m}_{nt}$, with $\mathbf{m}_{nt} = (\mathbf{m}'_{m,nt}, \mathbf{m}'_{k,nt})'$ a triangular array with a well-defined limit $\mathbf{m}_{n[nr]} \Rightarrow \mathbf{m}(r)$, where $\mathbf{m}(r) = (\mathbf{m}'_m(r), \mathbf{m}'_k(r))'$ is a full-ranked process, in the sense that $\int_0^1 \mathbf{m}(r)\mathbf{m}'(r)dr > 0$ a.s. Given the assumption on the structure of the deterministic component appearing in the cointegrating regression and the fact that the k-vector of stochastic regressors can be

---

[3] There could be some situations where this invariance principle is not enough, being necessary to have stronger approximations that involve explicit convergence rates. Thus, e.g., Park and Hahn (1999) prove that, under Assumption L, it is also verified that: (a) $\sup_{0\leq r\leq 1}|n^{-1/2}\sum_{t=1}^{[nr]}(\boldsymbol{\xi}_{0,t} - \mathbf{D}(1)\mathbf{e}_t)| = O_p(n^{-a})$, and (b) $\sup_{0\leq r\leq 1}|\mathbf{D}(1)n^{-1/2}\sum_{t=1}^{[nr]}\mathbf{e}_t - \mathbf{B}(r)| = O_p(n^{-a})$ for large n, where $a = (m-2)/2m$. This results imply that the convergence rate is faster if $\mathbf{e}_t$ has higher moments, with $n^a \to \sqrt{n}$ as $m\to\infty$.

[4] Given the partitioned structure of $\mathbf{D}(L) = (\mathbf{d}_0(L), \mathbf{D}'_k(L))'$, with $\mathbf{d}_0(L) = \sum_{i=0}^{\infty}\mathbf{d}_{0i}L^i$ and $\mathbf{D}_k(L) = \sum_{i=0}^{\infty}\mathbf{D}_{ki}L^i$, we then have $\sum_{j=0}^{\infty} j^a\|\mathbf{D}_j\| = \sum_{j=0}^{\infty} j^a Tr(\mathbf{d}_{0j}\mathbf{d}'_{0j}) + \sum_{j=0}^{\infty} j^a Tr(\mathbf{D}'_{kj}\mathbf{D}_{kj})$. Also, given recurrence relation $\mathbf{d}_{0j} = \mathbf{c}_j - \alpha\mathbf{c}_{j-1}$ for $j \geq 1$, we can write

$$\sum_{j=0}^{\infty} j^a Tr(\mathbf{d}_{0j}\mathbf{d}'_{0j}) = (1+\alpha^2)\sum_{j=0}^{\infty} j^a Tr(\mathbf{c}_j\mathbf{c}'_j) + \alpha^2\sum_{j=0}^{\infty} Tr(\mathbf{c}_j\mathbf{c}'_j) - \alpha\sum_{j=1}^{\infty} j^a Tr(\mathbf{c}_j\mathbf{c}'_{j-1} + \mathbf{c}_{j-1}\mathbf{c}'_j)$$

so that the summability condition on the coefficients $\mathbf{D}_j$, and hence for $\mathbf{d}_{0j}$, implies that of $\sum_{j=0}^{\infty} j^a Tr(\mathbf{c}_j\mathbf{c}'_j)$.

decomposed as $\mathbf{X}_{k,t} = \mathbf{A}_{k,m}\mathbf{\Gamma}_{m,n}^{-1}\boldsymbol{\tau}_{m,nt} + \mathbf{\Gamma}_{kk,n}^{-1}\mathbf{m}_{k,nt}$, with $\mathbf{m}_{k,nt} = \mathbf{\Gamma}_{kk,n}(\boldsymbol{\eta}_{k,t} + \mathbf{A}_{k,q}\boldsymbol{\tau}_{q,t})$, then a convenient choice for the weighting matrix $\mathbf{\Gamma}_{kk,n}$ is $\mathbf{\Gamma}_{kk,n} = \mathbf{W}_{kk,n}^{-1}\mathbf{C}'_{kk}$, so that $\mathbf{W}_n$ is given by

$$\mathbf{W}_n = \begin{pmatrix} \mathbf{\Gamma}_{m,n}^{-1} & \mathbf{0}_{m,k} \\ \mathbf{A}_{k,m}\mathbf{\Gamma}_{m,n}^{-1} & \mathbf{\Gamma}_{kk,n}^{-1} \end{pmatrix} \tag{2.4}$$

In the simplest case where $\mathbf{A}_{k,q} = \mathbf{0}_{k,q}$, which also includes the situation where $\mathbf{A}_{k,m} = \mathbf{0}_{k,m}$, then an obvious choice for the matrices composing $\mathbf{\Gamma}_{kk,n}$ is $\mathbf{W}_{kk,n} = \sqrt{n}\mathbf{I}_{kk}$ and $\mathbf{C}_{kk} = \mathbf{I}_{kk}$, so that

$$\mathbf{W}_n = \begin{pmatrix} \mathbf{\Gamma}_{m,n}^{-1} & \mathbf{0}_{m,k} \\ \mathbf{A}_{k,m}\mathbf{\Gamma}_{m,n}^{-1} & \sqrt{n}\mathbf{I}_{k,k} \end{pmatrix} \tag{2.5}$$

where $\mathbf{m}_{nt} = (\boldsymbol{\tau}'_{m,nt}, \boldsymbol{\eta}'_{k,nt})'$, with $\mathbf{m}_{k,nt} = \boldsymbol{\eta}_{k,nt} = n^{-1/2}\boldsymbol{\eta}_{k,t}$ and $\mathbf{m}_{k,n[nr]} \Rightarrow \mathbf{B}_k(r)$. On the other hand, when $\mathbf{A}_{k,q} \neq \mathbf{0}_{k,q}$ with $k > q$ and $\mathbf{A}_{k,q}$ is a full rank matrix, so that there are any additional deterministic term not included in the polynomial trend function of the specified regression, Hansen (1992a, b) proposes to use the weighting matrix $\mathbf{\Gamma}_{kk,n}$ given by

$$\mathbf{\Gamma}_{kk,n} = \begin{pmatrix} \mathbf{\Gamma}_{q,n} & \mathbf{0}_{k,k-q} \\ \mathbf{0}_{k-q,q} & (1/\sqrt{n})\mathbf{I}_{k-q,k-q} \end{pmatrix} \begin{pmatrix} \mathbf{C}'_{k,q} \\ \mathbf{C}'_{k,k-q} \end{pmatrix} \tag{2.6}$$

with $\mathbf{C}_{k,q} = \mathbf{A}_{k,q}(\mathbf{A}'_{k,q}\mathbf{A}_{k,q})^{-1}$, and $\mathbf{C}_{k,k-q} = \mathbf{A}^*_{k,k-q}(\mathbf{A}'^*_{k,k-q}\mathbf{\Omega}_{kk}\mathbf{A}^*_{k,k-q})^{-1/2}$ where $\mathbf{A}^*_{k,k-q}$ is the $k{\times}(k{-}q)$ matrix which spans the null space of $(\mathbf{A}_{k,m} : \mathbf{A}_{k,q})$, so that $\mathbf{C}_{kk} = (\mathbf{C}_{k,q}, \mathbf{C}_{k,k-q}) > 0$ and is well-defined and $\mathbf{m}_{k,nt}$ is now given by

$$\mathbf{m}_{k,nt} = \begin{pmatrix} \mathbf{\Gamma}_{q,n}\boldsymbol{\tau}_{q,t} + \sqrt{n}\mathbf{\Gamma}_{q,n}\mathbf{C}'_{k,q}\boldsymbol{\eta}_{k,nt} \\ \mathbf{C}'_{k,k-q}\boldsymbol{\eta}_{k,nt} \end{pmatrix} \tag{2.7}$$

with limit

$$\mathbf{m}_{k,n[nr]} \Rightarrow \mathbf{m}_k(r) = \begin{pmatrix} \boldsymbol{\tau}_q(r) \\ \mathbf{W}_{k-q}(r) \end{pmatrix} \tag{2.8}$$

where

$$\begin{aligned}\mathbf{W}_{k-q}(r) &= \mathbf{C}'_{k,k-q}\mathbf{B}_k(r) = (\mathbf{A}'^*_{k,k-q}\mathbf{\Omega}_{kk}\mathbf{A}^*_{k,k-q})^{-1/2}\mathbf{A}^*_{k,k-q}\mathbf{B}_k(r) \\ &= (\mathbf{A}'^*_{k,k-q}\mathbf{\Omega}_{kk}\mathbf{A}^*_{k,k-q})^{-1/2}\mathbf{A}^*_{k,k-q}\mathbf{\Omega}_{kk}^{1/2}\mathbf{W}_k(r) =^d \mathbf{BM}(\mathbf{I}_{k-q,k-q})\end{aligned}$$

so that the limit process $\mathbf{m}_k(r)$ is also full-ranked which allows the derivation of a nondegenerate asymptotic theory, but with a different limit and implications as in the standard case of $\mathbf{A}_{k,q} = \mathbf{0}_{k,q}$.

This set of basic results allows a convenient treatment of the OLS estimators and residuals from (2.3) under a wide variety of situations concerning the underlying structure of the integrated regressors and its effects on the corresponding limiting distributions. Thus, given $\hat{\boldsymbol{\theta}}_n = (\sum_{t=1}^n \mathbf{m}_t\mathbf{m}'_t)^{-1}\sum_{t=1}^n \mathbf{m}_t Y_t = \boldsymbol{\theta} + (\sum_{t=1}^n \mathbf{m}_t\mathbf{m}'_t)^{-1}\sum_{t=1}^n \mathbf{m}_t u_t$ the OLS estimator of $\boldsymbol{\theta} = (\boldsymbol{\alpha}'_m, \boldsymbol{\beta}'_k)'$, we get that the scaled and normalized vector of estimation errors is given by

$$\hat{\mathbf{\Theta}}_n = n^{\kappa}\mathbf{W}_n \begin{pmatrix} \hat{\boldsymbol{\alpha}}_{m,n} - \boldsymbol{\alpha}_m \\ \hat{\boldsymbol{\beta}}_{k,n} - \boldsymbol{\beta}_k \end{pmatrix} = \left((1/n)\sum_{t=1}^n \mathbf{m}_{nt}\mathbf{m}'_{nt} \right)^{-1} n^{-(1-\kappa)}\sum_{t=1}^n \mathbf{m}_{nt}u_t \tag{2.9}$$

where the exponent κ takes value $\pm 1/2$ depending on whether we assume cointegration or not. Thus, under cointegration (with $\kappa = 1/2$), and standard application of the weak limit of a sample covariance between the regression error and the k-vector of stochastic trend components, $\boldsymbol{\eta}_{k,t} = \boldsymbol{\eta}_{k,t-1} + \boldsymbol{\varepsilon}_{k,t}$, we get $n^{-1/2}\sum_{t=1}^{n}\boldsymbol{\eta}_{k,nt}u_t \Rightarrow \int_0^1 \mathbf{B}_k(r)dB_u(r) + \boldsymbol{\Delta}_{ku}$, with $\boldsymbol{\Delta}_{ku} = \sum_{j=0}^{\infty} E[\boldsymbol{\varepsilon}_{k,t-j}u_t]$ the one-sided long-run covariance between u_t and $\boldsymbol{\varepsilon}_{k,t}$. Then, the weak limit of the sample vector covariance in the last term of (2.9) is given by

$$n^{-1/2}\sum_{t=1}^{n}\mathbf{m}_{nt}u_t \Rightarrow \begin{pmatrix} \int_0^1 \boldsymbol{\tau}_m(r)dB_u(r) \\ \int_0^1 \mathbf{m}_k(r)dB_u(r) \end{pmatrix} + \begin{pmatrix} \mathbf{0}_m \\ \boldsymbol{\Phi}_{ku} \end{pmatrix} \tag{2.10}$$

where $\mathbf{m}_k(r) = \mathbf{B}_k(r)$ and $\boldsymbol{\Phi}_{ku} = \boldsymbol{\Delta}_{ku}$ when $\mathbf{A}_{k,q} = \mathbf{0}_{k,q}$, and $\mathbf{m}_k(r)$ is as in (2.8) with $\boldsymbol{\Phi}_{ku} = (\mathbf{0}_q', \boldsymbol{\Delta}_{ku}'\mathbf{C}_{k,k-q})'$. This last result implies that, besides the presence of nuisance parameters arising from the endogeneity of the stochastic regressors and the serial correlation in the regression error terms, the limiting null distribution of the OLS estimates of the model parameters strongly depends on the nature of the deterministic trend components in $\mathbf{X}_{k,t}$, which means that a correct use of (2.10) requires to know the trend properties of these variables. As has been proved in Hansen (1992a, b), the same applies to the asymptotically efficient estimates obtained when using the Fully Modified OLS (FM-OLS) estimator proposed by Phillips and Hansen (1990). Moreover, taking into account that the OLS residuals from (2.3), $\hat{u}_t = Y_t - \mathbf{m}_t'\hat{\boldsymbol{\theta}}_n$, that can be represented as

$$\hat{u}_t = u_t - n^{-\kappa}\mathbf{m}_{nt}'\hat{\boldsymbol{\Theta}}_n \tag{2.11}$$

although being consistent estimators (under cointegration) of the regression error terms, $\hat{u}_t = u_t + O_p(n^{-1/2})$, will also be dependent on the nature of the underlying deterministic trend component of the stochastic regressors.

Thus, the limiting distribution under cointegration of many different functionals based on the sequence of OLS residuals, such as the partial sum given by $\hat{U}_t = \sum_{j=1}^{t}\hat{u}_j = \sum_{j=1}^{t}u_j - n^{1-\kappa}(n^{-1}\sum_{j=1}^{t}\mathbf{m}_{nj}')\hat{\boldsymbol{\Theta}}_n$, will also depend on this feature. As a wide variety of semiparametric and parametric statistics, both for testing the null hypothesis of cointegration against no cointegration (as the ones proposed by Shin (1994), Xiao (1999), Xiao and Phillips (2002), and Wu and Xiao (2008)) and for the reserved hypothesis (see, e.g., the residual-based statistics proposed in Phillips (1987) and Phillips and Ouliaris (1990)), exploit the information content of these residuals and their limiting distributions basically depends on the number of stochastic and deterministic trend components in the estimated cointegrating regression, it could be quite common to use in practice a wrong set of critical values if we only rely on the specification of the cointegrating regression without paying special attention to the structure of $\mathbf{d}_{k,t}$. With this in mind, but still relying on the usefulness of the information contained in the sequence of OLS residuals, next section will present a new testing procedure that overcome the difficulties discussed above.

3. A new CUSUM of squares test statistic

In the context of testing for stationarity of a univariate time series, Xiao and Lima (2007) consider, as an extended source of information for determining this type of behavior, the existence of excessive fluctuations in the bivariate process $\mathbf{z}_t = (u_t, u_t^2 - \sigma_n^2)'$, with $\sigma_n^2 = n^{-1}\sum_{s=1}^{n}u_s^2$. In order to define a proper measure of excessive fluctuation in these two series, these authors propose to built the scaled partial sum of $\mathbf{z}_t$ as

$$n^{-1/2}\sum_{t=1}^{[nr]}\begin{pmatrix}u_t\\ v_t\end{pmatrix}=\begin{pmatrix}n^{-1/2}\sum_{t=1}^{[nr]}u_t\\ n^{-1/2}\sum_{t=1}^{[nr]}(u_t^2-\sigma_n^2)\end{pmatrix} \tag{3.1}$$

Under stationarity (i.e., under cointegration if u_t where the regression errors in (2.3)), the scaled partial sum of centered and squared errors admits the decomposition

$$n^{-1/2}\sum_{t=1}^{[nr]}v_t=n^{-1/2}\sum_{t=1}^{[nr]}(u_t^2-\sigma_u^2)-\frac{[nr]}{n}n^{-1/2}\sum_{t=1}^{n}(u_t^2-\sigma_u^2) \tag{3.2}$$

which, under quite general assumptions, weakly converges to a well-defined limit distribution. On the other hand, under no stationarity (i.e., under no cointegration) we have that

$$n^{-5/2}\sum_{t=1}^{[nr]}\begin{pmatrix}u_t\\ v_t\end{pmatrix}=\begin{pmatrix}(1/n)\left\{(1/n)\sum_{t=1}^{[nr]}(u_t/\sqrt{n})\right\}\\ (1/n)\sum_{t=1}^{[nr]}[(u_t/\sqrt{n})^2-(1/n)\sigma_n^2]\end{pmatrix}=\begin{pmatrix}O_p(n^{-1})\\ (1/n)\sum_{t=1}^{[nr]}[(u_t/\sqrt{n})^2-(1/n)\sigma_n^2]\end{pmatrix} \tag{3.3}$$

so that the behavior under no stationarity is dominated by the second component, reflecting the violation of the covariance stationarity assumption induced by the unit root. This result give us the idea to define the empirical version of (3.2) based on the squared and centered OLS residuals as the basis for building a relatively simple to compute statistic to test the null hypothesis of cointegration. First, given the sequence of OLS residuals in equation (2.11), $\hat{u}_t=u_t-n^{-\kappa}\mathbf{m}_{nt}'\hat{\mathbf{\Theta}}_n$, we can write

$$\begin{aligned}(1/\sqrt{n})\sum_{t=1}^{[nr]}\hat{v}_t&=(1/\sqrt{n})\sum_{t=1}^{[nr]}(\hat{u}_t^2-\hat{\sigma}_n^2)\\ &=(1/\sqrt{n})\sum_{t=1}^{[nr]}v_t+n^{1/2-2\kappa}\hat{\mathbf{\Theta}}_n'\left(n^{-1}\sum_{t=1}^{[nr]}\mathbf{m}_{nt}\mathbf{m}_{nt}'-\frac{[nr]}{n}n^{-1}\sum_{t=1}^{n}\mathbf{m}_{nt}\mathbf{m}_{nt}'\right)\hat{\mathbf{\Theta}}_n\\ &\quad-2n^{1/2-2\kappa}\hat{\mathbf{\Theta}}_n'\left(n^{-(1-\kappa)}\sum_{t=1}^{[nr]}\mathbf{m}_{nt}u_t-\frac{[nr]}{n}n^{-(1-\kappa)}\sum_{t=1}^{n}\mathbf{m}_{nt}u_t\right)\end{aligned} \tag{3.4}$$

with $\hat{\sigma}_n^2=(1/n)\sum_{s=1}^{n}\hat{u}_s^2$ the residual variance. Hence, taking the exponent κ = 1/2 under the assumption of cointegration, it is easy to show that the behaviour of (3.4) is asymptotically determined by (3.2), that is $(1/\sqrt{n})\sum_{t=1}^{[nr]}\hat{v}_t=(1/\sqrt{n})\sum_{t=1}^{[nr]}v_t+O_p(n^{-1/2})$, so that the limiting distribution of (3.4) under cointegration will be invariant to the structure and nature of the regressors in (2.3). In order to correctly characterize this limiting null distribution, we only have to define an augmented version of the stationary error vector $\boldsymbol{\xi}_t$, $\boldsymbol{\xi}_t=(u_t,v_t,\boldsymbol{\varepsilon}_{k,t}')'$, that also satisfy an invariance principle such as

$$(1/\sqrt{n})\sum_{t=1}^{[nr]}\begin{pmatrix}u_t\\ v_t\\ \boldsymbol{\varepsilon}_{k,t}\end{pmatrix}\Rightarrow\begin{pmatrix}B_u(r)\\ B_v(r)\\ \mathbf{B}_k(r)\end{pmatrix}=\mathbf{BM}(\mathbf{\Omega}_1) \tag{3.5}$$

where, from the upper triangular Cholesky decomposition of the long-run covariance matrix $\mathbf{\Omega}_1$, we get $B_v(r)=\omega_{v.k}W_v(r)+\boldsymbol{\gamma}_{kv}'\mathbf{B}_k(r)$, with $\omega_{v.k}^2=\omega_v^2-\boldsymbol{\omega}_{kv}'\mathbf{\Omega}_{kk}^{-1}\boldsymbol{\omega}_{kv}=\omega_v^2-\boldsymbol{\omega}_{kv}'\boldsymbol{\gamma}_{kv}$ the conditional long-run variance of v_t given $\boldsymbol{\varepsilon}_{k,t}$, and $\boldsymbol{\gamma}_{kv}=\mathbf{\Omega}_{kk}^{-1}\boldsymbol{\omega}_{kv}$ with $\boldsymbol{\omega}_{kv}$ the long-run covariance between v_t and $\boldsymbol{\varepsilon}_{k,t}$. Following, e.g., Phillips and Solo (1992), and

Ibragimov and Phillips (2008), for a univariate stationary sequence u_t given by a linear process on an iid or martingale difference sequence e_t, such as $u_t = c(L)e_t = \sum_{i=0}^{\infty} c_i e_{t-i}$, with $c(L) = \sum_{i=0}^{\infty} c_i L^i$, $\sum_{i=0}^{\infty} i c_i^2 < \infty$, $c(1) \neq 0$, and $E[|e_0|^{2+m}] < \infty$, $m > 2$, the limiting behaviour of the sample covariance is given by

$$(1/\sqrt{n})\sum_{t=1}^{[nr]}(u_t u_{t+h} - \gamma_u(h)) \Rightarrow B_h(r) = \kappa(h)W_h(r)$$

for any $h = 0,1,...$, where $\gamma_u(h) = E[u_t u_{t+h}] = g_h(1)\sigma_e^2$, $g_h(1) = \sum_{i=0}^{\infty} c_i c_{i+h}$, $c_i = 0$ for $i < 0$, $\kappa^2(h) = g_h^2(1)E[(e_0^2 - \sigma_e^2)^2] + \sigma_e^4 \sum_{s=1}^{\infty}(g_{h+s}(1) + g_{h-s}(1))^2$, with

$$\kappa^2(0) = g_0^2(1)E[(e_0^2 - \sigma_e^2)^2] + 4\sigma_e^4 \sum_{s=1}^{\infty} g_s^2(1)$$

when $h = 0$, with $g_{-s}(1) = g_s(1)$. In our case, given that from Assumption 2.2, we have define $u_t = \mathbf{c}'(L)\mathbf{e}_t = \sum_{j=0}^{\infty} \mathbf{c}'_j \mathbf{e}_{t-j}$, with $\sigma_u^2 = \mathbf{c}'(1)\mathbf{\Sigma}_e \mathbf{c}(1)$, then we obtain the following decomposition of $v_{0,t} = u_t^2 - \sigma_u^2$

$$v_{0,t} = u_t^2 - \sigma_u^2 = \sum_{j=0}^{\infty} \mathbf{c}'_j(\mathbf{e}_{t-j}\mathbf{e}'_{t-j} - \mathbf{\Sigma}_e)\mathbf{c}_j + 2\sum_{s=1}^{\infty}\sum_{j=0}^{\infty} \mathbf{c}'_j(\mathbf{e}_{t-j}\mathbf{e}'_{t-j-s})\mathbf{c}_{j+s}$$

$$= \sum_{j=0}^{\infty}(\mathbf{c}'_j \otimes \mathbf{c}'_j)\mathrm{vec}(\mathbf{e}_{t-j}\mathbf{e}'_{t-j} - \mathbf{\Sigma}_e) + 2\sum_{s=1}^{\infty}\sum_{j=0}^{\infty}(\mathbf{c}'_{j+s} \otimes \mathbf{c}'_j)\mathrm{vec}(\mathbf{e}_{t-j}\mathbf{e}'_{t-j-s})$$

so that the scaled partial sum process of $v_{0,t}$ is given by

$$\begin{aligned} n^{-1/2}\sum_{t=1}^{[nr]} v_{0,t} &= \sum_{j=0}^{\infty}(\mathbf{c}'_j \otimes \mathbf{c}'_j)n^{-1/2}\sum_{t=1}^{[nr]}\mathrm{vec}(\mathbf{e}_{t-j}\mathbf{e}'_{t-j} - \mathbf{\Sigma}_e) \\ &\quad + 2\sum_{s=1}^{\infty}\sum_{j=0}^{\infty}(\mathbf{c}'_{j+s} \otimes \mathbf{c}'_j)n^{-1/2}\sum_{t=1}^{[nr]}\mathrm{vec}(\mathbf{e}_{t-j}\mathbf{e}'_{t-j-s}) \end{aligned} \tag{3.6}$$

By application of a multivariate version of the invariance principle for the iid sequence $\mathbf{e}_t$ and under cointegration, (3.6) will have a well-defined limiting distribution given by the Brownian process $B_v(r)$ appearing in (3.5), so that

$$(1/\sqrt{n})\sum_{t=1}^{[nr]}\hat{v}_t = (1/\sqrt{n})\sum_{t=1}^{[nr]} v_t + O_p(n^{-1/2}) \Rightarrow V_v(r) = B_v(r) - rB_v(1) \tag{3.7}$$

where $V_v(r)$ a first-level Brownian Bridge process, which can be decomposed as

$$V_v(r) = \omega_{v.k}(W_v(r) - rW_v(1)) + \boldsymbol{\gamma}'_{kv}(\mathbf{B}_k(r) - r\mathbf{B}_k(1)) \tag{3.8}$$

Thus, combining (3.7) and (3.8), we define our test statistic as the maximum absolute fluctuation of a modified version of the CUSUM of squared and centered OLS residuals and it is given by

$$C\hat{S}_n = \frac{1}{\sqrt{n}\hat{\omega}_{v.k,n}(m_n)} \max_{t=1,\ldots,n}\left|\sum_{j=1}^{t}\hat{v}_j - \hat{\boldsymbol{\gamma}}'_{kv,n}(m_n)(\mathbf{X}_{k,t} - (t/n)\mathbf{X}_{k,n})\right| \tag{3.9}$$

where $\hat{\omega}_{v.k,n}^2(m_n) = \hat{\omega}_{v,n}^2(m_n) - \hat{\boldsymbol{\omega}}'_{kv,n}(m_n)\hat{\boldsymbol{\gamma}}_{kv,n}(m_n)$ is the usual plug-in kernel estimate of $\omega_{v.k}^2$, with

$$\hat{\omega}_{v,n}^2(m_n) = \sum_{j=-(n-1)}^{n-1} w(jm_n^{-1})\left(n^{-1}\sum_{t=j+1}^{n}\hat{v}_t\hat{v}_{t-j}\right) \tag{3.10}$$

and $\hat{\boldsymbol{\gamma}}_{kv,n}(m_n) = \hat{\mathbf{\Omega}}_{kk,n}^{-1}(m_n)\hat{\boldsymbol{\omega}}_{kv,n}(m_n)$, the kernel estimators of the corresponding variances and covariances, with kernel function $w(\cdot)$ and bandwidth m_n. It is assumed that both components of these estimators satisfy the regularity conditions stated in Jansson (2002) in order to obtain consistent estimates of the corresponding parameters. Also, following Hansen (1992a,b), the estimates $\hat{\mathbf{\Omega}}_{kk,n}(m_n)$ and $\hat{\boldsymbol{\omega}}_{kv,n}(m_n)$ are computed by using the sequence of OLS-detrended observations of the integrated regressors given by the sequence of OLS residuals in the multivariate regression $\Delta\mathbf{X}_{k,t} = \mathbf{A}_k\Delta\boldsymbol{\tau}_t + \boldsymbol{\varepsilon}_{k,t}$, with $\mathbf{A}_k = (\mathbf{A}_{k,m}, \mathbf{A}_{k,q})$, and $\boldsymbol{\tau}_t = (\boldsymbol{\tau}'_{m,t}, \boldsymbol{\tau}'_{q,t})'$. The following results establish the limiting null distribution of (3.9) and the rate of divergence under no cointegration.

Proposition 3.1. *Limiting null distribution of the maximum absolute fluctuation of the modified CUSUM of squares statistic based on OLS residuals*

Under Assumption 2.2, with finite four moments for the innovations driving the linear process for u_t, and with a kernel function and bandwidth parameter satisfying the conditions stated in Jansson (2002), then under cointegration

$$C\hat{S}_n \Rightarrow \sup_{r\in[0,1]} |W_v(r) - rW_v(1)| \tag{3.11}$$

which is the supremum of a standard Brownian Bridge

Proof. This result follows directly from the above results and standard application of the Continuous Mapping Theorem.

Proposition 3.2. *Consistency rate under no cointegration*

Under the same conditions as in Proposition 3.1, but under a nonstationary behavior of the regression error term u_t, that is, under no cointegration, we have that:

(a) $(1/\sqrt{n})\sum_{t=1}^{[nr]} \hat{v}_t = O_p(n\sqrt{n})$

(b) $\hat{\omega}^2_{v.k,n}(m_n) = O_p(m_n n^2), \hat{\boldsymbol{\omega}}_{kv,n}(m_n) = O_p(m_n\sqrt{n})$

and

(c) $C\hat{S}_n = O_p(\sqrt{n/m_n})$

Proof. The above divergence rates can be obtained by making use of standard results in cointegration analysis based on OLS estimation of a single-equation cointegrating regression model, as the one specified in (2.3). In any case, the formal proof can be requested from the author.

4. Finite sample alternative distribution, size and power

The DGP used for the simulation experiment is based on $u_t = \alpha u_{t-1} + \upsilon_t$ and $\boldsymbol{\eta}_{k,t} = \boldsymbol{\eta}_{k,t-1} + \boldsymbol{\varepsilon}_{k,t}$, $\boldsymbol{\varepsilon}_{k,t} = \phi\boldsymbol{\varepsilon}_{k,t-1} + \mathbf{e}_{k,t}$, where $\boldsymbol{\xi}_{0,t} = (\upsilon_t, \mathbf{e}'_{k,t})' \sim iidN(\mathbf{0}_{k+1}, \mathbf{\Sigma}_{k+1})$, with

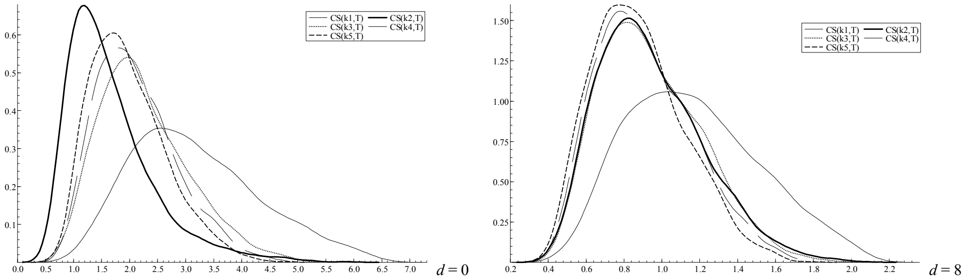
$$\mathbf{\Sigma}_{k+1} = \begin{pmatrix} \sigma_0^2 & \boldsymbol{\sigma}_{0k} \\ \boldsymbol{\sigma}_{k0} & \mathbf{\Sigma}_{kk} \end{pmatrix}, \text{ and } \sigma^2_{0.k} = \sigma_0^2 - \boldsymbol{\sigma}_{0k}\mathbf{\Sigma}^{-1}_{kk}\boldsymbol{\sigma}_{k0}$$

With these error terms we compute the OLS residuals from (2.3) without specifying any particular value for the model parameters, where all the results are computed by generating 5000 draws from the discrete time approximation (direct simulation) to the limiting random variables based on *n* steps, with *k* = 1, …, 5. In this preliminary version of the paper, and to save space, we only present the more relevant results related to the power behavior and are displayed graphically.

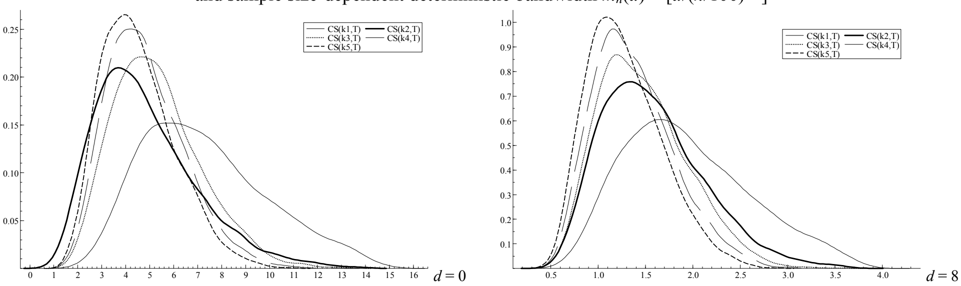
First, figure 1 display kernel estimates of the density function characterizing the distribution of the CUSUM of squares statistic under the alternative hypothesis of no cointegration. For different choices of the parameters determining the degree of endogeneity and serial correlation of the integrated regressors, we observe a markedly different behaviour for low, medium and high dimensional regression models in terms of *k*.

Figure 1. *Nonparametric kernel estimation of the density function of the CUSUM-of squares test statistics computed under the alternative of no cointegration*

Case A. No deterministic component, with sample size $n = 200$, $(\sigma_{k\upsilon}, \phi) = (0.75, 0.50)$, and sample size-dependent deterministic bandwidth $m_n(d) = [d/(n/100)^{1/4}]$



Case B. No deterministic component, with sample size $n = 1000$, $(\sigma_{k\upsilon}, \phi) = (0.75, 0.50)$, and sample size-dependent deterministic bandwidth $m_n(d) = [d/(n/100)^{1/4}]$



Additionally, figures 2 and 3 display the power profile for relatively small, medium and large samples sizes for different choices of the magnitude of the bandwidth parameter in computing the kernel estimates of the long-run variances and covariances appearing in (3.9) and (3.10).

Figure 2. *Power profile as a function of the number of integrated regressors (k), the magnitude of the deterministic bandwidth parameter* $m_n(d) = [d(n/100)^{1/4}]$*, and the degree of serial correlation in the errors driving the stochastic trend components* (ϕ)

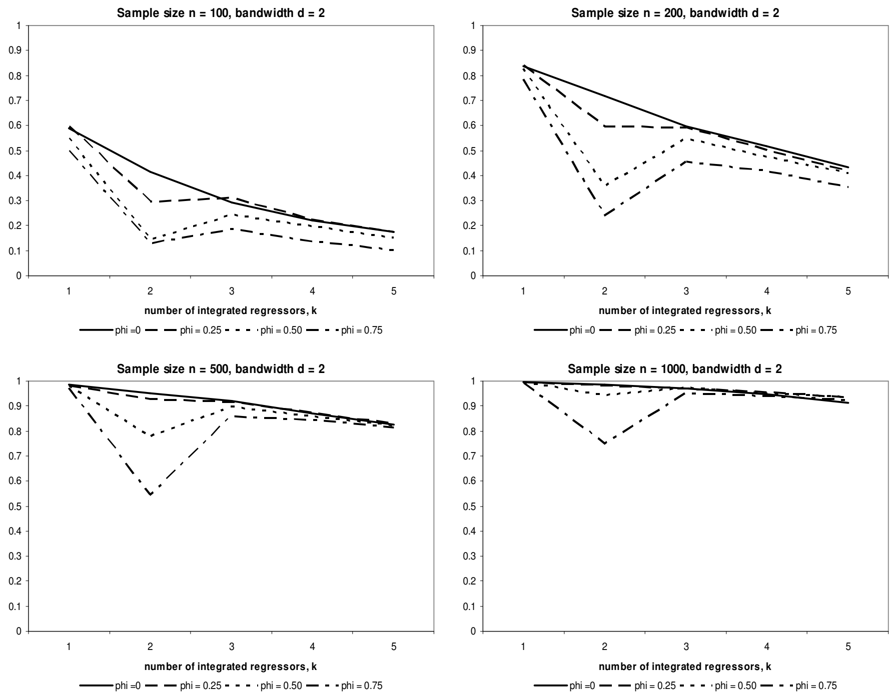
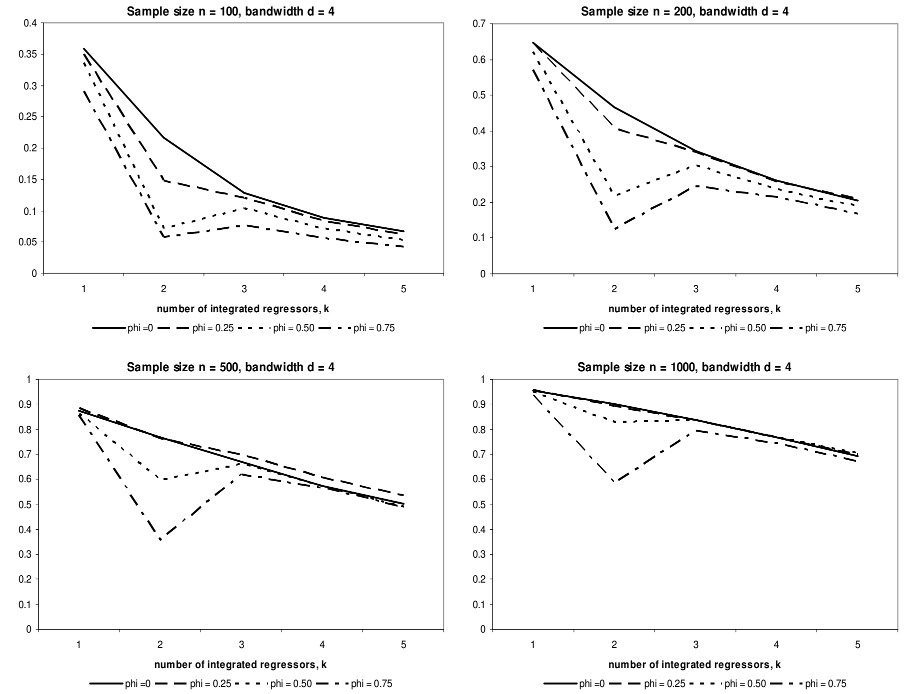


Figure 3. *Power profile as a function of the number of integrated regressors (k), the magnitude of the deterministic bandwidth parameter* $m_n(d) = [d(n/100)^{1/4}]$*, and the degree of serial correlation in the errors driving the stochastic trend components* (ϕ)



Power results displayed in figures 2, 3 corresponds to the case of no deterministic component in the estimated cointegrating regression. For the case of inclusion of a constant term, or a constant term and a linear trend, the profiles are quite similar, but with a slight loss of power, which is a feature commonly shared by many of the existing testing procedures in this framework. Also, another common feature, also displayed here is that the power is a decreasing function of the number of stochastic trends, *k*. For high dimensional models, as can be seen from figure 1, the limiting distribution under the alternative becomes right skewed. However, for low dimensional models and moderate sample sizes, the power is quite high and can be compared favourably with that of others existing testing procedures.

References

Engle, R.F., C.W.J. Granger (1987). Co-integration and error correction: representation, estimation, and testing. Econometrica, 55(2), 251-276.

Granger, C.W.J. (1981). Some properties of time series data and their use in econometric model specification. Journal of Econometrics, 16(1), 121-130.

Hansen, B.E. (1992a). Efficient estimation and testing of cointegrating vectors in the presence of deterministic trends. Journal of Econometrics, 53(1-2), 87-121.

Hansen, B.E. (1992b). Tests for parameter instability in regressions with I(1) processes. Journal of Business and Economic Statistics, 20(1), 45-59.

Harris, D., B. McCabe, S. Leybourne (2003). Some limit theory for autocovariances whose order depends on sample size. Econometric Theory, 19(5), 829-864.

Ibragimov, R., P.C.B. Phillips (2008). Regression asymptotics using martingale convergence methods. Econometric Theory, 24(4), 888-947.

Jansson, M. (2002). Consistent covariance matrix estimation for linear processes. Econometric Theory, 18(6), 1449-1459.

Park, J.Y., S.B. Hahn (1999). Cointegrating regressions with time varying coefficients. Econometric Theory, 15(5), 664-703.

Phillip, P.C.B., V. Solo (1992). Asymptotics for linear processes. The Annals of Statistics, 20(2), 971-1001.

Phillips, P.C.B. (1987). Time series regression with a unit root. Econometrica, 55(2), 277-301.

Phillips, P.C.B. (1989). Partially identified econometric models. Econometric Theory, 5(2), 181-240.

Phillips, P.C.B., B.E. Hansen (1990). Statistical inference in instrumental variables regression with I(1) processes. The Review of Economic Studies, 57(1), 99-125.

Phillips, P.C.B., S. Ouliaris (1990). Asymptotic properties of residual based tests for cointegration. Econometrica, 58(1), 165-193.

Shin, Y. (1994). A residual-based test of the null of cointegration against the alternative of no cointegration. Econometric Theory, 10(1), 91-115.

Wu, G., Z. Xiao (2008). Are there speculative bubbles in stock markets? Evidence from an alternative approach. Statistics and its Interface, 1, 307-320.

Xiao, Z. (1999). A residual based test for the null hypothesis of cointegration. Economics Letters, 64(2), 133-141.

Xiao, Z., L.R. Lima (2007). Testing covariance stationarity. Econometric Reviews, 26(6), 643-667.

Xiao, Z., P.C.B. Phillips (2002). A CUSUM test for cointegration using regression residuals. Journal of Econometrics, 108(1), 43-61.