A new model invariant CUSUM of squares-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 (Ordinary Least Squares) 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: *Cointegration, OLS estimators and residuals, covariance stationarity, cointegration, CUSUM of Squares-type statistic*

JEL classification: *C12, C22*

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.

Section 2 presents an extensive analysis of the specification and estimation methods available for the cointegrating regression model, under the assumption of the existence of at most one cointegration relationship among a set of k+1, $k \geq 1$, integrated variables in the general case where the regressors are possibly endogenous and the generating mechanism for the observations of all these variables may contain a deterministic component usually parameterized as a polynomial time trend function of a certain order. On the basis of this estimation results, we also review some of the more commonly used testing procedures for cointegration in terms of the dependence of its limiting distributions on the model specification and the structure of the data generating process for the observations of the stochastic regressors, particularly relevant in the case of deterministically trending integrated regressors.

Section 3 presents the main contribution of the paper in terms of a CUSUM of squares-type statistic, that is relatively simple to compute and only requires the use of the OLS residuals, for testing the null hypothesis of cointegration in a single-equation cointegrating regression model admitting an arbitrary number of integrated regressors and a very general form of the deterministic trend component, and 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. We conclude with a simple empirical illustration involving testing for cointegration in a low dimensional system given by an aggregate consumption function, using US quarterly macroeconomic data.

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

$$\begin{pmatrix} y_t \\ \mathbf{x}_{k,t} \end{pmatrix} = \begin{pmatrix} d_{0,t} \\ \mathbf{d}_{k,t} \end{pmatrix} + \begin{pmatrix} \eta_{0,t} \\ \boldsymbol{\eta}_{k,t} \end{pmatrix} \qquad (2.1)$$

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, $\boldsymbol{\Omega}_{\boldsymbol{\varepsilon}} = \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}} + \boldsymbol{\Lambda}_{\boldsymbol{\varepsilon}} + \boldsymbol{\Lambda}'_{\boldsymbol{\varepsilon}}$, $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}} = E[\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}'_t]$, $\boldsymbol{\Lambda}_{\boldsymbol{\varepsilon}} = \sum_{j=1}^{\infty} E[\boldsymbol{\varepsilon}_{t-j} \boldsymbol{\varepsilon}'_t]$, which is partitioned as $\boldsymbol{\Omega}_{\boldsymbol{\varepsilon}} = ([\omega_0^2, \boldsymbol{\omega}'_{k0}]' : [\boldsymbol{\omega}_{k0}, \boldsymbol{\Omega}_{kk}]')$, with $\boldsymbol{\omega}_{k0} = \boldsymbol{\omega}'_{0k}$. Then, if there exists a k-vector $\boldsymbol{\beta}_k$ such that

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

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 $\boldsymbol{\kappa} = (1, -\boldsymbol{\beta}'_k)'$, and $\boldsymbol{\beta}_k = \boldsymbol{\Omega}_{kk}^{-1} \boldsymbol{\omega}_{k0}$.[2] By combining (2.1) and (2.2) we get the regression equation

$$y_t = d_t + \boldsymbol{\beta}'_k \mathbf{x}_{k,t} + u_t,$$

based on the observed variables with $d_t = d_{0,t} - \boldsymbol{\beta}'_k \mathbf{d}_{k,t}$ the deterministic component. Obviating in this paper the case of possible cancellation of similar components in this linear combination of the deterministic trend functions (which is called cotrending), to obtain an operative version of (2.2) based on the generating mechanism of 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 $0 \leq p_1 < ... < p_{m+q}$, and $q \geq 0$, where, whenever the trend coefficient matrices $\mathbf{A}_{k,m} = (\boldsymbol{\alpha}_{1,m}, ..., \boldsymbol{\alpha}_{k,m})'$ and $\mathbf{A}_{k,q} = (\boldsymbol{\alpha}_{1,q}, ..., \boldsymbol{\alpha}_{k,q})'$ 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 $E[\mathbf{e}_t \mathbf{e}'_t] = \boldsymbol{\Sigma}_e > 0$ and (2+m)th-order

[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.

[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, $v_t = \Delta u_t = \boldsymbol{\kappa}' \Delta \boldsymbol{\eta}_t = \boldsymbol{\kappa}' \boldsymbol{\varepsilon}_t$, is zero, i.e. $\omega_v^2 = \sum_{j=-\infty}^{\infty} E[v_t v_{t-j}] = 0$. Given that $\det(\boldsymbol{\Omega}_{\boldsymbol{\varepsilon}}) = \omega_{0.k}^2 \det(\boldsymbol{\Omega}_{kk})$, where $\boldsymbol{\Omega}_{kk} > 0$ (excludes the possibility of additional cointegrating relationships among the stochastic trend components of the regressors), with $\omega_{0.k}^2 = \omega_0^2 - \boldsymbol{\omega}'_{k0} \boldsymbol{\Omega}_{kk}^{-1} \boldsymbol{\omega}_{k0} = \omega_0^2 (1 - \rho_{k0}^2)$, $\rho_{k0}^2 = \boldsymbol{\omega}'_{k0} (\omega_0^2 \boldsymbol{\Omega}_{kk})^{-1} \boldsymbol{\omega}_{k0}$, the conditional long-run variance of $\varepsilon_{0,t}$ given $\boldsymbol{\varepsilon}_{k,t}$, then $\omega_v^2 = \boldsymbol{\kappa}' \boldsymbol{\Omega}_{\boldsymbol{\varepsilon}} \boldsymbol{\kappa} = \omega_{0.k}^2 = 0$, and hence $\boldsymbol{\Omega}_{\boldsymbol{\varepsilon}}$ is a singular covariance matrix.

finite moment, $E[|\mathbf{e}_t|^{2+m}] = E[(\mathbf{e}_t'\mathbf{e}_t)^{2+m}] < \infty$ for some $m \geq 0$. 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 $\sum_{j=0}^{\infty} j^a \|\mathbf{D}_j\|^2 < \infty$, $a \geq 2$, 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 $0 \leq \alpha \leq 1$.

Next, we make some additional comments on these assumptions. First, from Assumption 2.1 we have to mention that the leading cases in many applications are giving by the inclusion of only a constant term in the estimable form of the cointegrating regression, i.e. $d_t = \alpha_0$, or a constant and a linear trend, $d_t = \alpha_0 + \alpha_1 t$, that correspond to the choice of m = 1 or 2, with $p_1 = 0$ and $p_2 = 1$, respectively. In any case, and under the general specification of a polynomial trend function component, it is necessary to introduce the scaling matrices $\mathbf{\Gamma}_{mm,n} = \text{diag}(n^{-p_1}, ..., n^{-p_m})$ and $\mathbf{\Gamma}_{qq,n} = \text{diag}(n^{-p_{m+1}}, ..., n^{-p_{m+q}})$ such that the components of the polynomial trend functions can be scaled to be bounded as $\boldsymbol{\tau}_{m,nt} = \mathbf{\Gamma}_{mm,n}\boldsymbol{\tau}_{m,t} \in (0,1]^m$ and $\boldsymbol{\tau}_{q,nt} = \mathbf{\Gamma}_{qq,n}\boldsymbol{\tau}_{q,t} \in (0,1]^q$, implying that certain functionals such as $n^{-1}\sum_{t=1}^{[nr]} \boldsymbol{\tau}_{m,nt}$ and $\mathbf{Q}_{mm,[nr]} = n^{-1}\sum_{t=1}^{[nr]} \boldsymbol{\tau}_{m,nt}\boldsymbol{\tau}_{m,nt}'$ have well defined limits, such as $n^{-1}\sum_{t=1}^{[nr]} \boldsymbol{\tau}_{m,nt} \to \int_0^r \boldsymbol{\tau}_m(s)ds$ and $\mathbf{Q}_{mm,[nr]} \to \mathbf{Q}_{mm}(r) = \int_0^r \boldsymbol{\tau}_m(s)\boldsymbol{\tau}_m'(s)ds$, with $[x]$ the integer part of x and similarly for $\boldsymbol{\tau}_{q,nt}$.

Second, from part (A) of Assumption 2.2 and the use of the well-known Beveridge-Nelson (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 $\mathbf{D}(1)n^{-1/2}\sum_{t=1}^{[nr]} \mathbf{e}_t \Rightarrow \mathbf{B}(r) = \mathbf{BM}(\mathbf{\Omega}_0)$, $0 \leq r \leq 1$ 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 $\mathbf{\Omega}_0 = \mathbf{D}(1)\Sigma_e\mathbf{D}(1)'$. Once established this result, and taking into account that $\sup_{0\leq r\leq 1}|(1/\sqrt{n})\sum_{t=1}^{[nr]}(\boldsymbol{\xi}_{0,t} - \mathbf{D}(1)\mathbf{e}_t)| \leq 2\max_{0\leq t\leq n}|(1/\sqrt{n})\tilde{\mathbf{e}}_t| = o_p(1)$, 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 $0 \leq \alpha < 1$ (that is, under the cointegration assumption), we can define the k+1-vector

[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 $m\to\infty$.

$$\boldsymbol{\xi}_t = \begin{pmatrix} u_t \\ \boldsymbol{\varepsilon}_{k,t} \end{pmatrix} = \begin{pmatrix} (1-\alpha L)^{-1}\upsilon_t \\ \boldsymbol{\varepsilon}_{k,t} \end{pmatrix} = \mathbf{C}(L)\mathbf{e}_t = \begin{pmatrix} \mathbf{c}_0'(L)\mathbf{e}_t \\ \mathbf{D}_k(L)\mathbf{e}_t \end{pmatrix} \tag{2.3}$$

with $\mathbf{c}_0(L) = (1-\alpha L)^{-1}\mathbf{d}_0(L)$, that also satisfy the same summability conditions as for the coefficients in $\mathbf{d}_0(L)$,[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 $\boldsymbol{\Omega}_{\boldsymbol{\xi}} = \mathbf{C}(1)\Sigma_e\mathbf{C}(1)'$, with

$$\mathbf{C}(1) = \begin{pmatrix} \mathbf{c}_0'(1) \\ \mathbf{D}_k(1) \end{pmatrix} = \begin{pmatrix} (1-\alpha)^{-1}\mathbf{d}_0'(1) \\ \mathbf{D}_k(1) \end{pmatrix}$$

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 long-run variance of u_t conditional on $\boldsymbol{\varepsilon}_{k,t}$, 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. The rest of this section is dedicated to complete the specification of the estimating cointegrating regression model and to the review the estimation of the model parameters with the final aim to obtain the residuals as the main tool used in the construction of many feasible statistics to test for cointegration.

2.1. The cointegrating regression model and estimation

Once stated 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

$$y_t = \boldsymbol{\alpha}_m'\boldsymbol{\tau}_{m,t} + \boldsymbol{\beta}_k'\mathbf{x}_{k,t} + u_t \tag{2.4}$$

where $\boldsymbol{\alpha}_m = \boldsymbol{\alpha}_{0,m} - \mathbf{A}_{k,m}'\boldsymbol{\beta}_k$, that only incorporates the common structure of the deterministic components of each series under the assumption that $-\mathbf{A}_{k,q}'\boldsymbol{\beta}_k = \mathbf{0}_q$. Next, defining the complete vector of all the trending regressors in (2.4) 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 decomposed as $\mathbf{x}_{k,t} = \mathbf{A}_{k,m}\boldsymbol{\Gamma}_{mm,n}^{-1}\boldsymbol{\tau}_{m,nt} + \boldsymbol{\Gamma}_{kk,n}^{-1}\mathbf{m}_{k,nt}$, with $\mathbf{m}_{k,nt} = \boldsymbol{\Gamma}_{kk,n}(\boldsymbol{\eta}_{k,t} + \mathbf{A}_{k,q}\boldsymbol{\tau}_{q,t})$, then

[4] By writing $\mathbf{D}(L)$ as $\mathbf{D}(L) = (\mathbf{d}_0(L), \mathbf{D}_k'(L))'$, 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 have $\sum_{j=0}^\infty j^a\|\mathbf{D}_j\| = \sum_{j=0}^\infty j^a Tr(\mathbf{d}_{0j}\mathbf{d}_{0j}') + \sum_{j=0}^\infty j^a Tr(\mathbf{D}_{kj}'\mathbf{D}_{kj})$. Then, by the recurrence relation $\mathbf{d}_{0j} = \mathbf{c}_{0,j} - \alpha\mathbf{c}_{0,j-1}$ for $j \geq 1$, the first term above can be decomposed as $\sum_{j=0}^\infty j^a Tr(\mathbf{d}_{0j}\mathbf{d}_{0j}') = (1+\alpha^2)\sum_{j=0}^\infty j^a Tr(\mathbf{c}_{0,j}\mathbf{c}_{0,j}')$ $+\alpha^2\sum_{j=0}^\infty Tr(\mathbf{c}_{0,j}\mathbf{c}_{0,j}') - \alpha\sum_{j=1}^\infty j^a Tr(\mathbf{c}_{0,j}\mathbf{c}_{0,j-1}' + \mathbf{c}_{0,j-1}\mathbf{c}_{0,j}')$, so that the summability condition on the coefficients $\mathbf{D}_j$, and hence for $\mathbf{d}_{0j}$, implies that of $\sum_{j=0}^\infty j^a Tr(\mathbf{c}_{0,j}\mathbf{c}_{0,j}')$.

a convenient choice for the weighting matrix $\boldsymbol{\Gamma}_{kk,n}$ is $\boldsymbol{\Gamma}_{kk,n} = \mathbf{W}_{kk,n}^{-1}\mathbf{C}_{kk}'$, so that $\mathbf{W}_n$ is given by

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

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 the scaling matrix $\boldsymbol{\Gamma}_{kk,n}$ is $\mathbf{W}_{kk,n} = \sqrt{n}\mathbf{I}_{kk}$ and $\mathbf{C}_{kk} = \mathbf{I}_{kk}$, so that

$$\mathbf{W}_n = \begin{pmatrix} \boldsymbol{\Gamma}_{mm,n}^{-1} & \mathbf{0}_{m,k} \\ \mathbf{A}_{k,m}\boldsymbol{\Gamma}_{mm,n}^{-1} & n^{1/2}\mathbf{I}_{k,k} \end{pmatrix} \tag{2.6}$$

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 $\boldsymbol{\Gamma}_{kk,n}$ given by

$$\boldsymbol{\Gamma}_{kk,n} = \begin{pmatrix} \boldsymbol{\Gamma}_{qq,n} & \mathbf{0}_{q,k-q} \\ \mathbf{0}_{k-q,q} & n^{-1/2}\mathbf{I}_{k-q,k-q} \end{pmatrix} \begin{pmatrix} \mathbf{C}_{k,q}' \\ \mathbf{C}_{k,k-q}' \end{pmatrix} \tag{2.7}$$

with $\mathbf{C}_{k,q} = \mathbf{A}_{k,q}(\mathbf{A}_{k,q}'\mathbf{A}_{k,q})^{-1}$, and $\mathbf{C}_{k,k-q} = \overline{\mathbf{A}}_{k,k-q}(\overline{\mathbf{A}}_{k,k-q}'\boldsymbol{\Omega}_{kk}\overline{\mathbf{A}}_{k,k-q})^{-1/2}$ where $\overline{\mathbf{A}}_{k,k-q}$ is a full rank $k\times(k-q)$ matrix which spans the null space of $\mathbf{A}_{k,q}$ such that $\overline{\mathbf{A}}_{k,k-q}'\mathbf{A}_{k,q} = \mathbf{0}_{k-q,q}$, with $\mathbf{C}_{kk} = (\mathbf{C}_{k,q}, \mathbf{C}_{k,k-q}) > 0$ and well-defined. In this case, the k-dimensional triangular array $\mathbf{m}_{k,nt}$ is now given by

$$\mathbf{m}_{k,nt} = \begin{pmatrix} \boldsymbol{\tau}_{q,nt} + n^{1/2}\boldsymbol{\Gamma}_{qq,n}\mathbf{C}_{k,q}'\boldsymbol{\eta}_{k,nt} \\ \mathbf{C}_{k,k-q}'\boldsymbol{\eta}_{k,nt} \end{pmatrix} \tag{2.8}$$

where $\boldsymbol{\tau}_{q,nt} = \boldsymbol{\Gamma}_{qq,n}\boldsymbol{\tau}_{q,t}$, with weak limit

$$\mathbf{m}_{k,n[nr]} \Rightarrow \mathbf{m}_k(r) = \begin{pmatrix} \boldsymbol{\tau}_q(r) \\ \mathbf{W}_{k-q}(r) \end{pmatrix} \tag{2.9}$$

where $\mathbf{W}_{k-q}(r) = \mathbf{C}_{k,k-q}'\mathbf{B}_k(r)$ and satisfies the following distributional equivalence

$$\mathbf{W}_{k-q}(r) = (\overline{\mathbf{A}}_{k,k-q}'\boldsymbol{\Omega}_{kk}\overline{\mathbf{A}}_{k,k-q})^{-1/2}\boldsymbol{\upsilon}_{k-q}(r) =^d \mathbf{BM}(\mathbf{I}_{k-q,k-q})$$

with $\boldsymbol{\upsilon}_{k-q}(r) = \overline{\mathbf{A}}_{k,k-q}'\mathbf{B}_k(r) = \overline{\mathbf{A}}_{k,k-q}'\boldsymbol{\Omega}_{kk}^{1/2}\mathbf{W}_k(r) =^d \mathbf{BM}(\overline{\mathbf{A}}_{k,k-q}'\boldsymbol{\Omega}_{kk}\overline{\mathbf{A}}_{k,k-q})$, so that the limit process $\mathbf{m}_k(r)$ in (2.9) is 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 for a convenient treatment of the asymptotic distributional aspects of the OLS estimators and residuals from (2.4) under a wide variety of situations concerning the underlying deterministic structure of the integrated regressors in (2.1) and its effects on the corresponding limiting distributions. Thus, given $\hat{\boldsymbol{\theta}}_n = (\sum_{t=1}^n \mathbf{m}_t\mathbf{m}_t')^{-1}\sum_{t=1}^n \mathbf{m}_t y_t = \boldsymbol{\theta} + (\sum_{t=1}^n \mathbf{m}_t\mathbf{m}_t')^{-1}\sum_{t=1}^n \mathbf{m}_t u_t$ as the OLS estimator of $\boldsymbol{\theta} = (\boldsymbol{\alpha}_m', \boldsymbol{\beta}_k')'$, we get that the scaled and normalized vector of OLS estimation errors from (2.4) is given by

$$\hat{\boldsymbol{\Theta}}_n = n^{\kappa}\mathbf{W}_n'(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) = \left(n^{-1}\sum_{t=1}^{n}\mathbf{m}_{nt}\mathbf{m}_{nt}' \right)^{-1} n^{-(1-\kappa)}\sum_{t=1}^{n}\mathbf{m}_{nt}u_t \tag{2.10}$$

where $\hat{\boldsymbol{\Theta}}_n = (\hat{\boldsymbol{\Theta}}_{m,n}', \hat{\boldsymbol{\Theta}}_{k,n}')$, so that the normalized and scaled OLS estimators of $\boldsymbol{\alpha}_m$ and $\boldsymbol{\beta}_k$ are $\hat{\boldsymbol{\Theta}}_{m,n} = n^{\kappa}\boldsymbol{\Gamma}_{mm,n}^{-1}[\hat{\boldsymbol{\alpha}}_{m,n} - \boldsymbol{\alpha}_m - \mathbf{A}_{k,m}'(\hat{\boldsymbol{\beta}}_{k,n} - \boldsymbol{\beta}_k)]$ and $\hat{\boldsymbol{\Theta}}_{k,n} = n^{\kappa}\boldsymbol{\Gamma}_{kk,n}^{-1}(\hat{\boldsymbol{\beta}}_{k,n} - \boldsymbol{\beta}_k)$ respectively, and with the exponent κ taking values $\pm 1/2$ 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 (2.10) is given by

$$n^{-1/2}\sum_{t=1}^{n}\mathbf{m}_{nt}u_t \Rightarrow \int_0^1 \mathbf{m}(r)dB_u(r) + \boldsymbol{\Phi} = \int_0^1 \begin{pmatrix} \boldsymbol{\tau}_m(r) \\ \mathbf{m}_k(r) \end{pmatrix} dB_u(r) + \begin{pmatrix} \mathbf{0}_m \\ \boldsymbol{\Phi}_{ku} \end{pmatrix} \tag{2.11}$$

where $\mathbf{m}_k(r) = \mathbf{B}_k(r)$ and $\boldsymbol{\Phi}_{ku} = \boldsymbol{\Delta}_{ku}$ when $\mathbf{A}_{k,q} = \mathbf{0}_{k,q}$, while that $\mathbf{m}_k(r)$ is as in (2.9) with $\boldsymbol{\Phi}_{ku} = (\mathbf{0}_q', \boldsymbol{\Delta}_{ku}'\mathbf{C}_{k,k-q})'$ in the case $\mathbf{A}_{k,q} \neq \mathbf{0}_{k,q}$. 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 (2.11) requires to know the trend properties of these variables through the limiting representation of the k-vector $\mathbf{m}_k(r)$. Despite these difficulties, another important result arising from (2.10) and (2.11) is the consistent estimation of all the model parameters in $\boldsymbol{\theta} = (\boldsymbol{\alpha}_m', \boldsymbol{\beta}_k')'$ under cointegration. Particularly, it is remarkable the superconsistent estimation, at the rate n, of the cointegrating vector $\boldsymbol{\beta}_k$ under cointegration, as can see from the representation

$$\begin{aligned}\hat{\boldsymbol{\Theta}}_{k,n} = n^{\kappa}\boldsymbol{\Gamma}_{kk,n}^{-1}(\hat{\boldsymbol{\beta}}_{k,n} - \boldsymbol{\beta}_k) &= \left(n^{-1}\sum_{t=1}^{n}\hat{\mathbf{m}}_{k,nt}\hat{\mathbf{m}}_{k,nt}' \right)^{-1} n^{-(1-\kappa)}\sum_{t=1}^{n}\hat{\mathbf{m}}_{k,nt}u_{t,m} \\ &= \hat{\mathbf{Q}}_{kk,n}^{-1} n^{-(1-\kappa)}\sum_{t=1}^{n}\hat{\mathbf{m}}_{k,nt}u_t\end{aligned} \tag{2.12}$$

with $\hat{\mathbf{m}}_{k,nt} = \mathbf{m}_{k,nt} - \mathbf{Q}_{km,n}\mathbf{Q}_{mm,n}^{-1}\boldsymbol{\tau}_{m,nt}$ and $u_{t,m} = u_t - n^{-\kappa}\boldsymbol{\tau}_{m,nt}'\mathbf{Q}_{mm,n}^{-1}n^{-(1-\kappa)}\sum_{j=1}^{n}\boldsymbol{\tau}_{m,nj}u_j$ the OLS detrended observations of $\mathbf{m}_{k,nt}$ and u_t, respectively, $\mathbf{Q}_{km,n} = n^{-1}\sum_{j=1}^{n}\mathbf{m}_{k,nj}\boldsymbol{\tau}_{m,nj}'$, and where the second equality in (2.12) comes from the orthogonality between $\hat{\mathbf{m}}_{k,nt}$ and $\boldsymbol{\tau}_{m,nt}$, i.e. $\sum_{t=1}^{n}\hat{\mathbf{m}}_{k,nt}\boldsymbol{\tau}_{m,nt}' = \mathbf{0}_{k,m}$. Given that $\hat{\mathbf{Q}}_{kk,n} \Rightarrow \hat{\mathbf{Q}}_{kk} = \int_0^1 \hat{\mathbf{m}}_k(r)\hat{\mathbf{m}}_k'(r)dr$, with $\hat{\mathbf{m}}_k(r) = \mathbf{m}_k(r) - \int_0^1 \mathbf{m}_k(s)\boldsymbol{\tau}_m'(s)ds\,\mathbf{Q}_{mm}^{-1}(1)\boldsymbol{\tau}_m(r)$ the detrended version of the limit process $\mathbf{m}_k(r)$, then under cointegration we get

$$n^{1/2}\boldsymbol{\Gamma}_{kk,n}^{-1}(\hat{\boldsymbol{\beta}}_{k,n} - \boldsymbol{\beta}_k) \Rightarrow \hat{\mathbf{Q}}_{kk}^{-1}\left(\int_0^1 \hat{\mathbf{m}}_k(r)dB_u(r) + \boldsymbol{\Phi}_{ku} \right) \tag{2.13}$$

in the general case, with $n(\hat{\boldsymbol{\beta}}_{k,n} - \boldsymbol{\beta}_k) \Rightarrow \hat{\mathbf{Q}}_{kk}^{-1}(\int_0^1 \hat{\mathbf{B}}_k(r)dB_u(r) + \boldsymbol{\Delta}_{ku})$ when $\mathbf{A}_{k,q} = \mathbf{0}_{k,q}$ and

$$\mathbf{C}_{kk}^{\prime -1}\begin{pmatrix} n^{1/2}\boldsymbol{\Gamma}_{qq,n}^{-1}(\hat{\boldsymbol{\beta}}_{q,n}-\boldsymbol{\beta}_q) \\ n(\hat{\boldsymbol{\beta}}_{k-q,n}-\boldsymbol{\beta}_{k-q}) \end{pmatrix} \Rightarrow \hat{\mathbf{Q}}_{kk}^{-1}\left(\int_0^1 \hat{\mathbf{m}}_k(r)dB_u(r)+\begin{pmatrix} \mathbf{0}_q \\ \mathbf{C}_{k,k-q}^{\prime}\boldsymbol{\Delta}_{ku} \end{pmatrix}\right)$$

if $\mathbf{A}_{k,q} \neq \mathbf{0}_{k,q}$, where $\hat{\boldsymbol{\beta}}_{k,n}$ has been partitioned as $\hat{\boldsymbol{\beta}}_{k,n}=(\hat{\boldsymbol{\beta}}_{q,n}^{\prime},\hat{\boldsymbol{\beta}}_{k-q,n}^{\prime})^{\prime}$. In the case of no cointegration when the regression error is unit-root nonstationary, also known as a spurious regression, and under our assumption on the generating mechanism of the dependent variable in the cointegrating regression such as $y_t=\boldsymbol{\alpha}_{0,m}^{\prime}\boldsymbol{\tau}_{m,t}+\eta_{0,t}$, then equation (2.10) can be rewritten as

$$n^{\kappa}\mathbf{W}_n^{\prime}(\hat{\boldsymbol{\theta}}_n-\boldsymbol{\theta})=-n^{\kappa}\mathbf{W}_n^{\prime}\boldsymbol{\theta}+\left(n^{-1}\sum_{t=1}^{n}\mathbf{m}_{nt}\mathbf{m}_{nt}^{\prime}\right)^{-1}n^{-(1-\kappa)}\sum_{t=1}^{n}\mathbf{m}_{nt}(\eta_{0,t}+\boldsymbol{\tau}_{m,nt}^{\prime}\boldsymbol{\Gamma}_{mm,n}^{-1}\boldsymbol{\alpha}_{0,m})$$

$$=-n^{\kappa}\left\{\mathbf{W}_n^{\prime}\boldsymbol{\theta}+\begin{pmatrix}\boldsymbol{\Gamma}_{mm,n}^{-1}\boldsymbol{\alpha}_{0,m}\\ \mathbf{0}_k\end{pmatrix}\right\}+\left(n^{-1}\sum_{t=1}^{n}\mathbf{m}_{nt}\mathbf{m}_{nt}^{\prime}\right)^{-1}n^{-(1-\kappa)}\sum_{t=1}^{n}\mathbf{m}_{nt}\eta_{0,t}$$

so that

$$n^{\kappa}\mathbf{W}_n^{\prime}\begin{pmatrix}\hat{\boldsymbol{\alpha}}_{m,n}-\boldsymbol{\alpha}_{0m}\\ \hat{\boldsymbol{\beta}}_{k,n}\end{pmatrix}=\left(n^{-1}\sum_{t=1}^{n}\mathbf{m}_{nt}\mathbf{m}_{nt}^{\prime}\right)^{-1}n^{-(1-\kappa)}\sum_{t=1}^{n}\mathbf{m}_{nt}\eta_{0,t} \quad (2.14)$$

where $n^{-3/2}\sum_{t=1}^{n}\mathbf{m}_{nt}\eta_{0,t}\Rightarrow\int_0^1\mathbf{m}(r)B_0(r)dr$, with $B_0(r)$ the weak limit of $n^{-1/2}\eta_{0,[nr]}$ when $\kappa=-1/2$. Alternatively, given that u_t can also be written as $u_t=\eta_{0,t}-\boldsymbol{\beta}_k^{\prime}\mathbf{m}_{k,t}$ and

$$n^{-(1-\kappa)}\sum_{t=1}^{n}\hat{\mathbf{m}}_{k,nt}u_t=n^{-(1-\kappa)}\sum_{t=1}^{n}\hat{\mathbf{m}}_{k,nt}\eta_{0,t}-n^{\kappa}\hat{\mathbf{Q}}_{kk,n}\boldsymbol{\Gamma}_{kk,n}^{-1}\boldsymbol{\beta}_k$$

then (2.12) is now given by

$$n^{\kappa}\boldsymbol{\Gamma}_{kk,n}^{-1}\hat{\boldsymbol{\beta}}_{k,n}=\hat{\mathbf{Q}}_{kk,n}^{-1}n^{-(1-\kappa)}\sum_{t=1}^{n}\hat{\mathbf{m}}_{k,nt}\eta_{0,t}=\hat{\mathbf{Q}}_{kk,n}^{-1}n^{-(1-\kappa)}\sum_{t=1}^{n}\hat{\mathbf{m}}_{k,nt}\hat{\eta}_{0,t} \quad (2.15)$$

with $\hat{\eta}_{0,t}=\eta_{0,t}-\boldsymbol{\tau}_{m,nt}^{\prime}\mathbf{Q}_{mm,n}^{-1}n^{-1}\sum_{j=1}^{n}\boldsymbol{\tau}_{m,nj}\eta_{0,j}$, which implies inconsistent estimation of $\boldsymbol{\beta}_k$ when $\mathbf{A}_{k,q}=\mathbf{0}_{k,q}$, $\hat{\boldsymbol{\beta}}_{k,n}=O_p(1)$, while that if $\mathbf{A}_{k,q}\neq\mathbf{0}_{k,q}$ then we get

$$\begin{pmatrix}n^{-1/2}\boldsymbol{\Gamma}_{qq,n}^{-1}\hat{\boldsymbol{\beta}}_{q,n}\\ \hat{\boldsymbol{\beta}}_{k-q,n}\end{pmatrix}=\mathbf{C}_{kk}^{\prime}\hat{\mathbf{Q}}_{kk,n}^{-1}n^{-(1-\kappa)}\sum_{t=1}^{n}\hat{\mathbf{m}}_{k,nt}\eta_{0,t}$$

indicating that in this general case, even under no cointegration, some elements of $\boldsymbol{\beta}_k$ can be still consistently estimated although we are dealing with a nonsense regression as $\boldsymbol{\beta}_k=\mathbf{0}_k$ by definition. As shown in Hansen (1992a, b), similar results and conclusions are attributed to asymptotically efficient estimates obtained when using, e.g., the Fully Modified OLS (FM-OLS) estimator proposed by Phillips and Hansen (1990) as a way to simultaneously correct for the two sources of finite-sample bias appearing in the last terms of (2.11) and $\int_0^1\mathbf{m}(r)dB_u(r)=\int_0^1\mathbf{m}(r)dB_{u.k}(r)+\int_0^1\mathbf{m}(r)d\mathbf{B}_k^{\prime}(r)\cdot\boldsymbol{\Omega}_{kk}^{-1}\boldsymbol{\omega}_{ku}$ caused by the endogeneity of the integrated regressors. Motivated by the fact that $n^{-1/2}\sum_{t=1}^{n}\boldsymbol{\eta}_{k,nt}z_t\Rightarrow\int_0^1\mathbf{B}_k(r)dB_{u.k}(r)+\boldsymbol{\Delta}_{ku}^{+}$ under cointegration, where $z_t=u_t-\boldsymbol{\gamma}_{ku}^{\prime}\boldsymbol{\varepsilon}_{k,t}$ and $\boldsymbol{\Delta}_{ku}^{+}=\boldsymbol{\Delta}_{ku}-\boldsymbol{\Delta}_{kk}\boldsymbol{\gamma}_{ku}$, with $\boldsymbol{\gamma}_{ku}=\boldsymbol{\Omega}_{kk}^{-1}\boldsymbol{\omega}_{ku}$ and $\boldsymbol{\Delta}_{kk}=\sum_{j=0}^{\infty}E[\boldsymbol{\varepsilon}_{k,t-j}\boldsymbol{\varepsilon}_{k,t}^{\prime}]$, the FM-OLS estimator of the model parameters in (2.4) is given by

$$\hat{\boldsymbol{\theta}}_n^{+}=\begin{pmatrix}\hat{\boldsymbol{\alpha}}_{m,n}^{+}\\ \hat{\boldsymbol{\beta}}_{k,n}^{+}\end{pmatrix}=\left(\sum_{t=1}^{n}\mathbf{m}_t\mathbf{m}_t^{\prime}\right)^{-1}\left(\sum_{t=1}^{n}\mathbf{m}_t y_t^{+}-n\begin{pmatrix}\mathbf{0}_m\\ \hat{\boldsymbol{\Delta}}_{ku,n}^{+}\end{pmatrix}\right)$$

where y_t^{+} are modified observations of the dependent variable defined as

$y_t^+ = y_t - \hat{\boldsymbol{\gamma}}'_{ku,n}\hat{\mathbf{z}}_{k,t} = \mathbf{m}'_t\boldsymbol{\theta} + u_t^+$, with $u_t^+ = u_t - \hat{\boldsymbol{\gamma}}'_{ku,n}\hat{\mathbf{z}}_{k,t}$, where $\hat{\boldsymbol{\gamma}}_{ku,n}$ and $\hat{\boldsymbol{\Delta}}^+_{ku,n}$ are consistent estimates of $\boldsymbol{\gamma}_{ku}$ and $\boldsymbol{\Delta}^+_{ku}$ under the cointegration assumption. The usual choice are kernel-type plug-in estimators $\hat{\boldsymbol{\gamma}}_{ku,n} = \hat{\boldsymbol{\Omega}}^{-1}_{kk,n}\hat{\boldsymbol{\omega}}_{ku,n}$ and $\hat{\boldsymbol{\Delta}}^+_{ku,n} = \hat{\boldsymbol{\Delta}}_{ku,n} - \hat{\boldsymbol{\Delta}}_{kk,n}\hat{\boldsymbol{\gamma}}_{ku,n}$, where $\hat{\boldsymbol{\Omega}}_{kk,n} = \hat{\boldsymbol{\Delta}}_{kk,n} + \hat{\boldsymbol{\Lambda}}'_{kk,n} = \sum_{j=0}^{n-1} w(jq_n^{-1})\hat{\mathbf{G}}_{kk,n}(j) + \sum_{j=1}^{n-1} w(jq_n^{-1})\hat{\mathbf{G}}_{kk,n}(-j)$, with $\hat{\mathbf{G}}_{kk,n}(j) = n^{-1}\sum_{t=j+1}^{n}\hat{\mathbf{z}}_{k,t-j}\hat{\mathbf{z}}'_{k,t}$ and $\hat{\mathbf{G}}_{kk,n}(-j) = \hat{\mathbf{G}}'_{kk,n}(j)$, while $\hat{\boldsymbol{\omega}}_{ku,n} = \hat{\boldsymbol{\Delta}}_{ku,n} + \hat{\boldsymbol{\Lambda}}'_{uk,n}$ where $\hat{\boldsymbol{\Delta}}_{ku,n} = \sum_{j=0}^{n-1} w(jq_n^{-1})\hat{\mathbf{g}}_{ku,n}(j)$ and $\hat{\boldsymbol{\Lambda}}'_{uk,n} = \sum_{j=1}^{n-1} w(jq_n^{-1})\hat{\mathbf{g}}_{ku,n}(-j)$, based on the sample serial covariances $\hat{\mathbf{g}}_{ku,n}(j) = n^{-1}\sum_{t=j+1}^{n}\hat{\mathbf{z}}_{k,t-j}\hat{u}_t$ and $\hat{\mathbf{g}}_{ku,n}(-j) = n^{-1}\sum_{t=j+1}^{n}\hat{\mathbf{z}}_{k,t}\hat{u}_{t-j}$, although $\hat{\boldsymbol{\Delta}}^+_{ku,n}$ can also be computed as $\hat{\boldsymbol{\Delta}}^+_{ku,n} = \sum_{j=0}^{n-1} w(jq_n^{-1})n^{-1}\sum_{t=j+1}^{n}\hat{\mathbf{z}}_{k,t-j}\hat{z}_t$, where $\hat{z}_t = \hat{u}_t - \hat{\boldsymbol{\gamma}}'_{ku,n}\hat{\mathbf{z}}_{k,t}$. Both y_t^+ and these estimators are based on the k-vector $\hat{\mathbf{z}}_{k,t}$ that represents the sequence of OLS residuals in the multivariate regression $\mathbf{z}_{k,t} = \Delta\mathbf{x}_{k,t} = \mathbf{A}_k\Delta\boldsymbol{\tau}_t + \boldsymbol{\varepsilon}_{k,t} = \mathbf{B}_k\boldsymbol{\tau}_t + \boldsymbol{\varepsilon}_{k,t}$, with $\mathbf{B}_k = (\mathbf{A}_k, \mathbf{0}_k)$, $\mathbf{A}_k = (\mathbf{A}_{k,m}, \mathbf{A}_{k,q})$, and $\boldsymbol{\tau}_t = (\boldsymbol{\tau}'_{m,t}, \boldsymbol{\tau}'_{q,t})'$ in the case where the mechanism generating the stochastic regressors do contain deterministic components, while that when $\mathbf{d}_{k,t} = \mathbf{0}_k$, or $\mathbf{d}_{k,t} = \mathbf{A}_{k,1}\tau_{0,t}$ when $m+q = 1$ with $p_1 = 0$, then $\mathbf{z}_{k,t} = \boldsymbol{\varepsilon}_{k,t}$. Otherwise, for $m+q > 1$, the k-vector of OLS residuals is given by $\hat{\mathbf{z}}_{k,t} = \boldsymbol{\varepsilon}_{k,t} + \mathbf{F}_{k,nt}$, with $\mathbf{F}_{k,nt} = -n^{-1/2}(n^{-1/2}\sum_{j=1}^{n}\boldsymbol{\varepsilon}_{k,j}\boldsymbol{\tau}'_{nj})\mathbf{Q}_n^{-1}\boldsymbol{\tau}_{nt}$ or, alternatively as $\mathbf{F}_{k,nt} = -n^{1/2}(n^{-1}\sum_{j=1}^{n}\boldsymbol{\eta}_{k,nj}\boldsymbol{\tau}'_{nj})\mathbf{Q}_n^{-1}\Delta\boldsymbol{\tau}_{nt} = -n^{-1/2}(n^{-1}\sum_{j=1}^{n}\boldsymbol{\eta}_{k,nj}\boldsymbol{\tau}'_{nj})\mathbf{Q}_n^{-1}\mathbf{d}_{nt}$, given that $\Delta\boldsymbol{\tau}_{nt} = n^{-1}\mathbf{d}_{nt}$, when based on the first difference of the k-vector of OLS residuals computed from the multivariate regression $\mathbf{x}_{k,t} = \mathbf{A}_k\boldsymbol{\tau}_t + \boldsymbol{\eta}_{k,t}$, with $\mathbf{Q}_n = n^{-1}\sum_{j=1}^{n}\boldsymbol{\tau}_{nj}\boldsymbol{\tau}'_{nj}$. In both cases, it can be shown that $\mathbf{F}_{k,nt} = O_p(n^{-1/2})$ implying the consistent estimation of $\boldsymbol{\varepsilon}_{k,t}$, the sequence of error terms driving the stochastic trend components of the integrated regressors. Similar to the OLS case, the scaled and normalized FM-OLS estimation error can be written as

$$\hat{\boldsymbol{\Theta}}_n^+ = n^{\kappa}\mathbf{W}'_n(\hat{\boldsymbol{\theta}}_n^+ - \boldsymbol{\theta}) = \left(n^{-1}\sum_{t=1}^{n}\mathbf{m}_{nt}\mathbf{m}'_{nt}\right)^{-1}\left(n^{-(1-\kappa)}\sum_{t=1}^{n}\mathbf{m}_{nt}u_t^+ - n^{\kappa}\mathbf{W}_n^{-1}\begin{pmatrix}\mathbf{0}_m\\ \hat{\boldsymbol{\Delta}}^+_{ku,n}\end{pmatrix}\right)$$

where $u_t^+ = u_t - \boldsymbol{\gamma}'_{ku}\boldsymbol{\varepsilon}_{k,t} - (\hat{\boldsymbol{\gamma}}'_{ku,n} - \boldsymbol{\gamma}'_{ku})\boldsymbol{\varepsilon}_{k,t} - \hat{\boldsymbol{\gamma}}'_{ku,n}\mathbf{F}_{k,nt}$, while that for the last term we have

$$n^{\kappa}\mathbf{W}_n^{-1}\begin{pmatrix}\mathbf{0}_m\\ \hat{\boldsymbol{\Delta}}^+_{ku,n}\end{pmatrix} = \begin{pmatrix}\mathbf{0}_m\\ n^{\kappa}\boldsymbol{\Gamma}_{kk,n}\hat{\boldsymbol{\Delta}}^+_{ku,n}\end{pmatrix}$$

with $\mathbf{W}_n$ the weighting matrix defined in (2.5), where $n^{\kappa}\boldsymbol{\Gamma}_{kk,n}\hat{\boldsymbol{\Delta}}^+_{ku,n} = \hat{\boldsymbol{\Delta}}^+_{ku,n}$ in the case of cointegration (with $\kappa = 1/2$) when $\mathbf{A}_{k,q} = \mathbf{0}_{k,q}$, and

$$n^{\kappa}\boldsymbol{\Gamma}_{kk,n}\hat{\boldsymbol{\Delta}}^+_{ku,n} = \begin{pmatrix}n^{\kappa}\boldsymbol{\Gamma}_{qq,n}\mathbf{C}'_{k,q}\hat{\boldsymbol{\Delta}}^+_{ku,n}\\ n^{-1/2+\kappa}\mathbf{C}'_{k,k-q}\hat{\boldsymbol{\Delta}}^+_{ku,n}\end{pmatrix}$$

when $\mathbf{A}_{k,q} \neq \mathbf{0}_{k,q}$, where the first term is asymptotically negligible both under cointegration and no cointegration and the second term is just $\mathbf{C}'_{k,k-q}\hat{\boldsymbol{\Delta}}^+_{ku,n}$ under cointegration which allows to cancel the bias term $\boldsymbol{\Phi}_{ku}$ in (2.11) as $n\to\infty$. Finally, given that $\hat{\boldsymbol{\gamma}}_{ku,n} - \boldsymbol{\gamma}_{ku} = \hat{\boldsymbol{\Omega}}^{-1}_{kk,n}[\hat{\boldsymbol{\omega}}_{ku,n} - \boldsymbol{\omega}_{ku} - (\hat{\boldsymbol{\Omega}}_{kk,n} - \boldsymbol{\Omega}_{kk})\boldsymbol{\Omega}^{-1}_{kk}\boldsymbol{\omega}_{ku}]$, with $\hat{\boldsymbol{\Omega}}_{kk,n} - \boldsymbol{\Omega}_{kk} = O_p(n^{-1/2})$ and

$\hat{\boldsymbol{\omega}}_{ku,n} - \boldsymbol{\omega}_{ku} = O_p(n^{-1/2})$ under cointegration and the correct detrending of the regressors, then we have $u_t^+ = z_t + O_p(n^{-1/2})$ which provides the desired limit result free of nuisance parameters. As a by-product of this estimation, the sequence of FM-OLS residuals, defined as $\hat{u}_t^+ = y_t^+ - \mathbf{m}_t \hat{\boldsymbol{\theta}}_n^+ = u_t^+ - n^{-\kappa}\mathbf{m}_{nt}\hat{\boldsymbol{\Theta}}_n^+$, can also be written as $\hat{u}_t^+ = z_t + O_p(n^{-1/2})$ under cointegration, where $z_t = (1, -\boldsymbol{\gamma}'_{ku})\boldsymbol{\xi}_t$ with $\boldsymbol{\xi}_t = (u_t, \boldsymbol{\varepsilon}'_{k,t})'$ so that the residual covariance of order h is decomposed as

$$n^{-1}\sum_{t=h+1}^{n} \hat{u}_{t-h}^+ \hat{u}_t^+ = (1, -\boldsymbol{\gamma}'_{ku}) n^{-1} \sum_{t=h+1}^{n} \boldsymbol{\xi}_{t-h}\boldsymbol{\xi}'_t \begin{pmatrix} 1 \\ -\boldsymbol{\gamma}_{ku} \end{pmatrix} + O_p(n^{-1/2}) \tag{2.16}$$

and hence the kernel-type estimator of the long-run variance of u_t^+ based on these residuals can be written as

$$\hat{\omega}^2_{u^+,n}(q_n) = (1, -\boldsymbol{\gamma}'_{ku})\boldsymbol{\Omega}_{\xi,n}(q_n)\begin{pmatrix} 1 \\ -\boldsymbol{\gamma}_{ku} \end{pmatrix} + O_p(n^{-1/2})$$

where

$$\boldsymbol{\Omega}_{\xi,n}(q_n) = \begin{pmatrix} \omega^2_{u,n} & \boldsymbol{\omega}_{uk,n} \\ \boldsymbol{\omega}_{ku,n} & \boldsymbol{\Omega}_{kk,n} \end{pmatrix} = n^{-1}\sum_{t=1}^{n}\boldsymbol{\xi}_t\boldsymbol{\xi}'_t + \sum_{h=1}^{n-1} w(h/q_n) n^{-1} \sum_{t=h+1}^{n} (\boldsymbol{\xi}_{t-h}\boldsymbol{\xi}'_t + \boldsymbol{\xi}_t\boldsymbol{\xi}'_{t-h})$$

which gives $\hat{\omega}^2_{u^+,n}(q_n) = \omega^2_{u.k,n} + (\boldsymbol{\omega}_{uk,n} - \boldsymbol{\gamma}'_{ku}\boldsymbol{\Omega}_{kk,n})\boldsymbol{\Omega}^{-1}_{kk,n}(\boldsymbol{\omega}_{ku,n} - \boldsymbol{\Omega}_{kk,n}\boldsymbol{\gamma}_{ku}) + O_p(n^{-1/2})$, where $\omega^2_{u.k,n} = \omega^2_{u,n} - \boldsymbol{\omega}_{uk,n}\boldsymbol{\Omega}^{-1}_{kk,n}\boldsymbol{\omega}_{ku,n}$, so that $\boldsymbol{\omega}_{ku,n} - \boldsymbol{\Omega}_{kk,n}\boldsymbol{\gamma}_{ku} \to^p \mathbf{0}_k$ and hence $\hat{\omega}^2_{u^+,n}(q_n) \to^p \omega^2_{u.k}$, the long-run variance of u_t conditional on $\boldsymbol{\varepsilon}_{k,t}$, under proper choice of the bandwidth, $q_n = o(n^{1/2})$.

Some other alternative estimation methods frequently used in practical applications, that also produce asymptotically efficient and equivalent results under proper choice of the required tuning parameters, are the canonical cointegrating regression (CCR) estimator by Park (1992), and the dynamic OLS (DOLS) estimator proposed by Phillips and Loretan (1991), Saikkonen (1991), and Stock and Watson (1993). The CCR estimation is similar in spirit to the FM-OLS estimation procedure but based on semiparametric transformations both of the dependent variable and the stochastic regressors in the cointegrating regression, while the DOLS estimator is based on the estimation of a dynamic version of the cointegrating regression model obtained by the addition of a number of leads and lags of $\Delta\boldsymbol{\eta}_{k,t} = \boldsymbol{\varepsilon}_{k,t}$ in the case where it is observed when $\mathbf{d}_{k,t} = \mathbf{0}_k$.[5]
Without going in further analysis of these last alternative estimation methods, it is important to remind that its usefulness crucially depends on the knowledge of the deterministic component which drive $\mathbf{x}_{k,t}$ and the proper choice of different tuning parameters conditioning its performance in finite samples that could substantially differ from what expected asymptotically in some important situations. For some studies evaluating these finite-sample properties see, e.g., Gonzalo (1994), Montalvo (1995), and more recently Kurozumi and Hayakawa (2009) and the references therein. Also, for a more complete analysis of the properties of the FM-OLS estimator see, e.g., Phillips

[5] Formally, under some regularity conditions, the regression error term u_t can be expressed as $u_t = r_t + w_t$, where $w_t = \sum_{j=-\infty}^{\infty} \boldsymbol{\pi}'_{k,j}\boldsymbol{\varepsilon}_{k,t-j}$ with $\sum_{j=-\infty}^{\infty} \|\boldsymbol{\pi}_{k,j}\|^2 < \infty$ and $E[\boldsymbol{\varepsilon}_{k,t-j} r_t] = \mathbf{0}_k$ for all j = 0, ±1, ±2, … Writing $u_t = \sum_{j=-q}^{q} \boldsymbol{\pi}'_{k,j}\boldsymbol{\varepsilon}_{k,t-j} + r_t(q)$, with $r_t(q) = r_t + \sum_{|j|>q} \boldsymbol{\pi}'_{k,j}\boldsymbol{\varepsilon}_{k,t-j}$, then the augmented version of the cointegrating regression model is $y_t = \boldsymbol{\theta}'\mathbf{m}_t + \sum_{j=-q}^{q} \boldsymbol{\pi}'_{k,j}\hat{\mathbf{z}}_{k,t-j} + r_t(q) - \sum_{j=-q}^{q} \boldsymbol{\pi}'_{k,j}\mathbf{F}_{k,nt-j}$, with $\hat{\mathbf{z}}_{k,t-j}$ and $\mathbf{F}_{k,nt-j}$ as defined before.

(1995) in the cointegrating regression model and its extension to the estimation of a vector autoregression with some unit roots.
To complete this section, we consider two additional estimation methods recently proposed that are mainly characterized by relying on fewer requirements thus making easiest the computation of the estimates.
First, although not designed to deal with the specification of the cointegrating regression considered here, it is worth to consider the so-called AIV estimator proposed by Harris, et.al. (2002, 2003) in the context of a generalized version of the heteroskedastic cointegration model first introduced by Hansen (1992c). This estimator utilize an IV technique based on $\mathbf{m}_{t-s} = (\boldsymbol{\tau}'_{m,t-s}, \mathbf{x}'_{k,t-s})'$, with $s > 0$, as an instrument designed to obtain consistent estimates of the model parameters when the regression error term can contain a certain type of highly persistent component only with a proper choice of the lag parameter s. This AIV estimator is then given by

$$\hat{\boldsymbol{\theta}}_n(s) = \left(\sum_{t=s+1}^{n} \mathbf{m}_{t-s}\mathbf{m}'_t \right)^{-1} \sum_{t=s+1}^{n} \mathbf{m}_{t-s} y_t$$

with associated scaled and normalized estimation error

$$\hat{\boldsymbol{\Theta}}_n(s) = n^{\kappa}\mathbf{W}'_n(\hat{\boldsymbol{\theta}}_n(s) - \boldsymbol{\theta}) = \left(n^{-1} \sum_{t=s+1}^{n} \mathbf{m}_{n(t-s)}\mathbf{m}'_{nt} \right)^{-1} n^{-(1-\kappa)} \sum_{t=s+1}^{n} \mathbf{m}_{n(t-s)} u_t$$

From (2.3), the standard BN decomposition of the linear process describing the generating mechanism of the regression error term under cointegration, $u_t = \mathbf{c}'_0(L)\mathbf{e}_t$, of the form $u_t = u_{0,t} - \Delta\tilde{u}_t$, with $u_{0,t} = \mathbf{c}'_0(1)\mathbf{e}_t$, $n^{-1/2}\sum_{t=1}^{[nr]} u_{0,t} \Rightarrow B_u(r)$ and $\tilde{u}_t = \tilde{\mathbf{c}}'_0(L)\mathbf{e}_t$, allows to decompose the second term in the right hand side above as

$$\begin{aligned} n^{-(1-\kappa)} \sum_{t=s+1}^{n} \mathbf{m}_{n(t-s)} u_t &= n^{-1/2}\sum_{t=1}^{n-s} \mathbf{m}_{nt} u_{0,t+s} + n^{-1/2}\sum_{t=1}^{n-s} (\mathbf{m}_{n(t+1)} - \mathbf{m}_{nt})\tilde{u}_{t+s} \\ &\quad + n^{-1/2}(\mathbf{m}_{n1}\tilde{u}_s - \mathbf{m}_{n(n-s+1)}\tilde{u}_n) \end{aligned}$$

where the last term is $O_p(n^{-1/2})$, while that for the first term we can write

$$\begin{aligned} & n^{-1/2}\sum_{t=1}^{n-s} (\mathbf{m}_{n(t+1)} - \mathbf{m}_{nt})\tilde{u}_{t+s} \\ &= n^{-1/2}\sum_{t=1}^{n-s} \begin{pmatrix} n^{-1}\mathbf{d}_{m,nt} \\ \boldsymbol{\Gamma}_{kk,n}(\boldsymbol{\varepsilon}_{k,t+1} + \mathbf{A}_{k,q}\boldsymbol{\Gamma}^{-1}_{qq,n} n^{-1}\mathbf{d}_{q,nt}) \end{pmatrix} \tilde{u}_{t+s} \\ &= \begin{pmatrix} n^{-1}\left(n^{-1/2}\sum_{t=1}^{n-s} \mathbf{d}_{m,nt}\tilde{u}_{t+s} \right) \\ \boldsymbol{\Gamma}_{kk,n}\left\{ n^{1/2}\left(n^{-1}\sum_{t=1}^{n-s} \boldsymbol{\varepsilon}_{k,t+1}\tilde{u}_{t+s} \right) + n^{-1}\mathbf{A}_{k,q}\boldsymbol{\Gamma}^{-1}_{qq,n}\left(n^{-1/2}\sum_{t=1}^{n-s} \mathbf{d}_{q,nt}\tilde{u}_{t+s} \right) \right\} \end{pmatrix} \end{aligned}$$

where we have used the fact that $\Delta\boldsymbol{\tau}_{m,n(t+1)} = n^{-1}\mathbf{d}_{m,nt}$ and $\Delta\boldsymbol{\tau}_{q,n(t+1)} = n^{-1}\mathbf{d}_{q,nt}$, with $\mathbf{d}_{m,nt}$ collecting terms of the form $(t/n)^{p_j-1}$ for j = 1, .., m such that $\mathbf{d}_{m,nt} \to (r^{p_1-1},\dots,r^{p_m-1})'$ uniformly in r and similarly for $\mathbf{d}_{q,nt}$ (see Hansen (1992a), equation (A.1)). By Theorem 1 and Lemma 2 in Harris, et.al. (2003), with $s\to\infty$ at least as fast as $n^{1/a}$, such as $s = s_n = o(n^b)$ $1/a \le b < 1$, $a \ge 2$ (see Assumption 2.2(A)), then $E[\boldsymbol{\varepsilon}_{k,t+1}\tilde{u}_{t+s}] \to \mathbf{0}_k$ when $s\to\infty$, and hence $n^{-1}\sum_{t=1}^{n-s} \boldsymbol{\varepsilon}_{k,t+1}\tilde{u}_{t+s} = O_p(n^{-1/2})$ given that $n^{-1/2}\sum_{t=1}^{n-s} \boldsymbol{\varepsilon}_{k,t+1}\tilde{u}_{t+s}$ weakly converges to a Brownian process (see Theorem 3 in Harris, et.al. (2003)), implying that all the terms above are asymptotically negligible in the case of standard stationary cointegration, i.e. $n^{-1/2}\sum_{t=1}^{n-s}(\mathbf{m}_{n(t+1)}-\mathbf{m}_{nt})\tilde{u}_{t+s}=o_p(1)$ as $n\to\infty$. To complete the result, as we can write $n^{-1/2}\sum_{t=1}^{n-s}\mathbf{m}_{nt}u_{0,t+s}=n^{-1/2}\sum_{t=1}^{n-s}\mathbf{m}_{nt}u_{0,t}-n^{-1/2}\sum_{t=1}^{n-s}(\mathbf{m}_{n,t+s}-\mathbf{m}_{nt})u_{0,t+s}+O_p(s/n^{1/2})$, then we get $n^{-1/2}\sum_{t=1}^{n-s}\mathbf{m}_{nt}u_{0,t+s}\Rightarrow\int_0^1\mathbf{m}(r)dB_u(s)$, implying that under standard stationary cointegration the use of the AIV estimator allows to eliminate the bias term $\mathbf{\Phi}_{ku}$ appearing in the limiting distribution in (2.11), but does not allows to completely correct for the endogeneity of the regressors.

Second, Vogelsang and Wagner (2014) propose the so-called Integrated Modified OLS (IM-OLS) estimator which is based on the OLS estimation of a simple modification, free of tuning parameters, of all the regressors appearing in the cointegrating regression model (2.4). The proposed transformation consists on taking partial sums of the variables in (2.4) and augmenting the resulting specification with the addition of the original observations of the stochastic regressors in such a way that the integrated modified version of the cointegrating regression is of the form

$$S_t=\boldsymbol{\alpha}'_m\mathbf{S}_{m,t}+\boldsymbol{\beta}'_k\mathbf{S}_{k,t}+\boldsymbol{\gamma}'_k\mathbf{x}_{k,t}+Z_t \tag{2.17}$$

where $S_t=\sum_{j=1}^t y_j$, $\mathbf{S}_{m,t}=\sum_{j=1}^t\boldsymbol{\tau}_{m,j}$, $\mathbf{S}_{k,t}=\sum_{j=1}^t\mathbf{x}_{k,j}$, and $Z_t=U_t-\boldsymbol{\gamma}'_k\mathbf{x}_{k,t}$, with $U_t=\sum_{j=1}^t u_j=O_p(n^{1-\kappa})$ for all t = 1, …, n. In more compact form, the IM cointegrating regression can be written as $S_t=\boldsymbol{\theta}'\mathbf{g}_t+Z_t$ where now $\boldsymbol{\theta}=(\boldsymbol{\alpha}'_m,\boldsymbol{\beta}'_k,\boldsymbol{\gamma}'_k)'$, and $\mathbf{g}_t=(\mathbf{S}'_{m,t},\mathbf{S}'_{k,t},\mathbf{x}'_{k,t})'$, and the IM-OLS estimator is obtained by applying the OLS estimation method, i.e. $\tilde{\boldsymbol{\theta}}_n=(\sum_{t=1}^n\mathbf{g}_t\mathbf{g}'_t)^{-1}\sum_{t=1}^n\mathbf{g}_tS_t=\boldsymbol{\theta}+(\sum_{t=1}^n\mathbf{g}_t\mathbf{g}'_t)^{-1}\sum_{t=1}^n\mathbf{g}_tZ_t$. Choi and Ahn (1995) also propose the OLS estimation of a simplified version of (2.17) in the case where the deterministic component characterizing the observations of the deterministically trending integrated regressors $\mathbf{x}_{k,t}$ is such as $\mathbf{d}_{k,t}=\mathbf{A}_{k,m}\boldsymbol{\tau}_{m,t}$ so that $\mathbf{A}_{k,q}=\mathbf{0}_{k,q}$, but without the addition of $\mathbf{x}_{k,t}$ as regressors and with the partial sums of y_t and $\mathbf{x}_{k,t}$ replaced by the integrated versions of the CCR transformation of such variables. Following the same line of analysis as from the estimation of (2.4), we can express the vector of regressors $\mathbf{g}_t$ as

$$\begin{aligned}\mathbf{g}_t&=\begin{pmatrix}n\mathbf{\Gamma}^{-1}_{mm,n}\mathbf{S}_{m,nt}\\ n\mathbf{A}_{k,m}\mathbf{\Gamma}^{-1}_{mm,n}\mathbf{S}_{m,nt}+n\mathbf{\Gamma}^{-1}_{kk,n}\mathbf{M}_{k,nt}\\ \mathbf{A}_{k,m}\mathbf{\Gamma}^{-1}_{mm,n}\boldsymbol{\tau}_{m,nt}+\mathbf{\Gamma}^{-1}_{kk,n}\mathbf{m}_{k,nt}\end{pmatrix}\\&=\begin{pmatrix}n\mathbf{\Gamma}^{-1}_{mm,n}&\mathbf{0}_{m,k}&\mathbf{0}_{m,k}\\ n\mathbf{A}_{k,m}\mathbf{\Gamma}^{-1}_{mm,n}&n\mathbf{\Gamma}^{-1}_{kk,n}&\mathbf{0}_{k,k}\\ \mathbf{0}_{k,m}&\mathbf{0}_{k,k}&\mathbf{\Gamma}^{-1}_{kk,n}\end{pmatrix}\begin{pmatrix}\mathbf{S}_{m,nt}\\ \mathbf{M}_{k,nt}\\ \mathbf{m}_{k,nt}\end{pmatrix}+\begin{pmatrix}\mathbf{0}_{m,m}&\mathbf{0}_{m,k}\\ \mathbf{0}_{k,m}&\mathbf{0}_{k,k}\\ \mathbf{A}_{k,m}\mathbf{\Gamma}^{-1}_{mm,n}&\mathbf{0}_{k,k}\end{pmatrix}\begin{pmatrix}\boldsymbol{\tau}_{m,nt}\\ \mathbf{m}_{k,nt}\end{pmatrix}\\&=\mathbf{W}_{1,n}\mathbf{g}_{nt}+\mathbf{W}_{2,n}\mathbf{m}_{nt}\end{aligned} \tag{2.18}$$

where $\mathbf{S}_{m,nt}=n^{-1}\sum_{j=1}^t\boldsymbol{\tau}_{m,nj}$ and $\mathbf{M}_{k,nt}=n^{-1}\sum_{j=1}^t\mathbf{m}_{k,nj}$, so that the components of the IM-OLS estimation error can be decomposed as

$$\begin{aligned}\sum_{t=1}^n\mathbf{g}_t\mathbf{g}'_t=n\mathbf{W}_{1,n}\Big\{&n^{-1}\sum_{t=1}^n\mathbf{g}_{nt}\mathbf{g}'_{nt}+n^{-1}\sum_{t=1}^n\mathbf{g}_{nt}\mathbf{m}'_{nt}\mathbf{W}'_{3,n}+\mathbf{W}_{3,n}n^{-1}\sum_{t=1}^n\mathbf{m}_{nt}\mathbf{g}'_{nt}\\&+\mathbf{W}_{3,n}n^{-1}\sum_{t=1}^n\mathbf{m}_{nt}\mathbf{m}_{nt}\mathbf{W}'_{3,n}\Big\}\mathbf{W}'_{1,n}\end{aligned} \tag{2.19}$$

and

$$\sum_{t=1}^{n}\mathbf{g}_t Z_t = n^{2-\kappa}\mathbf{W}_{1,n}\left\{n^{-1}\sum_{t=1}^{n}\mathbf{g}_{nt}(n^{-(1-\kappa)}Z_t)+\mathbf{W}_{3,n}n^{-1}\sum_{t=1}^{n}\mathbf{m}_{nt}(n^{-(1-\kappa)}Z_t)\right\} \tag{2.20}$$

where the weighting matrix $\mathbf{W}_{3,n}=\mathbf{W}_{1,n}^{-1}\mathbf{W}_{2,n}$ is of the form

$$\mathbf{W}_{3,n}=\begin{pmatrix}\mathbf{0}_{m,m} & \mathbf{0}_{m,k}\\ \mathbf{0}_{k,m} & \mathbf{0}_{k,k}\\ \mathbf{\Gamma}_{kk,n}\mathbf{A}_{k,m}\mathbf{\Gamma}_{mm,n}^{-1} & \mathbf{0}_{k,k}\end{pmatrix}$$

with the $k\times m$ matrix $\mathbf{\Gamma}_{kk,n}\mathbf{A}_{k,m}\mathbf{\Gamma}_{mm,n}^{-1}$ being asymptotically negligible when $m = 1$ with $p_1 = 0$ (that is, when the cointegrating regression (2.4) only contains a constant term) irrespective of whether $\mathbf{A}_{k,q}$ is zero or not, but diverges with the sample size for any other specification of the deterministic component in the regression model thus failing to provide finite limiting results for the terms in brackets in (2.19)-(2.20). Additionally, for the scaled regression error $n^{-(1-\kappa)}Z_t$ in (2.20) we have the following decomposition

$$n^{-(1-\kappa)}Z_t = n^{-(1-\kappa)}U_t - n^{-(1/2-\kappa)}\boldsymbol{\gamma}_k'\boldsymbol{\eta}_{k,nt} - n^{-(1-\kappa)}\boldsymbol{\gamma}_k'\mathbf{A}_k\mathbf{\Gamma}_n^{-1}\boldsymbol{\tau}_{nt} \tag{2.21}$$

with $\mathbf{A}_k=(\mathbf{A}_{k,m},\mathbf{A}_{k,q})$, $\mathbf{\Gamma}_n=\mathrm{diag}(\mathbf{\Gamma}_{mm,n},\mathbf{\Gamma}_{qq,n})$, and $\boldsymbol{\tau}_{nt}=(\boldsymbol{\tau}_{m,nt}',\boldsymbol{\tau}_{k,nt}')'$, so that the last term involving the normalized polynomial trend function underlying the observations of the stochastic regressors $\mathbf{x}_{k,t}$ will vanish asymptotically only in the case $m = 1$, $q = 0$ and $p_1 = 0$, that is when $\mathbf{d}_{k,t}$ only contains k constant terms, $\mathbf{d}_{k,t}=\mathbf{A}_{k,1}=(\alpha_{11},\ldots,\alpha_{k1})'$, both under cointegration ($\kappa = 1/2$) and no cointegration ($\kappa = -1/2$). In any other case ($m > 1$, $p_m > 0$) and under cointegration, this term will diverges with the sample size and dominates the other two components. This analysis implies that the standard formulation of the IM cointegrating regression in (2.17) is not appropriate to deal with a deterministic component $\mathbf{d}_{k,t}$ including more than k constant terms. The extension of this formulation to the general case considered in Assumption 2.1 and some required modifications of the IM-OLS estimation method is not considered in the present paper and it is still under development by the author. The original specification in Vogelsang and Wagner (2014) corresponds to the case $\mathbf{A}_k=\mathbf{0}_{k,m+q}$, so that the components in $\mathbf{g}_{nt}$ are $\mathbf{M}_{k,nt}=\mathbf{H}_{k,nt}=n^{-1}\sum_{j=1}^{t}\boldsymbol{\eta}_{k,nj}$, $\mathbf{m}_{k,nt}=\boldsymbol{\eta}_{k,nt}$, and $Z_t=U_t-\boldsymbol{\gamma}_k'\boldsymbol{\eta}_{k,t}$, with $\mathbf{g}_t=\mathbf{W}_{0,n}\mathbf{g}_{nt}$, where the weighting matrix is $\mathbf{W}_{0,n}=\mathrm{diag}(n\mathbf{\Gamma}_{mm,n}^{-1},n^{3/2}\mathbf{I}_{k,k},n^{1/2}\mathbf{I}_{k,k})$, providing the following representation for the IM-OLS estimation error of the model parameters

$$\tilde{\mathbf{\Theta}}_n = n^{-(1-\kappa)}\mathbf{W}_{0,n}'(\tilde{\boldsymbol{\theta}}_n-\boldsymbol{\theta})=\left(n^{-1}\sum_{t=1}^{n}\mathbf{g}_{nt}\mathbf{g}_{nt}'\right)^{-1}n^{-1}\sum_{t=1}^{n}\mathbf{g}_{nt}(n^{-(1-\kappa)}Z_t) \tag{2.22}$$

where $\tilde{\mathbf{\Theta}}_n=(\tilde{\mathbf{\Theta}}_{m,n}',\tilde{\mathbf{\Theta}}_{\boldsymbol{\beta}_k,n}',\tilde{\mathbf{\Theta}}_{\boldsymbol{\gamma}_k,n}')'$ with components given by $\tilde{\mathbf{\Theta}}_{m,n}=n^{\kappa}\mathbf{\Gamma}_{mm,n}^{-1}(\tilde{\boldsymbol{\alpha}}_{m,n}-\boldsymbol{\alpha}_m)$, $\tilde{\mathbf{\Theta}}_{\boldsymbol{\beta}_k,n}=n^{1/2+\kappa}(\tilde{\boldsymbol{\beta}}_{k,n}-\boldsymbol{\beta}_k)$ and $\tilde{\mathbf{\Theta}}_{\boldsymbol{\gamma}_k,n}=n^{-(1/2-\kappa)}(\tilde{\boldsymbol{\gamma}}_{k,n}-\boldsymbol{\gamma}_k)$, which makes evident that this estimation method provides the same consistency rates for the estimators of $\boldsymbol{\alpha}_m$ and $\boldsymbol{\beta}_k$ under cointegration that the usual estimates. Also, taking into account that $n^{-1/2}Z_{[nr]}\Rightarrow B_u(r)-\boldsymbol{\gamma}_k'\mathbf{B}_k(r)=B_{u.k}(r)$ under cointegration and the definition of the centering parameter for $\tilde{\boldsymbol{\gamma}}_{k,n}$ as $\boldsymbol{\gamma}_k=\boldsymbol{\gamma}_{ku}=\mathbf{\Omega}_{kk}^{-1}\boldsymbol{\omega}_{ku}$ in case of endogeneity, then the limiting distribution of the estimates is given by

$$\tilde{\boldsymbol{\Theta}}_n \Rightarrow \left(\int_0^1 \mathbf{g}(r)\mathbf{g}(r)'dr\right)^{-1} \int_0^1 \mathbf{g}(r)B_{u.k}(r)dr \tag{2.23}$$

wnere $\mathbf{g}(r) = (\mathbf{g}_m'(r), \mathbf{g}_k'(r), \mathbf{B}_k'(r))'$, with $\mathbf{g}_m(r) = \int_0^r \boldsymbol{\tau}_m(s)ds$ and $\mathbf{g}_k(r) = \int_0^r \mathbf{B}_k(s)ds$, so that the limiting result in (2.23) has a compound normal limit distribution with nuisance parameters that can be cancelled through scaling. Also, Theorem 2 in Vogelsang and Wagner (2014) gives an alternative representation for the second term in the right hand side of (2.23) as $\int_0^1 \mathbf{g}(r)B_{u.k}(r)dr = \int_0^1[\mathbf{G}(1) - \mathbf{G}(r)]dB_{u.k}(r)$, where $\mathbf{G}(r) = \int_0^r \mathbf{g}(s)ds$. An important result stated in Proposition 2 by these authors is that the IM-OLS estimates are almost asymptotically efficient in the sense that are asymptotically less efficient than FM-OLS, although its conditional asymptotic covariance matrix ignores the impact of the long-run variance estimators on the sampling behavior of these estimates. An important by-product of the IM-OLS estimation is that the first difference of the residuals, $\tilde{Z}_t = Z_t - n^{1-\kappa}\mathbf{g}_{nt}'\tilde{\boldsymbol{\Theta}}_n$, can be written as

$$\tilde{z}_t = \Delta\tilde{Z}_t = (1, -\tilde{\boldsymbol{\gamma}}_{k,n}')\boldsymbol{\xi}_t - n^{-\kappa}(\boldsymbol{\tau}_{m,nt}', \boldsymbol{\eta}_{k,nt}')\begin{pmatrix}\tilde{\boldsymbol{\Theta}}_{m,n} \\ \tilde{\boldsymbol{\Theta}}_{\boldsymbol{\beta}_k,n}\end{pmatrix} \tag{2.24}$$

where $\tilde{\boldsymbol{\gamma}}_{k,n} = \boldsymbol{\gamma}_k + n^{1/2-\kappa}\tilde{\boldsymbol{\theta}}_{\boldsymbol{\gamma}_k,n}$, with $\tilde{\boldsymbol{\theta}}_{\boldsymbol{\gamma}_k,n}$ denoting the last *k* components of (2.22). This result provides a similar result to (2.16) for the residual autocovariances and the kernel-type estimator of the long-run variance computed from (2.24) under cointegration (κ = 1/2), given by $\tilde{\omega}_{z,n}^2(q_n) = (1, -\tilde{\boldsymbol{\gamma}}_{k,n}')\boldsymbol{\Omega}_{\boldsymbol{\xi},n}(q_n)(1, -\tilde{\boldsymbol{\gamma}}_{k,n}')' + O_p(n^{-1/2})$[6], so that as $n \to \infty$ we get $\tilde{\omega}_{z,n}^2(q_n) \Rightarrow \omega_z^2 = \omega_{u.k}^2(1 + \omega_{u.k}^{-2}\tilde{\boldsymbol{\theta}}_{\boldsymbol{\gamma}_k}'\boldsymbol{\Omega}_{kk}\tilde{\boldsymbol{\theta}}_{\boldsymbol{\gamma}_k}) = \omega_{u.k}^2(1 + \boldsymbol{\theta}_{\boldsymbol{\gamma}_k}'\boldsymbol{\theta}_{\boldsymbol{\gamma}_k})$ with $\tilde{\boldsymbol{\theta}}_{\boldsymbol{\gamma}_k} = \omega_{u.k}\boldsymbol{\Omega}_{kk}^{-1/2}\boldsymbol{\theta}_{\boldsymbol{\gamma}_k}$ the last *k* random components of (2.23), implying that $\tilde{\omega}_{z,n}^2(q_n)$ is inconsistent for $\omega_{u.k}^2$ with $\omega_z^2 > \omega_{u.k}^2$.

Once analyzed the properties of some alternative estimation methods of the cointegrating regression model, it is important to remind the important role played by the structure of the deterministic component underlying the generating mechanism of the observations of the stochastic regressors, characterizing the limiting results both under cointegration and no cointegration in terms of the weak limit of the normalized *k*-dimensional vector $\mathbf{m}_{k,nt} = \boldsymbol{\Gamma}_{kk,n}(\boldsymbol{\eta}_{k,t} + \mathbf{A}_{k,q}\boldsymbol{\tau}_{q,t})$ given in equation (2.9) when $\mathbf{A}_{k,q} \neq \mathbf{0}_{k,q}$ and these trend components are omitted from the estimated regression model. Other important situation that conditions all these results is given by the inclusion of a subset of stationary and/or cointegrated variables as regressors in $\mathbf{x}_{k,t}$. These two cases are quite different but produces similar results in terms of the properties of the estimates of the corresponding parameters, mainly in relation to the rate of consistency and the resulting limiting distribution of the estimates, as can be seen, e.g., in Theorem 5.3 in Park and Phillips (1989) and Theorem 4.1 in Phillips (1995). Either of these cases represents the situation where some of the parameters in $\boldsymbol{\theta} = (\boldsymbol{\alpha}_m', \boldsymbol{\beta}_k')'$ cannot be consistently estimated even under the cointegration assumption. This is the case where there exist a certain number $1 \leq k_2 < k$ of cointegrating relationships among the set of *k* integrated regressors (subcointegration) or, alternatively, when there are k_2 out of *k* of such stochastic regressors that are not I(1) and behave like stationary variables.

---

[6] Specifically we have $\tilde{\omega}_{z,n}^2(q_n) = \hat{\omega}_{u^+,n}^2(q_n) - 2\tilde{\theta}_{\boldsymbol{\gamma}_k,n}'(\boldsymbol{\omega}_{ku,n} - \boldsymbol{\Omega}_{kk,n}\boldsymbol{\gamma}_{ku}) + \tilde{\theta}_{\boldsymbol{\gamma}_k,n}'\boldsymbol{\Omega}_{kk,n}\tilde{\theta}_{\boldsymbol{\gamma}_k,n}$, where the terms involving the elements in $\boldsymbol{\Omega}_{\boldsymbol{\xi},n}(q_n)$ will converge to the corresponding population counterparts under traditional bandwidth assumptions.

Appendix A presents, in a unified framework, these two possible situations that can occur in some practical applications where the main result is that only the set of parameters related to the remaining k_1 integrated and no cointegrated regressors can be consistently estimated. This general result will substantially modify both the above analysis as well as some results in the next section, except in the particular case where the stochastic component characterizing the behavior of the stationary regressors or the error terms characterizing the cointegrating relationship among the two sets of stochastic integrated regressors are contemporaneously uncorrelated with the regression errors u_t, as it is assumed in McCabe et.al. (1997).

2.2. Residual-based tests for cointegration

This part of the section reviews some of the more commonly used testing procedures to test for the existence of a single cointegration relationship in the framework of the cointegrating regression model. This is not an exhaustive study of these procedures, but aims to analyze the behaviour of some of these test statistics, both for testing the null of no cointegration against the alternative of cointegration or for these same hypothesis in reverse order, with particular emphasis in the impact of the number and nature of the trending regressors appearing in the cointegrating regression on the limiting distributions of these test statistics. In what follows we consider the OLS versions of these test statistics based on the sequence of OLS residuals, $\hat{u}_t = y_t - \mathbf{m}_t'\hat{\boldsymbol{\theta}}_n$, that from (2.4), (2.10) and (2.12) can also be represented in either of the two following forms

$$\hat{u}_t = u_t - n^{-\kappa}\mathbf{m}_{nt}'\hat{\boldsymbol{\Theta}}_n = u_{t,m} - n^{-\kappa}\hat{\mathbf{m}}_{k,nt}'\hat{\boldsymbol{\Theta}}_{k,n}, \tag{2.25}$$

which shows that although being consistent estimators of the regression error terms under cointegration, i.e., $\hat{u}_t = u_t + O_p(n^{-1/2})$ when κ = 1/2, some other properties will depend on the model's dimension and the structure 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 these features. A wide variety of semiparametric and parametric statistics, both for testing the null hypothesis of cointegration against no cointegration (as, e.g., the ones proposed by Shin (1994)[7], $\hat{CI}_{n,m}(k)$, Xiao (1999) and Wu and Xiao (2008), $\hat{R}_{n,m}(k)$, and Xiao and Phillips (2002), $\hat{CS}_{n,m}(k)$, that will be presented below) and for testing the reserved hypothesis (see, e.g., the residual-based statistics proposed in Phillips (1987a) and Phillips and Ouliaris (1990)), exploit the information content of these residuals and their limiting distributions basically depends on the number and nature of the stochastic and deterministic trend components contained in the estimated cointegrating regression. However, 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 $\mathbf{d}_{k,t}$. Also, Hansen (1990) indicates some other important consequences of this dependence upon dimensionality. Given part

[7] This testing procedure was also independently proposed by Harris and Inder (1994) and Leybourne and McCabe (1994), and consists on adapting the so-called KPSS test for the null of stationarity of a univariate series by Kwiatkowski et.al. (1992) to the regression errors of a cointegrating regression model.

B of Assumption 2.2, the AR(1) structure of the regression error u_t is transferred to the OLS residuals such as

$$\hat{u}_t = \alpha \hat{u}_{t-1} + \varepsilon_t \tag{2.26}$$

where the error term ε_t is can be written as $\varepsilon_t = \upsilon_t - n^{-\kappa}(\mathbf{m}'_{nt} - \alpha \mathbf{m}'_{n(t-1)})\hat{\mathbf{\Theta}}_n$ with $\mathbf{m}_{nt} - \alpha \mathbf{m}_{n(t-1)} = (1-\alpha)\mathbf{m}_{nt} + \alpha \Delta \mathbf{m}_{nt}$, so that the OLS estimator of α, given by

$$\hat{\alpha}_n = \frac{\sum_{t=1}^{n} \hat{u}_{t-1}\hat{u}_t}{\sum_{t=1}^{n} \hat{u}_{t-1}^2} = \alpha + \frac{\sum_{t=1}^{n} \hat{u}_{t-1}\varepsilon_t}{-\hat{u}_n^2 + \sum_{t=1}^{n} \hat{u}_t^2} = \alpha + n^{-(1/2-\kappa)} \frac{n^{-1}\sum_{t=1}^{n} \hat{u}_{t-1}\varepsilon_t}{n^{-(3/2-\kappa)}\left(-\hat{u}_n^2 + \sum_{t=1}^{n} \hat{u}_t^2\right)} \tag{2.27}$$

with $\hat{u}_0 = 0$, does not converge to a constant, but stays random in the limit in the case of no cointegration (i.e., when α = 1 with κ = −1/2). In this case, taking into account that $n^{-1/2}\hat{u}_{[nr]} = n^{-1/2}u_{[nr]} - \mathbf{m}'_{n[nr]}\hat{\mathbf{\Theta}}_n = O_p(1)$ weakly converges to the random limit $n^{-1/2}\hat{u}_{[nr]} \Rightarrow B_u(r) - \mathbf{m}(r)\mathbf{\Theta}_1$, with $\mathbf{\Theta}_1$ the limiting distribution under no cointegration of the OLS estimates of $\boldsymbol{\theta}$, $\mathbf{\Theta}_1 = (\int_0^1 \mathbf{m}(s)\mathbf{m}'(s)ds)^{-1}\int_0^1 \mathbf{m}(s)B_u(s)ds$ (see also equation (2.14)), the limiting distribution of many statistics constructed from $\hat{u}_t$ to test the null hypothesis of a unit root (no cointegration) can be show to depend upon this random element, and hence on *m*, *q* and *k*, that is on the number and type of trend components in the system. A detailed inspection of the sample covariance in the numerator of the last equality of (2.27) gives the representation

$$\begin{aligned} n^{-1}\sum_{t=1}^{n}\hat{u}_{t-1}\varepsilon_t &= n^{-1}\sum_{t=1}^{n}u_{t-1}\upsilon_t - n^{-(1/2+\kappa)}\hat{\mathbf{\Theta}}'_n n^{-1/2}\sum_{t=1}^{n}\mathbf{m}_{n,t-1}\upsilon_t \\ &\quad - n^{-2\kappa}\hat{\mathbf{\Theta}}'_n\left\{n^{-(1-\kappa)}\sum_{t=1}^{n}\mathbf{m}_{n,t}u_{t-1} - n^{-1}\sum_{t=1}^{n}\mathbf{m}_{n,t}\mathbf{m}'_{n,t-1}\right\}\hat{\mathbf{\Theta}}_n \\ &\quad + \alpha n^{-(1-\kappa)}\mathbf{m}_{n,t}\hat{u}_n \end{aligned}$$

so that under cointegration (κ = 1/2) we have $n^{-1}\sum_{t=1}^{n}\hat{u}_{t-1}\varepsilon_t = n^{-1}\sum_{t=1}^{n}u_{t-1}\upsilon_t + O_p(n^{-1})$, which implies that $\hat{\alpha}_n - \alpha = O_p(1)$ and $\sqrt{n}(\hat{\alpha}_n - \alpha - \hat{\sigma}_{u,n}^{-2}E[u_{t-1}\upsilon_t]) = O_p(1)$, where $\hat{\sigma}_{u,n}^2 = n^{-1}\sum_{t=1}^{n}\hat{u}_t^2 \to^p \sigma_u^2$ and $E[u_{t-1}\upsilon_t] = \sum_{j=0}^{\infty}\mathbf{c}'_{0,j}\mathbf{\Sigma}_e\mathbf{d}_{0,j+1}$. This implies that the asymptotic behavior of $\hat{\alpha}_n$ is model-free and only depends on the dynamics of the regression error term. The same applies to the well-known Z tests proposed by Phillips and Ouliaris (1990) (PO), given by the normalized estimation error $\hat{Z}_{1,n} = n(\hat{\alpha}_n(\hat{\lambda}_{\varepsilon,n}) - 1)$ and the pseudo-T ratio test statistic $\hat{Z}_{2,n} = (\hat{\omega}_{\varepsilon,n}^2/\sum_{t=1}^{n}\hat{u}_{t-1}^2)^{-1/2}(\hat{\alpha}_n(\hat{\lambda}_{\varepsilon,n}) - 1)$, where $\hat{\omega}_{\varepsilon,n}^2 = \hat{\sigma}_{\varepsilon,n}^2 + 2\hat{\lambda}_{\varepsilon,n}(q_n)$, with $\hat{\sigma}_{\varepsilon,n}^2 = n^{-1}\sum_{t=1}^{n}\hat{\varepsilon}_t^2$ and $\hat{\lambda}_{\varepsilon,n}(q_n) = \sum_{h=1}^{n-1}w(h/q_n)n^{-1}\sum_{t=h+1}^{n}\hat{\varepsilon}_{t-h}\hat{\varepsilon}_t$ ($\hat{\varepsilon}_1 = 0$) are computed from the sequence of OLS residuals in (2.26), i.e. $\hat{\varepsilon}_t = \hat{u}_t - \hat{\alpha}_n\hat{u}_{t-1} = \varepsilon_t - \hat{u}_{t-1}(\hat{\alpha}_n - \alpha) = \varepsilon_t - n^{-(1/2-\kappa)}\hat{u}_{t-1}[n^{1/2-\kappa}(\hat{\alpha}_n - \alpha)]$, and $\hat{\alpha}_n(\hat{\lambda}_{\varepsilon,n})$ is the bias-corrected estimator of α, $\hat{\alpha}_n(\hat{\lambda}_{\varepsilon,n}) = \sum_{t=1}^{n}(\hat{u}_{t-1}\hat{u}_t - \hat{\lambda}_{\varepsilon,n}(q_n))/\sum_{t=1}^{n}\hat{u}_{t-1}^2$. These limiting distributions under non-stationarity shift away from the origin as the dimensionality of the model increases. Thus larger values for these test statistics are needed for rejection, implying that smaller estimated AR(1) parameters are needed with an expected reduction in the power, particularly in small and even moderate sample sizes with moderately large systems.

When focusing on testing the null hypothesis of cointegration, we found a quite similar effect as the one described by Hansen (1990) for the semiparametric statistics based on measures of no excessive fluctuation in the series of residuals compatible with the null hypothesis of stationarity of the regression error terms cited above, which are given by the following functionals

$$C\hat{I}_{n,m}(k)=\frac{1}{n^2\hat{\omega}^2_{u,n}(q_n)}\sum_{t=1}^{n}\hat{U}_t^2 \tag{2.28}$$

$$\hat{R}_{n,m}(k)=\frac{1}{\hat{\omega}_{u,n}(q_n)\sqrt{n}}\max_{t=1,\dots,n}|\hat{U}_t-(t/n)\hat{U}_n| \tag{2.29}$$

and

$$C\hat{S}_{n,m}(k)=\frac{1}{\hat{\omega}_{u,n}(q_n)\sqrt{n}}\max_{t=1,\dots,n}|\hat{U}_t| \tag{2.30}$$

with $\hat{\omega}^2_{u,n}(q_n)=\sum_{h=-(n-1)}^{n-1}w(hq_n^{-1})n^{-1}\sum_{t=|h|+1}^{n}\hat{u}_t\hat{u}_{t-|h|}$, and $\hat{R}_{n,m}(k)=C\hat{S}_{n,m}(k)$ for $m\geq 1$ when the deterministic term appearing in the estimated cointegrating regression contains at least a constant term, so that $\hat{U}_n=0$. The main component of all these statistics used to find empirical evidence compatible with a stationary behavior of the regression error term is the CUSUM-type fluctuation measure given by

$$n^{-1/2}\hat{U}_{[nr]}=n^{-1/2}\sum_{t=1}^{[nr]}\hat{u}_t=n^{1/2-\kappa}\left\{n^{-(1-\kappa)}\sum_{t=1}^{[nr]}u_t-n^{-1}\sum_{t=1}^{[nr]}\mathbf{m}'_{nt}\hat{\mathbf{\Theta}}_n\right\}$$

which, under cointegration, has the following limiting distribution representation

$$n^{-1/2}\hat{U}_{[nr]}\Rightarrow B_u(r)-\int_0^r\mathbf{m}'(s)ds\left(\int_0^1\mathbf{m}(s)\mathbf{m}'(s)ds\right)^{-1}\left(\int_0^1\mathbf{m}(s)dB_u(s)+\mathbf{\Phi}\right) \tag{2.31}$$

that, besides the dependence on the bias term $\mathbf{\Phi}=(\mathbf{0}'_m,\mathbf{\Phi}'_{ku})$, is a function of the deterministic and stochastic trend components in the system through the random vector $\mathbf{m}(r)=(\boldsymbol{\tau}'_m(r),\mathbf{m}'_k(r))'$. As a numerical illustration of the dependence on model's dimensionality even in the simplest case, Table 1 below presents the finite-sample quantiles of the null distribution under cointegration of the OLS versions of the test statistics $C\hat{I}_{n,m}(k)$ and $\hat{R}_{n,m}(k)$, for a sample size of n = 250 observations and 10000 independent replications under the assumption of strictly exogenous stochastic regressors.

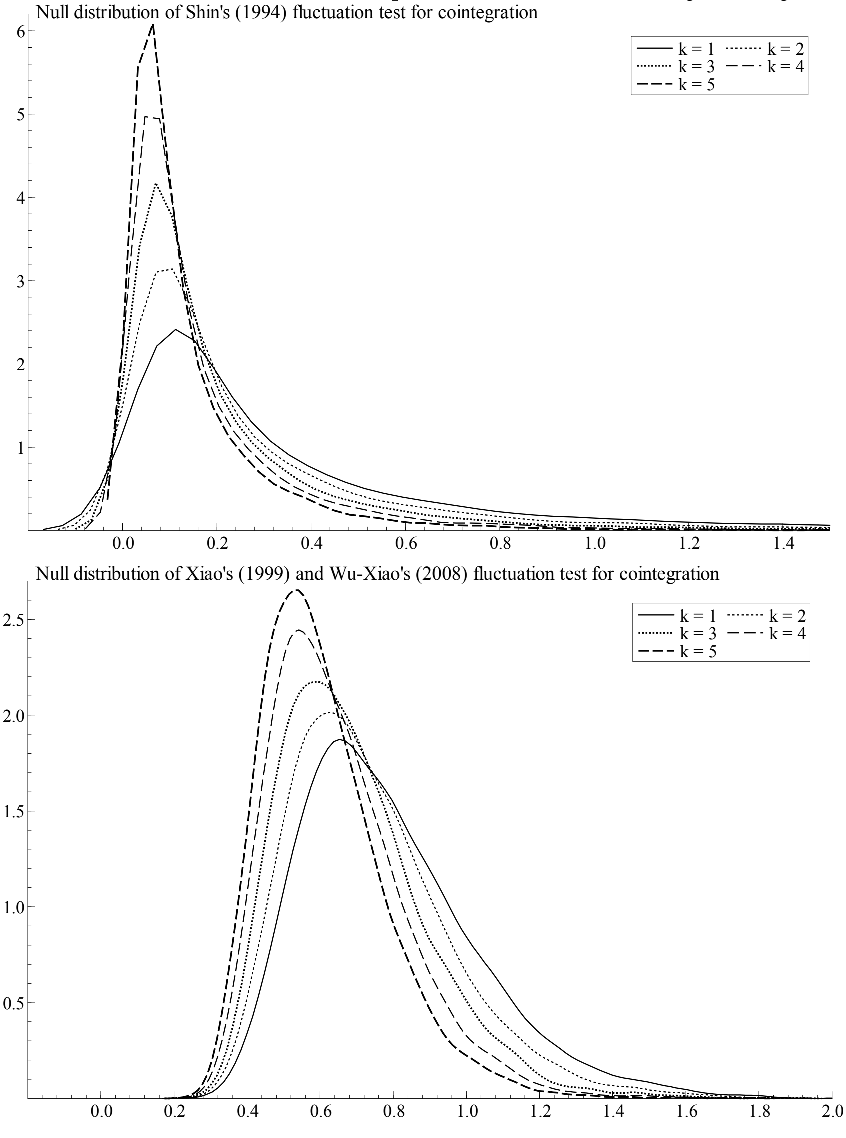
Table 1. Finite-sample quantiles of the null distribution under cointegration for the fluctuation-type test statistics. Case of no deterministic component and k = 1, ..., 5 integrated regressors

Significance	Shin (1994) Test					Xiao (1999), Wu and Xiao (2008) Test				
Level	k = 1	2	3	4	5	k = 1	2	3	4	5
0.01	0.0289	0.0254	0.0218	0.0192	0.0175	0.4008	0.3810	0.3633	0.3466	0.3357
0.025	0.0356	0.0305	0.0266	0.0233	0.0207	0.4365	0.4155	0.3938	0.3742	0.3587
0.05	0.0442	0.0368	0.0317	0.0277	0.0249	0.4731	0.4448	0.4250	0.4036	0.3851
0.1	0.0586	0.0470	0.0398	0.0340	0.0307	0.5191	0.4892	0.4635	0.4398	0.4206
0.25	0.0978	0.0758	0.0634	0.0528	0.0461	0.6126	0.5735	0.5410	0.5110	0.4856
0.5	0.1993	0.1510	0.1197	0.0982	0.0832	0.7496	0.6995	0.6558	0.6140	0.5798
0.75	0.4410	0.3299	0.2514	0.2072	0.1707	0.9251	0.8577	0.7977	0.7488	0.7007
0.9	0.8780	0.6469	0.4804	0.3863	0.3184	1.1141	1.0302	0.9561	0.8874	0.8321
0.95	1.2598	0.9326	0.6863	0.5491	0.4407	1.2470	1.1504	1.0629	0.9881	0.9219
0.975	1.6199	1.2346	0.9254	0.7438	0.5989	1.3725	1.2622	1.1487	1.0784	1.0196
0.99	2.2143	1.6575	1.2705	1.0106	0.8169	1.5261	1.4331	1.3007	1.1949	1.1278

A quick inspection of these results reveals a quite similar effect as the one described by

Hansen (1990) for testing procedures of the null of no cointegration, which is the requirement of fluctuations of lower magnitude for models of largest dimension for not rejection of the null hypothesis of cointegration, resulting in an expected loss of power for high dimensional systems as a consequence of the very different shape of these distributions depending on *k*, the number of integrated regressors. Also, Figure 1 below displays these distributions through kernel-density estimation of the OLS-based fluctuation-type test statistics with strictly exogenous integrated regressors, where it can be appreciate that irrespective of the value of *k*, these are right skewed distributions and are more concentrated for increasing values of *k*.

Figure 1. Kernel-density estimates of the null distribution under cointegration of fluctuation-type tests statistics. Case of no deterministic component and $k = 1, \ldots, 5$ integrated regressors



To complete these results in the alternative situation of no cointegration, that is when the regression errors u_t is also an I(1) process with $\alpha = 1$ in Assumption 2.2(B) and $\kappa = -1/2$, the normalized OLS estimator of $\boldsymbol{\beta}_k$ can also be written as

$$\hat{\boldsymbol{\Theta}}_{k,n} = n^{-1/2}\boldsymbol{\Gamma}_{kk,n}^{-1}\hat{\boldsymbol{\beta}}_{k,n} = \left((1/n)\sum_{t=1}^{n}\hat{\mathbf{m}}_{k,nt}\hat{\mathbf{m}}'_{k,nt}\right)^{-1} n^{-3/2}\sum_{t=1}^{n}\hat{\mathbf{m}}_{k,nt}\hat{\eta}_{0,t}$$

where $\hat{\eta}_{0,t} = \eta_{0,t} - n^{1/2}(n^{-3/2}\sum_{j=1}^{n}\eta_{0,j}\boldsymbol{\tau}'_{m,nj})\mathbf{Q}_{mm,n}^{-1}\boldsymbol{\tau}_{m,nt}$ are the OLS detrended observations of the dependent variable $y_t = \boldsymbol{\alpha}'_{0,m}\boldsymbol{\tau}_{m,t} + \eta_{0,t}$, so that the OLS residuals can be alternatively represented as

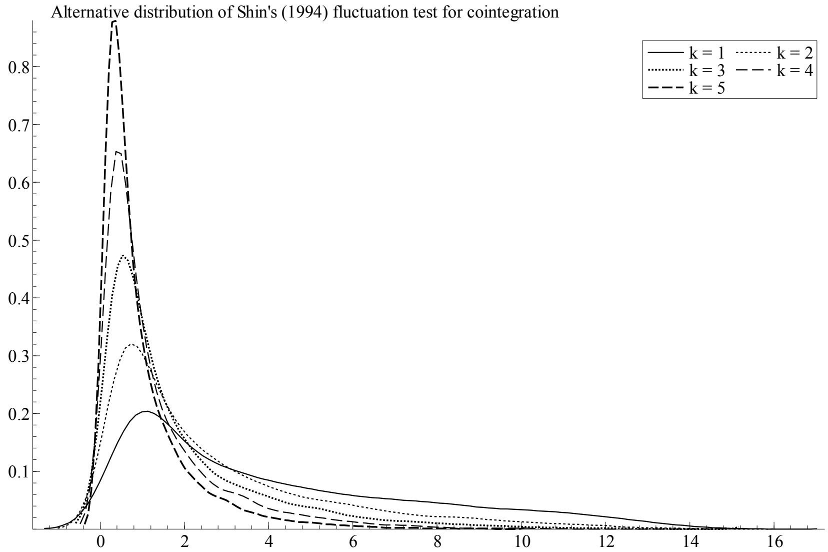
$$\hat{u}_t = \hat{\eta}_{0,t} - \hat{\boldsymbol{\beta}}'_{k,n}\hat{\mathbf{m}}_{k,t} = \sqrt{n}(1, -\hat{\boldsymbol{\Theta}}'_{k,n})\begin{pmatrix}\hat{\eta}_{0,nt} \\ \hat{\mathbf{m}}_{k,nt}\end{pmatrix} \tag{2.32}$$

and their scaled partial sum admits the representation

$$n^{-1/2}\hat{U}_{[nr]} = n^{-1/2}\sum_{t=1}^{[nr]}\hat{u}_t = n\left\{(1, -\hat{\boldsymbol{\Theta}}'_{k,n})n^{-1}\sum_{t=1}^{[nr]}\begin{pmatrix}\hat{\eta}_{0,nt} \\ \hat{\mathbf{m}}_{k,nt}\end{pmatrix}\right\} = O_p(n)$$

which is a fundamental partial result for determining the consistency of these testing procedures under no cointegration, but with the limitations described before. For a more detailed analysis of all these results see, e.g., the work by Phillips (1989). Figure 2 below displays the kernel-density estimation of the distribution of the OLS-based Shin's (1994) test under the alternative of no cointegration, with similar shapes as under cointegration for different number of integrated regressors.

Figure 2. Kernel-density estimates of the alternative distribution under no cointegration of fluctuation-type tests statistics. Case of no deterministic component and $k = 1, \ldots, 5$ integrated regressors



This represents an important difference with respect to, e.g. the PO tests, where the limiting distribution under the alternative hypothesis does not depend on the characteristics of the estimated model. Some other existing cointegration tests, such as the Choi and Ahn (1995) LM-type statistics and the statistic proposed by Jansson (2005), also display the same characteristics although we do not consider their study in this paper.

Also, besides these theoretical considerations, in practical applications it is worth to mention the effects on the size and power properties of the methods used to adjust for serial correlation and endogeneity of the stochastic regressors.

Finally, much of these results are also of application in more complex models allowing to capture some non-linear effects characterizing the potential cointegrating relationship, such as, e.g., structural breaks affecting the parameters of (2.4). Thus, augmenting the specification of the basic cointegrating regression (2.4) as

$$y_t = \boldsymbol{\theta}'_0\mathbf{m}_t + \boldsymbol{\theta}'_1\mathbf{m}_t h_t(\tau_0) + u_t = \boldsymbol{\theta}'\mathbf{A}_t(\tau_0) + u_t$$

where $\boldsymbol{\theta}_0 = (\boldsymbol{\alpha}'_m, \boldsymbol{\beta}'_k)'$, $\boldsymbol{\theta}_1 = (\boldsymbol{\lambda}'_m, \boldsymbol{\pi}'_k)'$, $\boldsymbol{\theta} = (\boldsymbol{\theta}'_0, \boldsymbol{\theta}'_1)'$, and

$$\mathbf{A}_t(\tau_0) = \begin{pmatrix} \mathbf{m}_t \\ \mathbf{m}_t h_t(\tau_0) \end{pmatrix}$$

with $h_t(\tau_0) = I(t > [n\tau_0])$ the indicator function of the break point in the sample, where $\tau_0 \in (0,1)$ is the true break fraction, this model allows for systematic, abrupt and permanent changes in the values of the model parameters. This general formulation allows for a great variety of different specifications with changes affecting the trend function and/or the cointegrating vector. In the case of an unknown break point, the estimating model is given by $y_t = \boldsymbol{\theta}'\mathbf{A}_t(\tau) + e_t$, with regression errors defined as $e_t = u_t + \boldsymbol{\theta}_1'\mathbf{m}_t(h_t(\tau_0) - h_t(\tau))$ so that the scaled OLS estimation error is given by

$$n^{\kappa}\mathbf{D}_n(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) = \left(n^{-1}\sum_{t=1}^{n} \mathbf{A}_{nt}(\tau)\mathbf{A}_{nt}'(\tau) \right)^{-1}$$
$$\times \left\{ n^{-(1-\kappa)}\sum_{t=1}^{n} \mathbf{A}_{nt}(\tau)u_t + n^{-1}\sum_{t=1}^{n} \mathbf{A}_{nt}(\tau)\mathbf{m}_{nt}'(h_t(\tau_0) - h_t(\tau))\, n^{\kappa}\mathbf{W}_n'\boldsymbol{\theta}_1 \right\}$$

with scaling matrix $\mathbf{D}_n = \mathrm{diag}(\mathbf{W}_n, \mathbf{W}_n)$, and

$$n^{\kappa}\mathbf{W}_n'\boldsymbol{\theta}_1 = n^{\kappa} \begin{pmatrix} \boldsymbol{\Gamma}_{m,n}^{-1}(\boldsymbol{\lambda}_m + \mathbf{A}_{k,m}'\boldsymbol{\pi}_k) \\ \boldsymbol{\Gamma}_{kk,n}^{-1}\boldsymbol{\pi}_k \end{pmatrix}$$

Both in the case of a known break point, i.e. $\tau = \tau_0$, or with a wrong determination of the break fraction with relatively small changes in the magnitude of the shifts in the model parameters such as $n^{\kappa}\mathbf{W}_n'\boldsymbol{\theta}_1$ is asymptotically negligible, then all the model parameters are consistently estimated by OLS under cointegration, but with the same nuisance parameters as in the linear case affecting their limiting null distribution.

With all these results 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 many of 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 $\mathbf{z}_t = (u_t, v_{nt})'$, where $v_{nt} = u_t^2 - \sigma_{u,n}^2$, with $\sigma_{u,n}^2 = n^{-1}\sum_{t=1}^{n} u_t^2$. 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

$$n^{-1/2}\sum_{t=1}^{[nr]} \begin{pmatrix} u_t \\ v_{nt} \end{pmatrix} = \begin{pmatrix} n^{-1/2}\sum_{t=1}^{[nr]} u_t \\ n^{-1/2}\sum_{t=1}^{[nr]} (u_t^2 - \sigma_{u,n}^2) \end{pmatrix} \tag{3.1}$$

Under stationarity (i.e., under cointegration if u_t were the regression errors in (2.4)), the scaled partial sum of centered and squared errors admits the decomposition

$$n^{-1/2}\sum_{t=1}^{[nr]} v_{nt} = n^{-1/2}\sum_{t=1}^{[nr]} v_t - \frac{[nr]}{n} n^{-1/2}\sum_{t=1}^{n} v_t \tag{3.2}$$

with $v_t = u_t^2 - \sigma_u^2$, that under quite general assumptions weakly converges to a well-defined limiting distribution. On the other hand, under non-stationarity (i.e., under no cointegration) we have that

$$n^{-5/2}\sum_{t=1}^{[nr]}\begin{pmatrix} u_t \\ v_{nt} \end{pmatrix} = \begin{pmatrix} n^{-1}\left\{ n^{-1}\sum_{t=1}^{[nr]}(n^{-1/2}u_t) \right\} \\ n^{-1}\sum_{t=1}^{[nr]}[(n^{-1/2}u_t)^2 - n^{-1}\sigma_{u,n}^2] \end{pmatrix} = \begin{pmatrix} O_p(n^{-1}) \\ n^{-1}\sum_{t=1}^{[nr]}u_{n,t}^2 - \frac{[nr]}{n} n^{-1}\sum_{t=1}^{n}u_{n,t}^2 \end{pmatrix} \tag{3.3}$$

where $u_{n,t} = n^{-1/2}u_t$, so that the behavior under non-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 (2.25), $\hat{u}_t = u_t - n^{-\kappa}\mathbf{m}'_{nt}\hat{\boldsymbol{\Theta}}_n$, we have that the scaled partial sum of squared and centered residuals is given by

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

with $\hat{\sigma}_{u,n}^2 = n^{-1}\sum_{t=1}^{n}\hat{u}_t^2$ the OLS-based residual variance. Note that this functional, $n^{-1/2}\sum_{t=1}^{[nr]}\hat{v}_t = n^{-1/2}\sum_{t=1}^{[nr]}(\hat{u}_t^2 - \hat{\sigma}_{u,n}^2)$, is the base to built the cumulative sum (CUSUM) of squares statistics used to test for structural stability in linear regression models with stationary regressors and errors (see Deng and Perron (2008) for a recent review of the conditions required on these components to obtain consistent results). In a non-stationary framework, Lee et.al. (2003) study its properties to test for a variance change in a unstable AR(q) model while Nielsen and Sohkanen (2011) also generalize it use to the case of a non-stationary autoregressive distributed lag model with deterministic time trends. For the cointegrated regression model proposed by Maekawa et.al. (1996), with cointegrated regressors and inconsistent OLS estimation of the model parameters, Lu et.al. (2008) studied how to built a CUSUM of squares statistic for testing structural stability based on the first difference of OLS residuals. When using (3.4) for the purpose of testing for cointegration, taking κ = 1/2, the behaviour of (3.4) is asymptotically equivalent to that of (3.2), that is

$$n^{-1/2}\sum_{t=1}^{[nr]}\hat{v}_t = n^{-1/2}\sum_{t=1}^{[nr]}v_{nt} + O_p(n^{-1/2}),$$

so that the limiting distribution of (3.4) under cointegration will be invariant to the structure and nature of the regressors in (2.4). In order to correctly characterize this limiting null distribution, we have to consider an augmented version of the error vector $\boldsymbol{\xi}_t$ given in (2.3) as $\boldsymbol{\zeta}_t = (u_t, v_t, \boldsymbol{\varepsilon}'_{k,t})'$, so that under the assumptions stated on the generating mechanism of $\boldsymbol{\xi}_t$ it also satisfies an invariance principle such as

$$n^{-1/2}\sum_{t=1}^{[nr]}\boldsymbol{\zeta}_t = n^{-1/2}\sum_{t=1}^{[nr]}\begin{pmatrix} u_t \\ u_t^2 - \sigma_u^2 \\ \boldsymbol{\varepsilon}_{k,t} \end{pmatrix} \Rightarrow \mathbf{B}_{\zeta}(r) = \begin{pmatrix} B_u(r) \\ B_v(r) \\ \mathbf{B}_k(r) \end{pmatrix} = \boldsymbol{\Omega}_{\zeta}^{1/2}\mathbf{W}_{\zeta}(r) \tag{3.5}$$

with covariance matrix

$$\mathbf{\Omega}_{\boldsymbol{\zeta}} = \begin{pmatrix} \omega_u^2 & \omega_{uv} & \boldsymbol{\omega}_{uk} \\ & \omega_v^2 & \boldsymbol{\omega}_{vk} \\ & & \mathbf{\Omega}_{kk} \end{pmatrix}$$

where $\boldsymbol{\omega}_{ku} = \boldsymbol{\omega}'_{uk}$ and $\boldsymbol{\omega}_{kv} = \boldsymbol{\omega}'_{vk} = \sum_{h=0}^{\infty} E[\boldsymbol{\varepsilon}_{k,t-h} u_t^2] + \sum_{h=1}^{\infty} E[\boldsymbol{\varepsilon}_{k,t} u_{t-h}^2]$ is the long-run covariance between $\boldsymbol{\varepsilon}_{k,t}$ and $u_t^2 - \sigma_u^2$. By the upper triangular Cholesky decomposition of $\mathbf{\Omega}_{\boldsymbol{\zeta}}$, we have that

$$\mathbf{\Omega}_{\boldsymbol{\zeta}}^{1/2} = \begin{pmatrix} \omega_{u.k}\sqrt{1-\left(\dfrac{\omega_{uv} - \boldsymbol{\omega}_{uk}\mathbf{\Omega}_{kk}^{-1}\boldsymbol{\omega}'_{vk}}{\omega_{u.k}\omega_{v.k}}\right)^2} & \omega_{u.k}\left(\dfrac{\omega_{uv} - \boldsymbol{\omega}_{uk}\mathbf{\Omega}_{kk}^{-1}\boldsymbol{\omega}'_{vk}}{\omega_{u.k}\omega_{v.k}}\right) & \omega_{u.k}\boldsymbol{\omega}_{uk}(\omega_{u.k}^2\mathbf{\Omega}_{kk})^{-1/2} \\ 0 & \omega_{v.k} & \omega_{v.k}\boldsymbol{\omega}_{vk}(\omega_{v.k}^2\mathbf{\Omega}_{kk})^{-1/2} \\ \mathbf{0}_k & \mathbf{0}_k & \mathbf{\Omega}_{kk}^{1/2} \end{pmatrix}$$

where $\omega_{u.k}^2 = \omega_u^2(1-\rho_{uk}^2)$ and $\omega_{v.k}^2 = \omega_v^2(1-\rho_{vk}^2) = \omega_v^2 - \boldsymbol{\omega}'_{kv}\boldsymbol{\gamma}_{kv}$, with $\boldsymbol{\gamma}_{kv} = \mathbf{\Omega}_{kk}^{-1}\boldsymbol{\omega}_{kv}$, are the long-run variances of u_t and $v_t = u_t^2 - \sigma_u^2$ conditional on $\boldsymbol{\varepsilon}_{k,t}$, respectively, and hence $B_v(r)$ in (3.5) can be decomposed as $B_v(r) = \omega_{v.k}W_v(r) + \boldsymbol{\gamma}'_{kv}\mathbf{B}_k(r)$.

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

$$n^{-1/2}\sum_{t=1}^{[nr]}(u_t u_{t+h} - \gamma_u(h)) \Rightarrow B_h(r) = \omega_h W_h(r)$$

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 $i < 0$, and $\omega_h^2 = g_h^2(1)E[(e_0^2 - \sigma_e^2)^2] + \sigma_e^4 \sum_{s=1}^{\infty}(g_{h+s}(1) + g_{h-s}(1))^2$, with

$$\omega_0^2 = g_0^2(1)E[(e_0^2 - \sigma_e^2)^2] + 4\sigma_e^4 \sum_{s=1}^{\infty} g_s^2(1)$$

for $h = 0$ with $g_{-s}(1) = g_s(1)$.

In our case, from Assumption 2.2 we have $u_t = \mathbf{c}'_0(L)\mathbf{e}_t = \sum_{j=0}^{\infty}\mathbf{c}'_{0,j}\mathbf{e}_{t-j}$, with $\sigma_u^2 = \sum_{j=0}^{\infty}\mathbf{c}'_{0,j}\mathbf{\Sigma}_e\mathbf{c}_{0,j} = \sum_{j=0}^{\infty}(\mathbf{c}'_{0,j}\otimes\mathbf{c}'_{0,j})\mathrm{vec}(\mathbf{\Sigma}_e)$, so that $v_t = u_t^2 - \sigma_u^2$ can be decomposed as

$$\begin{aligned} v_t = u_t^2 - \sigma_u^2 &= \sum_{j=0}^{\infty}\mathbf{c}'_{0,j}(\mathbf{e}_{t-j}\mathbf{e}'_{t-j} - \mathbf{\Sigma}_e)\mathbf{c}_{0,j} + 2\sum_{i=1}^{\infty}\sum_{j=0}^{\infty}\mathbf{c}'_{0,j}(\mathbf{e}_{t-j}\mathbf{e}'_{t-j-i})\mathbf{c}_{0,j+i} \\ &= \sum_{j=0}^{\infty}(\mathbf{c}'_{0,j}\otimes\mathbf{c}'_{0,j})\mathrm{vec}(\mathbf{e}_{t-j}\mathbf{e}'_{t-j} - \mathbf{\Sigma}_e) + 2\sum_{i=1}^{\infty}\sum_{j=0}^{\infty}(\mathbf{c}'_{0,j+i}\otimes\mathbf{c}'_{0,j})\mathrm{vec}(\mathbf{e}_{t-j}\mathbf{e}'_{t-j-i}) \\ &= \mathbf{g}'_0(L)\mathrm{vec}(\mathbf{e}_t\mathbf{e}'_t - \mathbf{\Sigma}_e) + 2\sum_{i=1}^{\infty}\mathbf{g}'_i(L)\mathrm{vec}(\mathbf{e}_t\mathbf{e}'_{t-i}) \end{aligned}$$

where $\mathbf{g}_i(L) = \sum_{j=0}^{\infty}(\mathbf{c}_{0,j+i}\otimes\mathbf{c}_{0,j})L^i = \sum_{j=0}^{\infty}\mathbf{g}_{i,j}L^i$, with $\mathbf{g}_{i,j} = (\mathbf{c}_{0,j+i}\otimes\mathbf{c}_{0,j})$ for i, j = 0, 1, …, so that the scaled partial sum process of v_t can be expressed as

$$n^{-1/2}\sum_{t=1}^{[nr]}v_t = \sum_{j=0}^{\infty}\mathbf{g}'_{0,j}n^{-1/2}\sum_{t=1}^{[nr]}\mathrm{vec}(\mathbf{e}_{t-j}\mathbf{e}'_{t-j} - \mathbf{\Sigma}_e) + 2\sum_{s=1}^{\infty}\sum_{j=0}^{\infty}\mathbf{g}'_{s,j}n^{-1/2}\sum_{t=1}^{[nr]}\mathrm{vec}(\mathbf{e}_{t-j}\mathbf{e}'_{t-j-s}) \tag{3.6}$$

By application of a multivariate version of the invariance principle for the sample variances $n^{-1/2}\sum_{t=1}^{[nr]}\text{vec}(\mathbf{e}_{t-j}\mathbf{e}'_{t-j}-\mathbf{\Sigma}_e)$ and covariances $n^{-1/2}\sum_{t=1}^{[nr]}\text{vec}(\mathbf{e}_{t-j}\mathbf{e}'_{t-j-s})$ above defined in terms of 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

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

where $V_v(r)$ a first-level Brownian Bridge process based on $B_v(r)$, with long-run variance

$$\omega_v^2=\mathbf{g}'_0(1)E[\text{vec}(\mathbf{e}_t\mathbf{e}'_t-\mathbf{\Sigma}_e)\text{vec}(\mathbf{e}_t\mathbf{e}'_t-\mathbf{\Sigma}_e)']+4\sum_{i=1}^{\infty}\mathbf{g}'_i(1)E[\text{vec}(\mathbf{e}_t\mathbf{e}'_t)\text{vec}(\mathbf{e}_t\mathbf{e}'_t)']\mathbf{g}_i(1)$$

where $E[\text{vec}(\mathbf{e}_t\mathbf{e}'_t)\text{vec}(\mathbf{e}_t\mathbf{e}'_t)']=\mathbf{\Sigma}_e^0\otimes\mathbf{\Sigma}_e^0$, $\mathbf{\Sigma}_e^0=\text{diag}(\sigma_0^2,\sigma_1^2,\ldots,\sigma_k^2)$ and $\sigma_j^2=E[e_{j,t}^2]$, and $E[\text{vec}(\mathbf{e}_t\mathbf{e}'_t-\mathbf{\Sigma}_e)\text{vec}(\mathbf{e}_t\mathbf{e}'_t-\mathbf{\Sigma}_e)']$ is a square $(k+1)^2\times(k+1)^2$ matrix involving the fourth-order central moments $\eta_j=E[(e_{j,t}^2-\sigma_j^2)^2]$, j = 0, 1, …, k, and products of distinct variances $\sigma_j^2\sigma_i^2$, j, i = 0, 1, … k, $j\neq i$. Also, given $\mathbf{\Omega}_{\boldsymbol{\zeta}}^{1/2}$, the limiting distribution in (3.7) admits the decomposition

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

with $E[\mathbf{B}_k(r)W_v(r)]=\mathbf{\Omega}_{kk}^{1/2}E[\mathbf{W}_k(r)W_v(r)]=\mathbf{0}_k$. Thus, taking together (3.4), (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 statistic as

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

where $\hat{\omega}_{v.k,n}^2(q_n)=\hat{\omega}_{v,n}^2(q_n)-\hat{\boldsymbol{\omega}}'_{kv,n}(q_n)\hat{\boldsymbol{\gamma}}_{kv,n}(q_n)$ is a plug-in kernel estimate of $\omega_{v.k}^2$, the long-run variance of $v_t=u_t^2-\sigma_u^2$ conditional on $\boldsymbol{\varepsilon}_{k,t}$ under cointegration, given by

$$\hat{\omega}_{v,n}^2(q_n)=\sum_{h=-(n-1)}^{n-1}w(h/q_n)n^{-1}\sum_{t=|h|+1}^{n}\hat{v}_{t-|h|}\hat{v}_t \tag{3.10}$$

with $\hat{\boldsymbol{\gamma}}_{kv,n}(q_n)=\hat{\mathbf{\Omega}}_{kk,n}^{-1}(q_n)\hat{\boldsymbol{\omega}}_{kv,n}(q_n)$, and

$$\hat{\boldsymbol{\omega}}_{kv,n}(q_n)=n^{-1}\sum_{t=1}^{n}\hat{\mathbf{z}}_{k,t}\hat{v}_t+\sum_{h=1}^{n-1}w(h/q_n)n^{-1}\sum_{t=h+1}^{n}(\hat{\mathbf{z}}_{k,t-h}\hat{v}_t+\hat{\mathbf{z}}_{k,t}\hat{v}_{t-h})$$

the kernel-based estimators of the corresponding variances and covariances, with kernel function $w(\cdot)$ and bandwidth q_n.[8] 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 under the assumption of cointegration.

Next Proposition 3.1 establish the limiting null distribution of (3.9) in the cases $\mathbf{d}_{k,t}=\mathbf{0}_k$ or, at most, $\mathbf{d}_{k,t}=\mathbf{A}_{k,1}\neq\mathbf{0}_k$ (that is m = 1 and q = 0) relating to the structure of

[8] Observe that, alternatively, the conditional long-run variance estimator $\hat{\omega}_{v.k,n}^2(q_n)$ can also be computed as $\hat{\omega}_{v.k,n}^2(q_n)=\sum_{j=-(n-1)}^{n-1}w(jq_n^{-1})(n^{-1}\sum_{t=|j|+1}^{n}\hat{v}_{t,k}\hat{v}_{t-|j|,k})$, with $\hat{v}_{t,k}=\hat{v}_t-\hat{\boldsymbol{\gamma}}'_{kv,n}(q_n)\hat{\mathbf{z}}_{k,t}$ a Fully Modified (FM)-type correction of $\hat{v}_t=\hat{u}_t^2-\hat{\sigma}_n^2$.

the deterministic component underlying the observed integrated regressors.

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 and (3.5), 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 and integrated regressors containing at most a constant term the limiting distribution of the modified CUSUM of squares statistic in (3.9) is given by*

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

which is the supremum of the absolute value of a standard (first-level) Brownian Bridge, $J_v(r) = W_v(r) - rW_v(1)$.

Proof. This result follows directly from the above results, standard application of the Continuous Mapping Theorem, the consistency of the kernel estimates of the long-run variances and covariances, i.e., $\hat{\omega}^2_{v.k,n}(q_n) \to^p \omega^2_{v.k}$ and $\hat{\boldsymbol{\gamma}}_{kv,n}(q_n) \to^p \boldsymbol{\gamma}_{kv,n}$, and the fact that $\mathbf{x}_{k,t} - (t/n)\mathbf{x}_{k,n} = \boldsymbol{\eta}_{k,t} - (t/n)\boldsymbol{\eta}_{k,n} + \mathbf{d}_{k,t} - (t/n)\mathbf{d}_{k,n}$, with $\mathbf{d}_{k,t} - (t/n)\mathbf{d}_{k,n} = \mathbf{0}_k$ and $n^{-1/2}(\mathbf{d}_{k,t} - (t/n)\mathbf{d}_{k,n}) = O(n^{-1/2})$ when the deterministic component underlying the generating mechanism of the regressors contains at most a constant term.

Next we make some comments on this basic result. The first one refers to an important limitation of this result in more general situations related to the structure of the deterministic component underlying the observed deterministically trending integrated regressors when $m > 1$, in which case we propose the convenient modification required to account for this characteristics. The second one has to do with the correction for endogeneity in (3.9), and in particular with the computation of the correction factor $\hat{\boldsymbol{\gamma}}'_{kv,n}(q_n)(\mathbf{x}_{k,t} - (t/n)\mathbf{x}_{k,n})$ and the conditional long-run variance $\hat{\omega}^2_{v.k,n}(q_n)$ which depend on the long-run covariance $\hat{\boldsymbol{\omega}}_{kv,n}(q_n)$. Finally, the third comment refers to the properties of the CUSUM of squares measure based on the sequence of residuals obtained from alternative estimation methods other than OLS such as, e.g., FM-OLS estimates.

Remark 3.1. Given the general representation for the deterministically trending integrated regressors given in (2.1) and Assumption 2.1, $\mathbf{x}_{k,t} - (t/n)\mathbf{x}_{k,n} = \boldsymbol{\eta}_{k,t} - (t/n)\boldsymbol{\eta}_{k,n} + \mathbf{A}_k\mathbf{d}_t$, with $\mathbf{d}_t = \boldsymbol{\tau}_t - (t/n)\boldsymbol{\tau}_n = \boldsymbol{\Gamma}^{-1}_n\mathbf{d}_{nt}$ and $\mathbf{d}_{nt} = \boldsymbol{\tau}_{nt} - (t/n)\boldsymbol{\tau}_{nn}$, so that the leading term in the numerator of (3.9) can be decomposed as

$$\begin{aligned} n^{-1/2}\left\{\sum_{j=1}^{t}\hat{v}_j - \hat{\boldsymbol{\gamma}}'_{kv,n}(q_n)(\mathbf{x}_{k,t} - (t/n)\mathbf{x}_{k,n})\right\} &= n^{-1/2}\sum_{j=1}^{t}\hat{v}_j - \hat{\boldsymbol{\gamma}}'_{kv,n}(q_n)(\boldsymbol{\eta}_{k,nt} - (t/n)\boldsymbol{\eta}_{k,nn}) \\ &\quad - \hat{\boldsymbol{\gamma}}'_{kv,n}(q_n)\mathbf{A}_k n^{-1/2}\boldsymbol{\Gamma}^{-1}_n\mathbf{d}_{nt} \end{aligned}$$

where $\boldsymbol{\Gamma}^{-1}_n = \mathrm{diag}(\boldsymbol{\Gamma}^{-1}_{m,n}, \boldsymbol{\Gamma}^{-1}_{q,n})$, which is clearly dominated by the last term for $m > 1$. Thus, if we define $\hat{V}_t = \sum_{j=1}^{t}\hat{v}_j$ and $\hat{V}_{t,k} = \hat{V}_t - \hat{\boldsymbol{\gamma}}'_{kv,n}(q_n)(\mathbf{x}_{k,t} - (t/n)\mathbf{x}_{k,n})$, then $\hat{V}_{t,k}$ is taken as the dependent variable in the auxiliary regression

$$\hat{V}_{t,k} = \boldsymbol{\alpha}'\mathbf{d}_t + s_t \tag{3.12}$$

where $s_t = \hat{V}_t - \hat{\boldsymbol{\gamma}}'_{kv,n}(q_n)(\boldsymbol{\eta}_{k,t} - (t/n)\boldsymbol{\eta}_{k,n})$, with limiting distribution under cointegration given by $n^{-1/2}s_{[nr]} \Rightarrow \omega_{v.k}J_v(r)$. Computing the sequence of OLS residuals from (3.12) as

$$\hat{s}_t = \hat{V}_{t,k} - \mathbf{d}_t'\hat{\boldsymbol{\alpha}}_n = s_t - n^{1/2}\mathbf{d}_{nt}'\left(n^{-1}\sum_{j=1}^{n}\mathbf{d}_{nj}\mathbf{d}_{nj}'\right)^{-1} n^{-3/2}\sum_{j=1}^{n}\mathbf{d}_{nj}s_j$$

then we propose to modify the test statistic in (3.9) as

$$C\hat{S}_n(m+q) = \frac{1}{\sqrt{n}\hat{\omega}_{v.k,n}(q_n)}\max_{t=1,\dots,n}|\hat{s}_t| \qquad (3.13)$$

with limiting distribution under cointegration given by the supremum of a demeaned Brownian Bridge process such as

$$C\hat{S}_n(m+q) \Rightarrow \sup_{r\in[0,1]}\left|J_v(r) - \boldsymbol{\tau}'(r)\left(\int_0^1\boldsymbol{\tau}(s)\boldsymbol{\tau}'(s)ds\right)^{-1}\int_0^1\boldsymbol{\tau}(s)J_v(s)ds\right| \qquad (3.14)$$

which depends on the order of the polinomial trend function $\boldsymbol{\tau}_t = (\boldsymbol{\tau}_{m,t}', \boldsymbol{\tau}_{q,t}')'$, but not on the number of integrated regressors, *k*. The quantiles of the limiting distribution in (3.14) are tabulated and presented in Table B.3 in Appendix B for different sample sizes and assumed orders for $\boldsymbol{\tau}_t$, with *m*+*q* = 1, ..., 5.

Remark 3.2. Given the definition of the long-run covariance between $\boldsymbol{\varepsilon}_{k,t}$ and $u_t^2 - \sigma_u^2$, $\boldsymbol{\omega}_{kv} = \sum_{h=0}^{\infty}E[\boldsymbol{\varepsilon}_{k,t-h}u_t^2] + \sum_{h=1}^{\infty}E[\boldsymbol{\varepsilon}_{k,t}u_{t-h}^2]$, under the linear process assumption for the sequence $\boldsymbol{\xi}_t = (u_t, \boldsymbol{\varepsilon}_{k,t}')'$ in (2.3) and the iid property of $\mathbf{e}_t = (e_{0,t}, e_{1,t}, \dots, e_{k,t})'$, we have that

$$E[\boldsymbol{\varepsilon}_{k,t-h}u_t^2] = \sum_{j=0}^{\infty}\mathbf{D}_{k,j}E[\mathbf{e}_t\text{vec}(\mathbf{e}_t\mathbf{e}_t')']\mathbf{g}_{0,h+j}$$

and

$$E[\boldsymbol{\varepsilon}_{k,t}u_{t-h}^2] = \sum_{j=0}^{\infty}\mathbf{D}_{k,h+j}E[\mathbf{e}_t\text{vec}(\mathbf{e}_t\mathbf{e}_t')']\mathbf{g}_{0,j}$$

where the $(k+1)^2\times(k+1)$ matrix $E[\text{vec}(\mathbf{e}_t\mathbf{e}_t')\mathbf{e}_t']$ is of the form

$$E[\text{vec}(\mathbf{e}_t\mathbf{e}_t')\mathbf{e}_t'] = (E[e_{0,t}^3]\mathbf{I}_{k+1}^{(1,1)}, E[e_{1,t}^3]\mathbf{I}_{k+1}^{(2,2)}, \dots, E[e_{k,t}^3]\mathbf{I}_{k+1}^{(k+1,k+1)})'$$

with $\mathbf{I}_{k+1}^{(j,j)}$, *j* = 1, 2, …, *k*+1 a square *k*+1 matrix made of zeros except for a unit value in the *j*th-position of the diagonal, so that all these expected values mainly depend on the characteristics of the distribution of the error terms $e_{j,t}$, *j* = 0, 1, ..., *k*, with $\boldsymbol{\omega}_{kv} = \mathbf{0}_k$ under symmetry of the distribution of the error terms $\mathbf{e}_t$. In any other case and from a practical point of view when using the estimator $\hat{\boldsymbol{\omega}}_{kv,n}(q_n)$, the phenomenon known as *supernormality* (see, e.g., Leybourne et.al. (1996)) implies that the residuals $\hat{\mathbf{z}}_{k,t}, \hat{u}_t$ tend to be symmetrized for large sample sizes even if the true innovations are not. However, in small or moderate samples we must not rely on this asymptotic property and perform the computation of the required elements for those corrections.

Remark 3.3. As an alternative to the OLS-version of the test statistic in (3.9), we consider the computation of the scaled partial sum of squared and centered FM-OLS residuals, $\hat{u}_t^+ = u_t^+ - n^{-\kappa}\mathbf{m}_{nt}'\hat{\boldsymbol{\Theta}}_n^+$, as was detailed in section 2.1. Taking into account that these estimates are consistent under cointegration with the same rates as with the OLS estimation, and that the FM-OLS correction of the regression errors is given by $u_t^+ = u_t - \hat{\boldsymbol{\gamma}}'_{ku,n}\hat{\mathbf{z}}_{k,t} = z_t - (\hat{\boldsymbol{\gamma}}_{ku,n} - \boldsymbol{\gamma}_{ku})'\boldsymbol{\varepsilon}_{k,t} - \hat{\boldsymbol{\gamma}}'_{ku,n}\mathbf{F}_{k,nt}$, with $z_t = u_t - \boldsymbol{\gamma}'_{ku}\boldsymbol{\varepsilon}_{k,t}$ and $\hat{\boldsymbol{\gamma}}_{ku,n} - \boldsymbol{\gamma}_{ku} = o_p(1)$ under cointegration, then the scaled partial sum process of squared and centered FM-OLS residuals can be decomposed as

$$
\begin{aligned}
n^{-1/2}\sum_{t=1}^{[nr]}\hat{v}_t^+ &= n^{-1/2}\sum_{t=1}^{[nr]}(z_t^2 - \sigma_n^2) \\
&\quad + (\hat{\boldsymbol{\gamma}}_{ku,n} - \boldsymbol{\gamma}_{ku})'\left\{n^{-1/2}\sum_{t=1}^{[nr]}(\boldsymbol{\varepsilon}_{k,t}\boldsymbol{\varepsilon}'_{k,t} - \boldsymbol{\Sigma}_{kk}) - \frac{[nr]}{n}n^{-1/2}\sum_{t=1}^{n}(\boldsymbol{\varepsilon}_{k,t}\boldsymbol{\varepsilon}'_{k,t} - \boldsymbol{\Sigma}_{kk})\right\}(\hat{\boldsymbol{\gamma}}_{ku,n} - \boldsymbol{\gamma}_{ku}) \\
&\quad + n^{1/2-2\kappa}\hat{\boldsymbol{\Theta}}_n'^{+}\left(n^{-1}\sum_{t=1}^{[nr]}\mathbf{m}_{nt}\mathbf{m}'_{nt} - \frac{[nr]}{n}n^{-1}\sum_{t=1}^{n}\mathbf{m}_{nt}\mathbf{m}'_{nt}\right)\hat{\boldsymbol{\Theta}}_n^{+} \\
&\quad - 2(\hat{\boldsymbol{\gamma}}_{ku,n} - \boldsymbol{\gamma}_{ku})'\left\{n^{-1/2}\sum_{t=1}^{[nr]}(\boldsymbol{\varepsilon}_{k,t}u_t - \boldsymbol{\sigma}_{ku}) - \frac{[nr]}{n}n^{-1/2}\sum_{t=1}^{n}(\boldsymbol{\varepsilon}_{k,t}u_t - \boldsymbol{\sigma}_{ku})\right\} \\
&\quad + 2(\hat{\boldsymbol{\gamma}}_{ku,n} - \boldsymbol{\gamma}_{ku})'\left\{n^{-1/2}\sum_{t=1}^{[nr]}(\boldsymbol{\varepsilon}_{k,t}\boldsymbol{\varepsilon}'_{k,t} - \boldsymbol{\Sigma}_{kk}) - \frac{[nr]}{n}n^{-1/2}\sum_{t=1}^{n}(\boldsymbol{\varepsilon}_{k,t}\boldsymbol{\varepsilon}'_{k,t} - \boldsymbol{\Sigma}_{kk})\right\}\boldsymbol{\gamma}_{ku} \\
&\quad - 2n^{1/2-2\kappa}\hat{\boldsymbol{\Theta}}_n'^{+}\left(n^{-(1-\kappa)}\sum_{t=1}^{[nr]}\mathbf{m}_{nt}z_t - \frac{[nr]}{n}n^{-(1-\kappa)}\sum_{t=1}^{n}\mathbf{m}_{nt}z_t\right) \\
&\quad + 2n^{-\kappa}\hat{\boldsymbol{\Theta}}_n'^{+}\left(n^{-1/2}\sum_{t=1}^{[nr]}\mathbf{m}_{nt}\boldsymbol{\varepsilon}'_{k,t} - \frac{[nr]}{n}n^{-1/2}\sum_{t=1}^{n}\mathbf{m}_{nt}\boldsymbol{\varepsilon}'_{k,t}\right)(\hat{\boldsymbol{\gamma}}_{ku,n} - \boldsymbol{\gamma}_{ku}) + O_p(n^{-\kappa})
\end{aligned}
$$

where $\sigma_n^2 = n^{-1}\sum_{t=1}^{n} z_t^2$ and the last term $O_p(n^{-\kappa})$ collects all the elements involving functionals of $\mathbf{F}_{k,nt}$. Also, all the terms involving the FM-OLS estimation errors, $\hat{\boldsymbol{\Theta}}_n^{+}$, are asymptotically negligible under cointegration, so that we can write

$$
n^{-1/2}\sum_{t=1}^{[nr]}\hat{v}_t^+ = n^{-1/2}\sum_{t=1}^{[nr]}(z_t^2 - \sigma_n^2) + O_p(n^{-1/2}) = n^{-1/2}\sum_{t=1}^{[nr]}(z_t^2 - \sigma_n^2) + o_p(1)
$$

where the first term can be decomposed as

$$
\begin{aligned}
n^{-1/2}\sum_{t=1}^{[nr]}(z_t^2 - \sigma_n^2) &= n^{-1/2}\sum_{t=1}^{[nr]}(u_t^2 - \sigma_u^2) - \frac{[nr]}{n}n^{-1/2}\sum_{t=1}^{n}(u_t^2 - \sigma_u^2) \\
&\quad + \boldsymbol{\gamma}'_{ku}\left\{n^{-1/2}\sum_{t=1}^{[nr]}(\boldsymbol{\varepsilon}_{k,t}\boldsymbol{\varepsilon}'_{k,t} - \boldsymbol{\Sigma}_{kk}) - \frac{[nr]}{n}n^{-1/2}\sum_{t=1}^{n}(\boldsymbol{\varepsilon}_{k,t}\boldsymbol{\varepsilon}'_{k,t} - \boldsymbol{\Sigma}_{kk})\right\}\boldsymbol{\gamma}_{ku} \\
&\quad - 2\boldsymbol{\gamma}'_{ku}\left\{n^{-1/2}\sum_{t=1}^{[nr]}(\boldsymbol{\varepsilon}_{k,t}u_t - \boldsymbol{\sigma}_{ku}) - \frac{[nr]}{n}n^{-1/2}\sum_{t=1}^{n}(\boldsymbol{\varepsilon}_{k,t}u_t - \boldsymbol{\sigma}_{ku})\right\} \\
&= n^{-1/2}\sum_{t=1}^{[nr]}v_{nt} + (\boldsymbol{\gamma}'_{ku}\otimes\boldsymbol{\gamma}'_{ku})\left\{n^{-1/2}\sum_{t=1}^{[nr]}\mathrm{vec}(\mathbf{E}_t) - \frac{[nr]}{n}n^{-1/2}\sum_{t=1}^{n}\mathrm{vec}(\mathbf{E}_t)\right\} \\
&\quad - 2\boldsymbol{\gamma}'_{ku}\left\{n^{-1/2}\sum_{t=1}^{[nr]}(\boldsymbol{\varepsilon}_{k,t}u_t - \boldsymbol{\sigma}_{ku}) - \frac{[nr]}{n}n^{-1/2}\sum_{t=1}^{n}(\boldsymbol{\varepsilon}_{k,t}u_t - \boldsymbol{\sigma}_{ku})\right\}
\end{aligned}
$$

with $\mathbf{E}_t = \boldsymbol{\varepsilon}_{k,t}\boldsymbol{\varepsilon}'_{k,t} - \boldsymbol{\Sigma}_{kk}$, implying that the resulting limiting distribution under cointegration will depend on some additional nuisance parameters such as $\boldsymbol{\gamma}_{ku} = \boldsymbol{\Omega}_{kk}^{-1}\boldsymbol{\omega}_{ku}$ and the number of integrated regressors through the dimensions of the vectors vec($\mathbf{E}_t$), $k^2\times 1$, and $\boldsymbol{\varepsilon}_{k,t}u_t - \boldsymbol{\sigma}_{ku}$ $k\times 1$. The same result follows when using the residuals obtained from any other asymptotically equivalent estimation method as, e.g., the IM-OLS residuals. This result implies a clear advantage of the use of OLS residuals for the computation of the proposed test statistic.

4. Some sources of size distortions and consistency analysis

The first part of this section is devoted to the analytic study of our test statistic in terms of evaluating the effects of some different sources of size distortions in finite samples caused by a highly persistent, but stationary, regression error term u_t following an AR(1) process with a root approaching unity at a moderate rate or, alternatively, by introducing a local-to-unity representation for a moving average (MA) root. We consider three of these types of representations that fall into the general class of summable processes of order ranging from [0, 1) characterized by the fact that u_t locally behaves as a stationary sequence. The concept of a summable stochastic processes has been recently formalized by Berenguer-Rico and Gonzalo (2014), where for a zero-mean stochastic process ζ_t the order of summability, γ, is the minimum real number that makes $n^{-(1/2+\gamma)}\sum_{t=1}^{n}\zeta_t$ bounded in probability, and it is denoted as $S(\gamma)$. It is clear that for a standard stationary process, I(0) = $S(0)$, while that for a integrated process, I(1) = $S(1)$. The second subsection deals with the consistency of the testing procedure against the alternative of no cointegration when the regression error term follows a fixed unit root process, although similar conclusions could be obtained under a standard local-to-unity representation for the autoregressive coefficient.

4.1 Size distortions

As for the testing procedures for the null of cointegration based on measures of excessive divergence of the partial sum of residuals from the estimation of the cointegrating regression not consistent with the stationarity of the regression error terms,[9] for our test statistic it is also expected that the main source of distortion of the empirical size in finite samples comes from the treatment of serial correlation. Palma and Zevallos (2004) study the correlation structure for the squares of a time series satisfying a linear filter, such as $u_t = c(L)e_t$ with e_t a sequence of uncorrelated but not necessarily independent variables, with mean zero, finite variance and kurtosis $\eta_e = E[e_t^4]/\sigma_e^4 < \infty$. From their results we have that

$$\gamma_v(h) = 2\gamma_u^2(h) + \sigma_e^4\left\{-2\eta_e\sum_{j=0}^{\infty}c_j^2c_{j+h}^2 + (\eta_e - 1)\sum_{j=0}^{\infty}\sum_{i=0}^{\infty}c_j^2c_i^2\rho_{e,2}(h+i-j)\right.$$
$$\left. +2(\eta_e - 1)\sum_{j=0}^{\infty}\sum_{i=0}^{\infty}c_jc_{j+h}c_ic_{i+h}\rho_{e,2}(i-j)\right\}$$

where $\gamma_v(h) = E[(u_t^2 - \sigma_u^2)(u_{t-h}^2 - \sigma_u^2)]$, with $\rho_{e,2}(j)$ the autocorrelation function of e_t^2, which reduces to

$$\gamma_v(h) = 2\gamma_u^2(h) + \sigma_e^4(\eta_e - 3)\sum_{j=0}^{\infty}c_j^2c_{j+h}^2 = 2\gamma_u^2(h) + \sigma_e^4(\eta_u - 3)\left(\sum_{j=0}^{\infty}c_j^2c_{j+h}^2 \Big/ \sum_{j=0}^{\infty}c_j^4\right)$$

under iid noise, with $\eta_u = 3 + (\eta_e - 3)g_0^{-2}(1)\sum_{j=0}^{\infty}c_j^4$ the kurtosis of u_t. These results

[9] For more detailed and exhaustive studies on these effects see, e.g., Carrion-i-Silvestre and Sansó (2006) and Müller (2005).

imply that, in the case of a short-memory stationary filter with $c_j \sim c^j$ for some $|c| < 1$, the decaying rate of the autocorrelation function of the squares is twice as fast as the decaying rate of the autocorrelation function of the original series, which determines an immediate impact on our results related to the computation of the long-run variances and covariances involved in the construction of the test statistic in (3.9). Simulation results, presented in Appendix C, seems to confirm this conjecture in terms of requiring a relatively small value of the bandwidth parameter, q_n, to obtain a very good performance of the test statistic under stationary and strongly autocorrelated regression errors.

Next we examine the expected behavior of the leading term in (3.9), $n^{-1/2}\sum_{j=1}^{t}\hat{v}_j = n^{-1/2}\sum_{j=1}^{t}(\hat{u}_t^2 - \hat{\sigma}_{u,n}^2)$, under two different constructions characterizing the local-to-stationarity behavior of the regression error terms.

4.1.1. The near stationary case

To introduce the effect of moderately serially correlated errors in a cointegrating regression, Kurozumi and Hayakawa (2009) consider the framework proposed by Giraitis and Phillips (2006), and Phillips and Magdalinos (2007a, b)[10] where the AR coefficient α is moderately close to 1. This can be modeled by the so-called m local-to-unity system, defined as $\alpha = \alpha_m = 1 - \lambda/m_n$, with $\lambda > 0$, $m_n \to \infty$, and $m_n/n \to 0$ as $n \to \infty$. A convenient parameterization of m_n is given by $m_n = n^{\gamma}$, with $\gamma \in (0,1)$. By Lemma 3.2 in Phillips and Magdalinos (2007b) and Lemma 1(d) in Kurozumi and Hayakawa (2009), $u_t = O_p(m_n^{1/2}) = O_p(n^{\gamma/2})$, and hence $n^{-1/2}u_t = O_p(n^{-(1-\gamma)/2}) = o_p(1)$ for any value $\gamma \in (0, 1)$, representing an intermediate case between standard stationarity and non-stationarity. Taking into account the following representations

$$(1-\alpha_n)\sum_{t=1}^{[nr]} u_t = \sum_{t=1}^{[nr]} \upsilon_t + \alpha_n n^{\gamma/2}[n^{-\gamma/2}(u_0 - u_{[nr]})]$$

$$(1-\alpha_n^2)\sum_{t=1}^{[nr]} u_t^2 = \sum_{t=1}^{[nr]} \upsilon_t^2 + 2\alpha_n n^{1/2+d}\left(n^{-(1/2+d)}\sum_{t=1}^{[nr]} u_{t-1}\upsilon_t\right) + \alpha_n^2 n^{\gamma}[n^{-\gamma}(u_0^2 - u_{[nr]}^2)]$$

and

$$n^{-1}\sum_{t=h+1}^{n} u_{t-h}u_t = \alpha_n^h n^{\gamma}\left(n^{-\gamma}\sigma_{u,n}^2 - n^{-1}\sum_{t=n-h+1}^{n}(n^{-\gamma/2}u_t)^2\right) + n^{d-1/2}\sum_{j=1}^{h}\alpha_n^{h-j} n^{-(1/2+d)}\sum_{t=1}^{n-h} u_t\upsilon_{t+j}$$

where $n^{-1}\sum_{t=n-h+1}^{n}(n^{-\gamma/2}u_t)^2 = O_p(h/n)$ and $n^{-(1/2+d)}\sum_{t=1}^{n-h} u_t\upsilon_{t+j} = O_p(1)$, with $d = \gamma/2$ when υ_t is an iid sequence, and $d = 1/2$ in the case of weakly dependent errors υ_t (see Phillips and Magdalinos (2007b)), then we obtain the limiting results

$$\frac{1}{m_n\sqrt{n}}\sum_{t=1}^{[nr]} u_t = \frac{1}{n^{1/2+\gamma}}\sum_{t=1}^{[nr]} u_t \Rightarrow (1/\lambda)B_{\upsilon}(r) \quad (4.1)$$

and

[10] Giraitis and Phillips (2006) and Phillips and Magdalinos (2007a) study the properties of the estimator of the first order autocorrelation for an observed univariate time series under near stationarity driven by iid noise, while Phillips and Magdalinos (2007b) extend the analysis to the case of weakly dependent stationary errors following a linear process.

$$(1/m_n)\sigma^2_{u,n} = \frac{1}{m_n \cdot n}\sum_{t=1}^{n} u_t^2 = \frac{1}{n^{1+\gamma}}\sum_{t=1}^{n} u_t^2 \to^p (1/2\lambda)\omega^2_\upsilon \qquad (4.2)$$

so that from (4.1), although stationary, u_t is summable of order $\gamma \in (0, 1)$. Also, from (2.10) and (4.1), we have that the normalized OLS estimation error of the model parameters is given now by

$$\tilde{\boldsymbol{\Theta}}_n = m_n^{-1}\hat{\boldsymbol{\Theta}}_n = n^{1/2-\gamma}\mathbf{W}'_n(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) = \left(n^{-1}\sum_{t=1}^{n}\mathbf{m}_{nt}\mathbf{m}'_{nt}\right)^{-1} n^{-(1/2+\gamma)}\sum_{t=1}^{n}\mathbf{m}_{nt}u_t$$

implying that the OLS estimator of $\boldsymbol{\beta}_k$ in (2.4) consistently estimate the cointegrating vector parameters but at a smaller rate than in the standard case, i.e. $n^{1-\gamma}(\hat{\boldsymbol{\beta}}_{k,n} - \boldsymbol{\beta}_k) = O_p(1)$. Thus, given the sequence of OLS residuals $\hat{u}_t = u_t - n^{-\kappa}\mathbf{m}'_{nt}\hat{\boldsymbol{\Theta}}_n$ where the index κ takes now the value $\kappa = 1/2 - \gamma$, we have that

$$\begin{aligned} n^{-1/2}\sum_{t=1}^{[nr]}(\hat{u}_t^2 - \hat{\sigma}^2_{u,n}) &= n^{-1/2}\sum_{t=1}^{[nr]}(u_t^2 - \sigma_n^2) \\ &+ n^{-(1/2-2\gamma)}\hat{\boldsymbol{\Theta}}'_n\left(n^{-1}\sum_{t=1}^{[nr]}\mathbf{m}_{nt}\mathbf{m}'_{nt} - \frac{[nr]}{n}n^{-1}\sum_{t=1}^{n}\mathbf{m}_{nt}\mathbf{m}'_{nt}\right)\hat{\boldsymbol{\Theta}}_n \\ &- 2n^{-(1/2-2\gamma)}\hat{\boldsymbol{\Theta}}'_n\left(n^{1/2+\gamma}\sum_{t=1}^{[nr]}\mathbf{m}_{nt}u_t - \frac{[nr]}{n}n^{1/2+\gamma}\sum_{t=1}^{n}\mathbf{m}_{nt}u_t\right) \end{aligned} \qquad (4.3)$$

where the last two terms in (4.3) are asymptotically negligible for $\gamma < 1/4$, while that for the first term we get

$$\begin{aligned} n^{-1/2}\sum_{t=1}^{[nr]}(u_t^2 - \sigma_n^2) &= \frac{n^\gamma}{2\lambda(1-\lambda/2n^\gamma)}\Bigg\{n^{-1/2}\sum_{t=1}^{[nr]}(\upsilon_t^2 - \sigma^2_{\upsilon,n}) \\ &+ \alpha_n^2 n^{\gamma-1/2}\left(n^{-\gamma}u_0^2\left(1 - \frac{[nr]}{n}\right) - n^{-\gamma}\left(u^2_{[nr]} - \frac{[nr]}{n}u_n^2\right)\right) \\ &+ 2\alpha_n n^d\left(n^{-(1/2+d)}\sum_{t=1}^{[nr]}u_{t-1}\upsilon_t - \frac{[nr]}{n}n^{-(1/2+d)}\sum_{t=1}^{n}u_{t-1}\upsilon_t\right)\Bigg\} \end{aligned} \qquad (4.4)$$

so that $n^{-1/2}\sum_{t=1}^{[nr]}(u_t^2 - \sigma_n^2) = O_p(n^{\gamma+d})$, and hence both $n^{-1/2}\sum_{t=1}^{[nr]}\hat{u}_t$ and $n^{-1/2}\sum_{t=1}^{[nr]}(\hat{u}_t^2 - \hat{\sigma}^2_{u,n})$ will diverge with the sample size. These results allows to explain and quantify the size distortions occurred in the presence of highly correlated regression error terms and reflected not only in the computation of the long-run variances and covariances required to built these test statistics but also in the behavior of the CUSUM and CUSUM of squares measures.

4.1.2. The case of a nearly integrated process with a local-to-unity MA root

As an alternative to the above construction, we next consider the formulation proposed by Nabeya and Perron (1994) as follows

$$u_t = \alpha_n u_{t-1} + \upsilon_t$$
$$\upsilon_t = \zeta_t + \theta_n \zeta_{t-1}$$

with $\alpha_n = 1 - \lambda_1 n^{-1}$, $\theta_n = -1 + \lambda_2 n^{-1/2}$ and ζ_t a zero mean stationary sequence with finite variance $\sigma^2_\zeta = E[\zeta_t^2]$, where, in the limit, the AR and MA roots cancel and the process is stationary. Simple manipulation of these terms allows to write u_t as

$$u_t = a_0\alpha_n^{t-1} + a_n\zeta_t + b_n\xi_t$$

with $a_0 = u_0\alpha_n + \zeta_0\theta_n$ which is assumed to be zero, $a_n = \alpha_n^{-1}(1-\lambda_2 n^{-1/2})$, $b_n = 1 - a_n$, and $\xi_t = \alpha_n\xi_{t-1} + \zeta_t$, where $n^{-1/2}\xi_{[nr]} \Rightarrow J_{\lambda_1}(r) = \int_0^r e^{-\lambda_1(r-s)}dB_\zeta(s)$ (see Phillips (1987b)). Taking these results, together with the fact that $\sqrt{n}b_n \to \lambda_2$, then we have that $u_t = O_p(1)$ but $n^{-1/2}\sum_{t=1}^{[nr]} u_t = O_p(n^{1/2})$, so that u_t is summable of order 1/2, with $n^{-1}\sum_{t=1}^{[nr]} u_t \Rightarrow \lambda_2\int_0^r J_{\lambda_1}(s)ds$, implying that the OLS estimator of the cointegrating vector is still consistently estimated at the usual rate $n^{1/2}$ (κ = 0 in (2.10)), but with a very different limiting distribution. Hence, all the CUSUM-type statistics described in section 2.2 must diverge with the sample size given that when based on OLS residuals we have $n^{-1}\sum_{t=1}^{[nr]}\hat{u}_t = n^{-1}\sum_{t=1}^{[nr]}u_t - n^{-1}\sum_{t=1}^{[nr]}\mathbf{m}'_{nt}\hat{\Theta}_n = O_p(1)$. Moreover, it is easy to show that the sample variance of u_t has the following weak limit $\sigma_{u,n}^2 \Rightarrow \sigma_\zeta^2 + \lambda_2^2\int_0^1 J_{\lambda_1}^2(s)ds$, and $n^{-1/2}\sum_{t=1}^{[nr]}(u_t^2 - \sigma_{u,n}^2) = O_p(n^{1/2})$, which implies that the CUSUM of squares measure for the sequence u_t diverges with the sample size at the same rate as the CUSUM for the levels. For this construction, the CUSUM of squares for the OLS residuals loses its main attractive feature of being independent from the components of the estimated model given that (3.4) is now given by

$$n^{-1}\sum_{t=1}^{[nr]}(\hat{u}_t^2 - \hat{\sigma}_{u,n}^2) = n^{-1}\sum_{t=1}^{[nr]}(u_t^2 - \sigma_{u,n}^2) + \hat{\Theta}'_n\left(n^{-1}\sum_{t=1}^{[nr]}\mathbf{m}_{nt}\mathbf{m}'_{nt} - \frac{[nr]}{n}n^{-1}\sum_{t=1}^{n}\mathbf{m}_{nt}\mathbf{m}'_{nt}\right)\hat{\Theta}_n$$
$$-2\hat{\Theta}'_n\left(n^{-1}\sum_{t=1}^{[nr]}\mathbf{m}_{nt}u_t - \frac{[nr]}{n}n^{-1}\sum_{t=1}^{n}\mathbf{m}_{nt}u_t\right)$$

where $n^{-1}\sum_{t=1}^{[nr]}(u_t^2 - \sigma_{u,n}^2) \Rightarrow \lambda_2^2(\int_0^r J_{\lambda_1}^2(s)ds - r\int_0^1 J_{\lambda_1}^2(s)ds)$. This construction represents an intermediate case between the near stationary formulation in 4.1.1 and the usual local-to-unity parameterization that follows when $\sigma_\zeta^2 = 0$, and hence could be used to evaluate the size performance of these test statistics in the vicinity of stationarity in finite samples.

4.2 Power behavior and consistency

To study the power performance of the proposed test statistic, we follow Jansson's (2005) proposal that considers a useful modification of the standard n local-to-unity system (that is, when $m_n = n$) to parameterize a local moving average (MA) unit root characterizing the behavior of the regression error term under near cointegration. To that end, if we take the first difference of u_t as $\Delta u_t = \alpha\Delta u_{t-1} + \Delta\upsilon_t$, then it can be rewritten as $\Delta u_t = (1-\rho L)\upsilon_t$ for α = 0 and ρ = 1. If instead of the fixed positive MA unit root we consider a local-to-unity representation as $\rho = \rho_n = 1 - \lambda n^{-1}$, for $\lambda \geq 0$, then the u_t admits the representation $u_t = u_0 + (1-\rho L)\sum_{i=0}^{t-1}\upsilon_{t-i} = a_0 + \rho_n\upsilon_t + (1-\rho_n)V_t$, with $a_0 = u_0 - \rho\upsilon_0$ and $V_t = \sum_{j=1}^{t}\upsilon_j$, then

$$n^{-1/2}\sum_{t=1}^{[nr]}u_t = a_0 n^{-1/2}[nr] + \rho_n n^{-1/2}\sum_{t=1}^{[nr]}\upsilon_t + (1-\rho_n)\sum_{t=1}^{[nr]}(n^{-1/2}V_t)$$
$$\Rightarrow B_{\lambda,\upsilon}(r) = B_\upsilon(r) + \lambda\int_0^r B_\upsilon(s)ds$$

under the necessary assumption $a_0 = 0$ on the initial values, so that λ = 0 corresponds to usual results under standard cointegration. For any value of λ > 0, the second term appearing in this limiting distribution determines a displacement of the limiting distributions under cointegration corresponding to a certain degree of excessive persistence of the regression error sequence. In any case, this framework does not cause any change in the rates of consistent estimation of the model parameters or the regression errors, so that the scaled partial sum of squared and centered OLS residuals in the leading term of the test statistic in (3.9) is governed by the behavior of

$$n^{-1/2}\sum_{t=1}^{[nr]} v_{nt} = \rho_n^2\left\{n^{-1/2}\sum_{t=1}^{[nr]}(\upsilon_t^2-\sigma_\upsilon^2)-\frac{[nr]}{n}n^{-1/2}\sum_{t=1}^{n}(\upsilon_t^2-\sigma_\upsilon^2)\right\}$$
$$+(1-\rho_n)^2 n^{3/2}\left\{n^{-1}\sum_{t=1}^{[nr]}(n^{-1/2}V_t)^2-\frac{[nr]}{n}n^{-1}\sum_{t=1}^{n}(n^{-1/2}V_t)^2\right\}$$
$$+2\rho_n(1-\rho_n)n^{1/2}\left\{n^{-1}\sum_{t=1}^{[nr]}V_t\upsilon_t-\frac{[nr]}{n}n^{-1}\sum_{t=1}^{n}V_t\upsilon_t\right\}$$

where $1-\rho_n=\lambda n^{-1}$ implying that the last two terms in the right-hand side of this expression are of order $O_p(n^{-1/2})$ and hence asymptotically negligible, determining the same weak limit result as in (3.7) irrespective of the value of λ. This result can be interpreted as a certain degree of robustness against this type of departure from the standard stationarity situation or, alternatively, that this construction is only appropriate to characterize local departures from the stationarity situation for the levels of the series. In fact, given the expressions

$$u_t^2=\rho_n^2\upsilon_t^2+n^{-1/2}(n^{-1}\lambda^2(n^{-1/2}V_t)+2\lambda\rho_n(n^{-1/2}V_t)\upsilon_t)$$

and

$$\sigma_{u,n}^2=\rho_n^2\sigma_{\upsilon,n}^2+n^{-1}\left\{n^{-1/2}\lambda^2 n^{-1}\sum_{t=!}^{n}(n^{-1/2}V_t)+2\lambda\rho_n n^{-1}\sum_{t=!}^{n}V_t\upsilon_t\right\}$$

it is immediate to observe that u_t^2 and $\sigma_{u,n}^2$ both asymptotically behave as υ_t^2 and $\sigma_{\upsilon,n}^2$, respectively. Alternatively, if we modify the assumption on the structure of the MA unit root as $\rho=\rho_n=1-\lambda n^{-\gamma}$, with $\lambda \geq 0$ and $\gamma \in (0, 1)$, this allow us to accommodate this structure for the squared series, implying that the CUSUM and CUSUM of squares theoretical measures based on u_t are now given by

$$n^{-1/2}\sum_{t=1}^{[nr]}u_t=\rho_n n^{-1/2}\sum_{t=1}^{[nr]}\upsilon_t+\lambda n^{1-\gamma}\left\{n^{-1}\sum_{t=1}^{[nr]}(n^{-1/2}V_t)\right\}=O_p(n^{1-\gamma})$$

and

$$n^{-1/2}\sum_{t=1}^{[nr]} v_{nt} = \rho_n^2\left\{n^{-1/2}\sum_{t=1}^{[nr]}(\upsilon_t^2-\sigma_\upsilon^2)-\frac{[nr]}{n}n^{-1/2}\sum_{t=1}^{n}(\upsilon_t^2-\sigma_\upsilon^2)\right\}$$
$$+\lambda^2 n^{-2\gamma+3/2}\left\{n^{-1}\sum_{t=1}^{[nr]}(n^{-1/2}V_t)^2-\frac{[nr]}{n}n^{-1}\sum_{t=1}^{n}(n^{-1/2}V_t)^2\right\}$$
$$+2\lambda\rho_n n^{-\gamma+1/2}\left\{n^{-1}\sum_{t=1}^{[nr]}V_t\upsilon_t-\frac{[nr]}{n}n^{-1}\sum_{t=1}^{n}V_t\upsilon_t\right\}$$

so that $n^{-1/2}\sum_{t=1}^{[nr]}u_t$ diverges at the given rate for any value of γ in the subset $\gamma \in (0, 1)$ and $\lambda > 0$, due to the closer proximity to non-stationarity for any sample size implied by this parameterization. Similarly $n^{-1/2}\sum_{t=1}^{[nr]}v_{nt}=O_p(n^{-2\gamma+3/2})$ and diverging with the sample size for any value of γ in the interval $0 < \gamma < 3/4$, while that for $3/4 < \gamma < 1$ the last two terms in the right-hand side are again asymptotically negligible. Only when γ takes exactly the value $\gamma = 3/4$ we get

$$n^{-1/2}\sum_{t=1}^{[nr]} v_{nt} = n^{-1/2}\sum_{t=1}^{[nr]}(\upsilon_t^2 - \sigma_\upsilon^2) - \frac{[nr]}{n} n^{-1/2}\sum_{t=1}^{n}(\upsilon_t^2 - \sigma_\upsilon^2)$$
$$+\lambda^2\left\{n^{-1}\sum_{t=1}^{[nr]}(n^{-1/2}V_t)^2 - \frac{[nr]}{n} n^{-1}\sum_{t=1}^{n}(n^{-1/2}V_t)^2\right\} + O_p(n^{-1/4})$$

with

$$n^{-1/2}\sum_{t=1}^{[nr]} v_{nt} \Rightarrow V_{v,\lambda}(r) = V_v(r) + \lambda^2\left\{\int_0^r B_\upsilon^2(s) - r\int_0^1 B_\upsilon^2(s)\right\}$$

where $V_v(r)$ is as in (3.7), with $B_v(r)$ the weak limit of $n^{-1/2}\sum_{t=1}^{[nr]}(\upsilon_t^2 - \sigma_\upsilon^2)$. We employ this construction at the end of section 5 to numerically compare the power performance of our test statistic with some other existing and popular testing procedures in practical applications.

Finally we study the behavior of the proposed statistic under the alternative of no cointegration, when the regression error term u_t contains a unit root (i.e. $\alpha = 1$). In this case, the sequence of OLS residuals admits the representation as in (2.32), that is

$$\hat{u}_t = \hat{\eta}_{0,t} - \hat{\boldsymbol{\beta}}'_{k,n}\hat{\mathbf{m}}_{k,t} = \sqrt{n}(1, -\hat{\boldsymbol{\Theta}}'_{k,n})\begin{pmatrix}\hat{\eta}_{0,nt}\\ \hat{\mathbf{m}}_{k,nt}\end{pmatrix} \tag{4.5}$$

so that $\hat{u}_t = O_p(\sqrt{n})$, and hence $n^{-3/2}\sum_{t=1}^{[nr]}\hat{u}_t = \hat{\boldsymbol{\kappa}}'_n n^{-1}\sum_{t=1}^{[nr]}\hat{\mathbf{m}}_{nt} = O_p(1)$, with $\hat{\boldsymbol{\kappa}}_n = (1, -\hat{\boldsymbol{\Theta}}'_{k,n})'$ and $\hat{\mathbf{m}}_{nt} = (\hat{\eta}_{0,nt}, \hat{\mathbf{m}}'_{k,nt})'$. Next result establish the rates of divergence of the different elements composing (3.9) under no cointegration and hence the divergence rate of the proposed test statistic under the alternative, that is, its consistency property.

Proposition 4.1. *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 with $\alpha = 1$, we have that:

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

(b) $\hat{\omega}_{v,n}^2(q_n) = O_p(q_n n^2)$, *and* $\hat{\boldsymbol{\omega}}_{kv,n}(q_n) = O_p(q_n\sqrt{n})$

so that $\hat{\omega}_{v.k,n}^2(q_n) = O_p(q_n n^2)$, *and*

(c) $\hat{CS}_n = O_p(\sqrt{n/q_n})$

Proof. First, given the above representation for the sequence of OLS residuals under no cointegration, we have that

$$n^{-1/2}\sum_{t=1}^{[nr]}\hat{v}_t = n^{3/2}(1, -\hat{\boldsymbol{\Theta}}'_{k,n})\left\{n^{-1}\sum_{t=1}^{[nr]}\hat{\mathbf{m}}_{nt}\hat{\mathbf{m}}'_{nt} - \frac{[nr]}{n} n^{-1}\sum_{t=1}^{n}\hat{\mathbf{m}}_{nt}\hat{\mathbf{m}}'_{nt}\right\}\begin{pmatrix}1\\ -\hat{\boldsymbol{\Theta}}_{k,n}\end{pmatrix}$$

that clearly yields the result in (a). Second, given that $\hat{u}_t^2 = O_p(n)$, $t = 1, ..., n$, and $\hat{\sigma}_n^2 = O_p(n)$, then $n^{-1}\sum_{t=1}^{n}\hat{v}_t^2 = O_p(n^2)$ and thus $\hat{\omega}_{v,n}^2(q_n) = O_p(n^2 q_n)$ in (3.10) following similar arguments as in Phillips (1991). Also, for the kernel estimate of the long-run variance $\boldsymbol{\omega}_{kv}$, given by

$$\hat{\boldsymbol{\omega}}_{kv,n}(q_n) = n^{-1}\sum_{t=1}^{n}\hat{\mathbf{Z}}_{k,t}\hat{v}_t + \sum_{h=1}^{n-1} w(hq_n^{-1})\left\{n^{-1}\sum_{t=h+1}^{n}(\hat{\mathbf{Z}}_{k,t-h}\hat{v}_t + \hat{\mathbf{Z}}_{k,t}\hat{v}_{t-h})\right\}$$

we have that

$$n^{-1}\sum_{t=1}^{n}\hat{\mathbf{Z}}_{k,t}\hat{v}_t = n^{-1}\sum_{t=1}^{n}\hat{\mathbf{Z}}_{k,t}\hat{u}_t^2 - \hat{\sigma}_n^2 n^{-1}\sum_{t=1}^{n}\hat{\mathbf{Z}}_{k,t}$$

$$= \sum_{t=1}^{n}\hat{\mathbf{Z}}_{k,t}\mathrm{vec}(\hat{\mathbf{m}}_{nt}\hat{\mathbf{m}}_{nt}')'(\hat{\boldsymbol{\kappa}}_n \otimes \hat{\boldsymbol{\kappa}}_n) - \sqrt{n}(n^{-1}\hat{\sigma}_n^2)\, n^{-1/2}\sum_{t=1}^{n}\hat{\mathbf{Z}}_{k,t}$$

$$= \sqrt{n}\left\{ n^{-1/2}\sum_{t=1}^{n}\hat{\mathbf{Z}}_{k,t}\mathrm{vec}(\hat{\mathbf{m}}_{nt}\hat{\mathbf{m}}_{nt}')'(\hat{\boldsymbol{\kappa}}_n \otimes \hat{\boldsymbol{\kappa}}_n) - (n^{-1}\hat{\sigma}_n^2)\, n^{-1/2}\sum_{t=1}^{n}\hat{\mathbf{Z}}_{k,t} \right\}$$

where we use the representation $\hat{u}_t^2 = n\hat{\boldsymbol{\kappa}}_n'\hat{\mathbf{m}}_{nt}\hat{\mathbf{m}}_{nt}'\hat{\boldsymbol{\kappa}}_n = n(\hat{\boldsymbol{\kappa}}_n' \otimes \hat{\boldsymbol{\kappa}}_n')\mathrm{vec}(\hat{\mathbf{m}}_{nt}\hat{\mathbf{m}}_{nt}')$ for the squared OLS residuals, with $\sum_{t=1}^{n}\hat{\mathbf{Z}}_{k,t} = \mathbf{0}_k$, for any $m \geq 1$ and $n^{-1/2}\sum_{t=1}^{n}\hat{\mathbf{Z}}_{k,t}$ $= n^{-1/2}\sum_{t=1}^{n}\boldsymbol{\varepsilon}_{k,t} = O_p(1)$ when the integrated regressors do not contain any deterministic component, so that $n^{-1}\sum_{t=1}^{n}\hat{\mathbf{Z}}_{k,t}\hat{v}_t = O_p(\sqrt{n})$. Similarly, for lag h = 1, ..., n–1, we can write

$$\sum_{t=h+1}^{n}\hat{\mathbf{Z}}_{k,t-h}\hat{v}_t = \sum_{t=h+1}^{n}\hat{\mathbf{Z}}_{k,t-h}\hat{u}_t^2 - \hat{\sigma}_n^2\left(\sum_{t=1}^{n}\hat{\mathbf{Z}}_{k,t} - \sum_{t=n-h+1}^{n}\hat{\mathbf{Z}}_{k,t}\right)$$

where $n^{-1/2}\sum_{t=n-h+1}^{n}\hat{\mathbf{Z}}_{k,t} = O_p(n^{-1/2}h)$, yielding $n^{-1}\sum_{t=1}^{n}\hat{\mathbf{Z}}_{k,t-h}\hat{v}_t = n^{-1}\sum_{t=1}^{n}\hat{\mathbf{Z}}_{k,t}\hat{v}_{t-h}$ $= O_p(\sqrt{n})$, and hence $\hat{\boldsymbol{\omega}}_{kv,n}(q_n) = O_p(q_n\sqrt{n})$ and $\hat{\boldsymbol{\gamma}}_{kv,n}(q_n) = \hat{\boldsymbol{\Omega}}_{kk,n}^{-1}\hat{\boldsymbol{\omega}}_{kv,n} = O_p(q_n\sqrt{n})$, so that $\hat{\omega}_{v.k,n}^2(q_n) = O_p(q_n n^2)$, and the leading term in the numerator of the test statistic (3.9) is given by

$$n^{-1/2}\hat{V}_t - \hat{\boldsymbol{\gamma}}_{kv,n}'(q_n)[n^{-1/2}(\mathbf{X}_{k,t} - (t/n)\mathbf{X}_{k,n})]$$

$$= n^{3/2}\left\{\frac{1}{n^2}\hat{V}_t - \frac{1}{n^{3/2}}\hat{\boldsymbol{\gamma}}_{kv,n}'(q_n)[n^{-1/2}(\mathbf{X}_{k,t} - (t/n)\mathbf{X}_{k,n})]\right\}$$

$$= n^{3/2}\left\{\frac{1}{n^2}\hat{V}_t + O_p(q_n/n)\right\}$$

which implies the final result in (c). One final comment about the divergence rate displayed by our testing procedure in part (c) of Proposition 4.1, which is the same as for the CUSUM-type tests proposed by Xiao (1999), Xiao and Phillips (2002) and Wu and Xiao (2008), and thus its power performance is comparable with that of these alternative testing procedures for the same null hypothesis of cointegration.

5. Finite sample alternative distribution, size and power

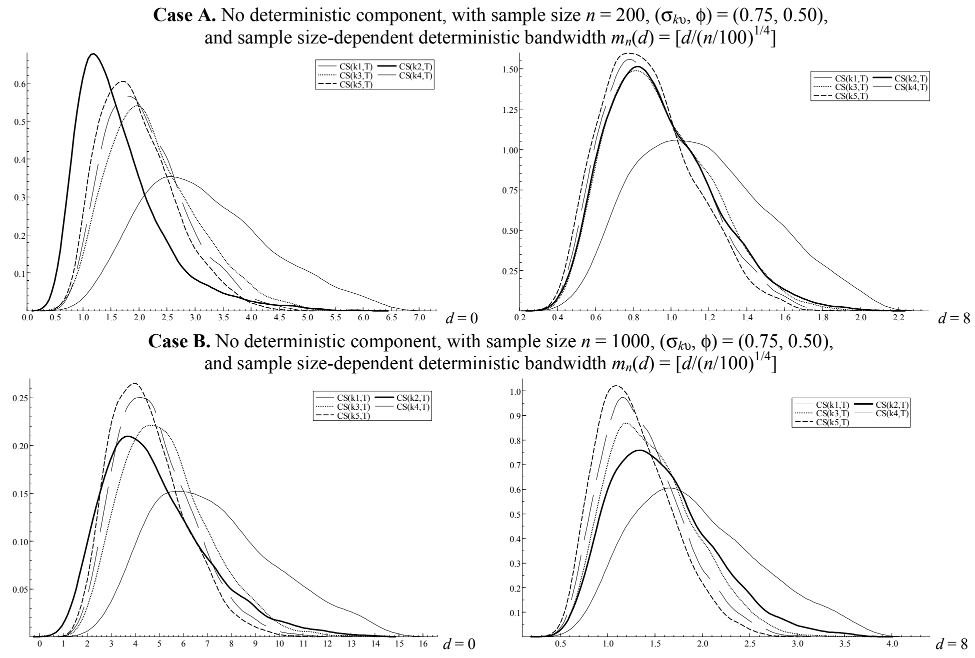
The DGP used for the simulation experiment is based on $u_t = \alpha u_{t-1} + \upsilon_t$ and $\boldsymbol{\eta}_{k,t} = \boldsymbol{\eta}_{k,t-1} + \boldsymbol{\varepsilon}_{k,t}$, $\boldsymbol{\varepsilon}_{k,t} = \phi\boldsymbol{\varepsilon}_{k,t-1} + \mathbf{e}_{k,t}$, where $\boldsymbol{\xi}_{0,t}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

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

With these error terms we compute the OLS residuals from (2.4) 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, except for calculating the quantiles of the null distribution with 20000 independent draws. Tables B.1 and B.2 in Appendix B present these quantiles for different samples sizes and for the cases of no deterministic term in the cointegrating regression, and for the inclusion of only a constant term or a constant term and a linear trend component. From these results it is remarkable the invariance of the null distribution of the test statistic to the structure and dimension of the estimated model even for very small sample sizes. Appendix C presents the results for the finite sample-adjusted empirical size of our testing procedure under quite general assumptions on the serial correlation of both the regression error term and the errors driving the endogenous stochastic trend components (endogenous integrated regressors), providing strong support for the theoretical robustness found in section 4 even for very small values of the bandwidth parameter in the computation of the kernel estimates of the long-run variances and covariances needed in the definition of the statistic in (3.9).

Finally, the results on power performance are presented graphically. First, figure 3 displays the 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, the number of integrated regressors.

Figure 3. *Nonparametric kernel estimation of the density function of the CUSUM-of squares test statistics computed under the alternative of no cointegration*



Additionally, figures 4 and 5 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 4. *Power profile as a function of the number of integrated regressors (k), the magnitude of the deterministic bandwidth parameter* $q_n(d) = [d(n/100)^{1/4}]$, *and the degree of serial correlation in the errors driving the stochastic trend components* (ϕ)

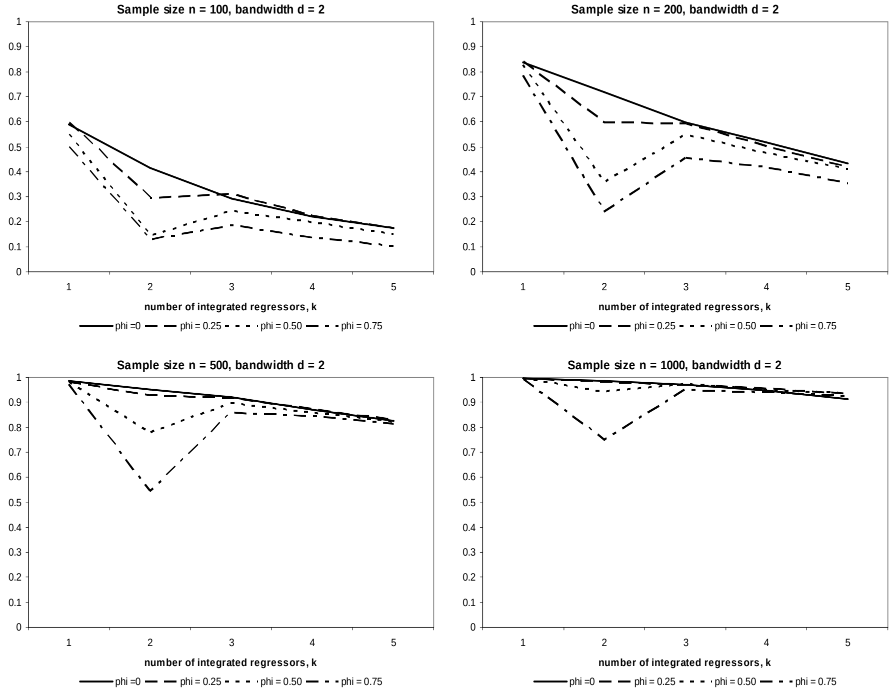
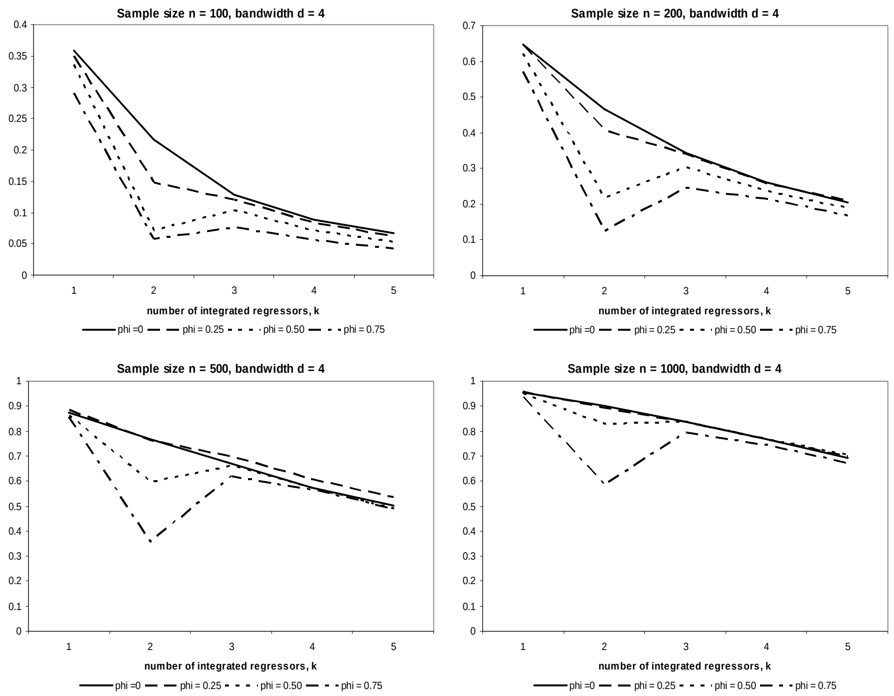


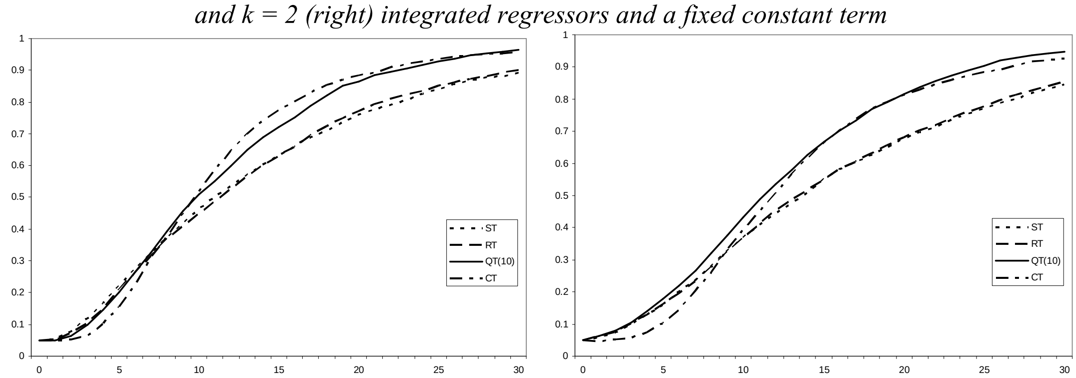
Figure 5. *Power profile as a function of the number of integrated regressors (k), the magnitude of the deterministic bandwidth parameter* $q_n(d) = [d(n/100)^{1/4}]$, *and the degree of serial correlation in the errors driving the stochastic trend components* (ϕ)



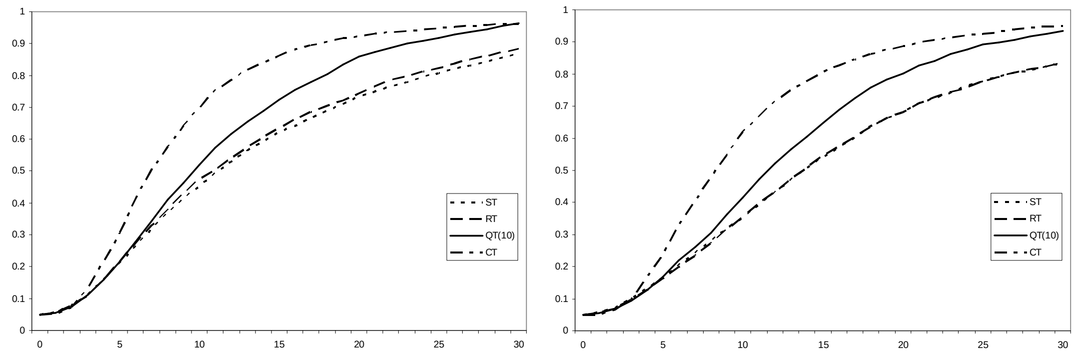
Power results displayed in figures 4, 5 correspond to the case of no deterministic component in the estimated cointegrating regression. For the case of inclusion of a constant term, or a constant term and a linear trend, the profiles are quite similar, but with a slight loss of power, which is a feature 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 3, 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.

Finally, we consider the following generalized structure for describing the local non-stationary behavior of the regression error terms $\Delta u_t = (1-\rho_n L)\upsilon_t$, where $\upsilon_t = \phi_0 \upsilon_{t-1} + e_{0,t}$, and $\rho_n = 1 - \lambda n^{-\gamma}$, to compare the power performance of our test statistic with the tests procedures proposed by Shin (1994), Xiao and Phillips (2002) and Jansson (2005). For these tests the index γ takes the unit value, $\gamma = 1$, while for our test we choose $\gamma = 3/4$, the only way to make these tests comparable under this construction. Figure 6 below represents the power profiles of these four tests for values of λ ranging from 0 to 30 in the case of including a constant term in the specification of the cointegrating regression, *k* = 1 or 2 integrated regressors, and values of $\phi_0 = 0, 0.5, 0.75$.

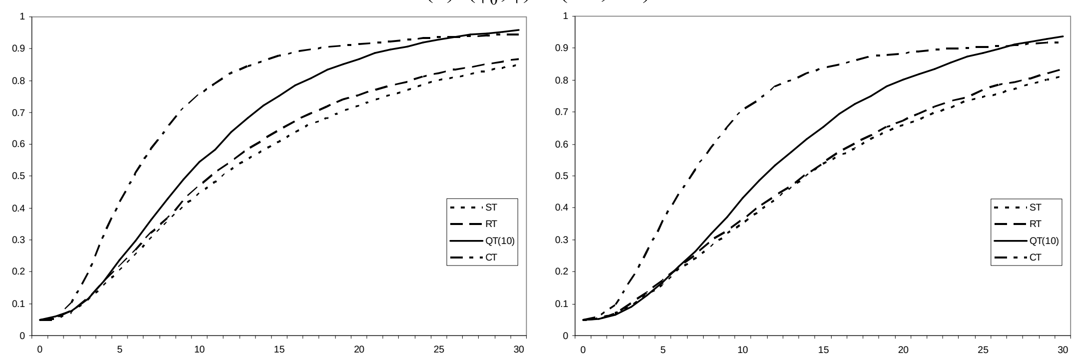
Figure 6. *Rejection rates under the local-to-unity MA root for the Shin (ST), Xiao and Phillips (RT), Jansson (QT(10)) and CUSUM of squares (CT) test statistics and a sample size n = 1000, with k = 1 (left) and k = 2 (right) integrated regressors and a fixed constant term*



(a) $(\phi_0, \phi) = (0, 0)$



(b) $(\phi_0, \phi) = (0.5, 0.5)$



(c) $(\phi_0, \phi) = (0.75, 0.25)$

The numerical results displayed in these figures clearly show the superiority of our testing procedures in terms of power, particularly for highly autocorrelated errors.

6. An application to the US aggregate consumption function and some concluding remarks

To illustrate the performance of the proposed test statistic as compared some other alternative testing procedures in the same framework of analysis, we use it to test for a stable long-run relationship between consumption and income using US quarterly data from 1947Q1 to 1991Q2 (n = 178), with the series defined as RPC, Real total personal consumption expenditure, RCNDS, Real consumption on non-durables and services, and RDPI, Real total disposable personal income in 2009 dollars. All these series were transformed into logs and are the same series used in Shin (1994) and McCabe et.al. (1997), except that in that cases are expressed in 1982 dollars.

Table 2 bellow shows the results of our test statistic including, for purposes of comparison, a simplified version to be used in the case of strictly exogenous regressors and given by

$$C\hat{S}_n^0(q_n) = n^{-1/2}\hat{\omega}_{v,n}^{-1}(q_n)\max_{t=1,\ldots,n} |\textstyle\sum_{j=1}^{t} \hat{v}_{j,p}(k)|$$

without the correction for endogeneity in (3.9), both in the numerator and denominator of the test statistic.

We also include the results of Shin's (1994)[11] and McCabe et.al. (MLS) (1997) tests for the null hypothesis of cointegration, where the MLS test is also based on a measure of excessive fluctuation under cointegration in the sequence of regression errors u_t similar to that of Shin's test, but with a parametric correction both for endogeneity and serially correlated regression error terms.

The estimated values of our test statistic and the Shin's test, $\hat{CI}_n(q)$, are obtained for certain different values of the bandwidth parameter to correct for serial correlation, while that the MLS test is also computed for different values of p, the order of the AR(p) model adjusted to the regression errors to obtain parametrically corrected error terms free of remaining serial dependence. From these results, we observe that our testing procedure clearly indicates evidence against the existence of a stable long-run relationship between consumption and income, with a little or small effect of the bandwidth choice. Shin establishes that each individual series is I(1), possibly with drift, and thus the versions of the tests in panel B of Table 2 (with inclusion of a constant term and a linear trend in the estimated regression model) seems more appropriate. Overall, there is a substantial amount of agreement between the outcomes of these three tests against the existence of a cointegration relationship between these variables.

[11] Formally, the general version of Shin's (1994) test is based on the residuals from the Dynamic OLS (DOLS) estimation of the cointegrating regression proposed by Saikkonen (1991), among others. However, it is not difficult to check that it is also valid when using any other existing asymptotically efficient testing estimator under endogenous regressors as the FM-OLS method by Phillips and Hansen.

Table 2. Application to US Aggregate consumption data, 1947Q1-1991Q2 (n = 178)

A. Constant term					$\hat{S}_n(p)$		
	Bandwidth	$C\hat{S}_n(q)$	$C\hat{S}_n^0(q)$	$\hat{C}I_n(q)$	$p=1$	2	3
RPC-RDPI	$q=1$	4.157 [a]	1.604 [a]	1.136 [a]	2.264 [a]	2.321 [a]	0.699 [a]
	2	3.496 [a]	1.395 [b]	0.790 [a]			
	5	2.695 [a]	1.093	0.393 [c]			
	10	2.278 [a]	0.906	0.250 [c]			
RCNDS-RDPI	$q=1$	3.613 [a]	1.536 [b]	1.518 [a]	2.963 [a]	3.028 [a]	3.194 [a]
	2	3.505 [a]	1.291 [c]	1.037 [a]			
	5	3.159 [a]	0.935	0.477 [b]			
	10	2.811 [a]	0.765	0.281 [c]			

B. Constant term and linear trend					$\hat{S}_n(p)$		
	Bandwidth	$C\hat{S}_n(q)$	$C\hat{S}_n^0(q)$	$\hat{C}I_n(q)$	$p=1$	2	3
RPC-RDPI	$q=1$	4.088 [a]	1.549 [b]	0.306 [a]	0.616 [a]	0.586 [a]	0.079
	2	3.131 [a]	1.364 [b]	0.217 [a]			
	5	2.489 [a]	1.156	0.188 [c]			
	10	2.397 [a]	1.077	0.083			
RCNDS-RDPI	$q=1$	4.356 [a]	1.377 [b]	0.374 [a]	0.507 [a]	0.739 [a]	0.123 [b]
	2	4.088 [a]	1.179 [c]	0.261 [a]			
	5	3.648 [a]	0.925	0.133 [b]			
	10	3.274 [a]	0.866	0.091			

Notes. [a, b, c]: Rejection at 1, 5, and 10%. $\hat{C}I_n(q)$ is the Shin's (1994) test based on FM-OLS estimates and residuals, while $\hat{S}_n(p)$ is the McCabe, Leybourne and Shin (MLS) (1997) test.

References

Berenguer-Rico, V., J. Gonzalo (2014). Summability of stochastic processes. A generalization of integration for non-linear processes. Journal of Econometrics, 178(2), 331-341.

Carrion-i-Silvestre, J.L., A. Sansó (2006). A guide to the computation of stationarity tests. Empirical Economics, 31(2), 433-448.

Choi, I., B.C. Ahn (1995). Testing for cointegration in a system of equations. Econometric Theory, 11(5), 952-983.

Deng, A., P. Perron (2008). The limit distribution of the CUSUM of squares test under general mixing conditions. Econometric Theory, 24(3), 809-822.

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

Giraitis, L., P.C.B. Phillips (2006). Uniform limit theory for stationary autoregression. Journal of Time Series Analysis, 27(1), 51-60.

Gonzalo, J. (1994). Five alternative methods of estimating long-run equilibrium relationships. Journal of Econometrics, 60(1-2), 203-233.

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. (1990). A powerful, simple test for cointegration using Cochrane-Orcutt. Working paper n.230, Rochester Center for Economic Research (RCER), University of Rochester.

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.

Hansen, B.E. (1992c). Heteroskedastic cointegration. Journal of Econometrics, 54(1-3), 139-158.

Harris, D., B. Inder (1994). A Test of the Null Hypothesis of Cointegration. In Nonstationary Time Series Analysis and Cointegration, edited by C. P. Hargreaves.Oxford Univeristy Press, 133-152. New York.

Harris, D., B. McCabe, S. Leybourne (2002). Stochastic cointegration: estimation and inference. Journal of Econometrics, 111(2), 363-384.

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.
Jansson, M. (2005). Tests of the null hypothesis of cointegration based on efficient tests for a unit MA root. In Andrews, D.W.K. and J.H. Stock (Eds): Identification and Inference for Econometric Models: Essays in Honor of T.J. Rothenberg, Chapter 15, 357-374. Cambridge University Press, Cambridge.
Jansson, M. (2005). Tests of the null hypothesis of cointegration based on efficient tests for a unit MA root. In D.W.K Andrews and J.H. Stock (Eds.): Identification and Inference in Econometric Models. Essays in Honour of T.J. Rothenberg, Chapter 15, 357-374. Cambridge University Press, Cambridge.
Kurozumi, E., K. Hayakawa (2009). Asymptotic properties of the efficient estimators for cointegrating regression models with serially dependent errors. Journal of Econometrics, 149(2), 118-135.
Kwiatkowski, D., P.C.B. Phillips, P. Schmidt, Y. Shin (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?. Journal of Econometrics, 54(1-3), 159-178.
Lee, S., O. Na, S. Na (2003). On the CUSUM of squares test for variance change in nonstationary and nonparametric time series models. Annals of the Institute of Statistical Mathematics, 55(3), 467-485.
Leybourne, S. J., B.P.M. McCabe (1994). A Simple Test for Cointegration. Oxford Bulletin of Economics and Statistics, 56(1), 97-103.
Leybourne, S.J., B.P.M. McCabe, A.R. Tremayne (1996). Can economic time series be differenced to stationarity?. Journal of Business and Economic Statistics, 14(4), 435-446.
Lu, X., K. Maekawa, S. Lee (2008). The CUSUM of squares test for the stability of regression models with non-stationary regressors. Economics Letters, 100(2), 234-237.
Maekawa, K. T. Yamamoto, Y. Takeuchi, M. Hatanaka (1996). Estimation in dynamic regression with an integrated process. Journal of Statistical Planning and Inference, 49(2), 279-393.
McCabe, B.P.M., S.J. Leybourne, Y. Shin (1997). A parametric approach to testing the null of cointegration. Journal of Time Series Analysis, 18(4), 395-413.
Montalvo, J.G. (1995). Comparing cointegrating regression estimators: Some additional Monte Carlo results. Economics Letters, 48(), 229-234.
Müller, U.K. (2005). Size and power of tests of stationarity in highly autocorrelated time series. Journal of Econometrics, 128(2), 195-213.
Nabeya, S., P. Perron (1994). Local asymptotic distribution related to the AR(1) model with dependent errors. Journal of Econometrics, 62(2), 229-264.
Nielsen, B., J.S. Sohkanen (2011). Asymptotic behavior of the CUSUM of squares test under stochastic and deterministic time trends. Econometric Theory, 27(4), 913-927.
Palma, W., M. Zevallos (2004). Analysis of the correlation structure of square time series. Journal of Time Series Analysis, 25(4), 529-550.
Park, J.Y. (1992). Canonical cointegrating regressions. Econometrica, 60(1), 119-143.
Park, J.Y., P.C.B. Phillips (1989). Statistical inference in regressions with integrated processes: Part 2. Econometric Theory, 5(1), 95-131.
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. (1987a). Time series regression with a unit root. Econometrica, 55(2), 277-301.
Phillips, P.C.B. (1987b). Towards a unified asymptotic theory for autoregression. Biometrika, 74(3), 535-547.
Phillips, P.C.B. (1989). Partially identified econometric models. Econometric Theory, 5(2), 181-240.
Phillips, P.C.B. (1991). Spectral regression for cointegrated time series. In: Barnett, W., J. Powell, G. Tauchen (Eds.), Nonparametric and Semiparametric Methods in Econometrics and Statistics. Cambridge University Press, Cambridge.
Phillips, P.C.B. (1995). Fully modified least squares and vector autoregression. Econometrica, 63(5), 1023-1078.
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., M. Loretan (1991). Estimating long-run economic equilibria. The Review of Economic Studies, 58(3), 407-436.
Phillips, P.C.B., S. Ouliaris (1990). Asymptotic properties of residual based tests for cointegration. Econometrica, 58(1), 165-193.

Phillips, P.C.B., T. Magdalinos (2007a). Limit theory for moderate deviations from a unit root. Journal of Econometrics, 136(1), 115-130.
Phillips, P.C.B., T. Magdalinos (2007b). Limit theory for moderate deviations from a unit root under weak dependence. In: Phillips, G.D.A., E. Tzavalis (Eds.), The Refinement of Econometric Estimation and Test Procedures: Finite Sample and Asymptotic Analysis, Chapter 5, 123-162. Cambridge University Press, Cambridge
Saikkonen, P. (1991). Asymptotically efficient estimation of cointegration regressions. Econometric Theory, 7(1), 1-21.
Shin, Y. (1994). A residual-based test of the null of cointegration against the alternative of no cointegration. Econometric Theory, 10(1), 91-115.
Stock, J.H., M.W. Watson (1993). A simple estimator of cointegrating vectors in higher order integrated systems. Econometrica, 61(4), 783-820.
Vogelsang, T.J., M. Wagner (2014). Integrated modified OLS estimation and fixed-b inference for cointegrating regressions. Journal of Econometrics, 178(2), 741-760.
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.

Appendix A. *Sources of inconsistent estimation of a subset of coefficients of the cointegration vector: Subcointegration and stationary stochastic regressors*

Let us consider the cointegrating regression model in (2.4) written as $\hat{Y}_t = \boldsymbol{\beta}_k' \hat{\mathbf{X}}_{k,t} + u_{t,m}$, where $\hat{Y}_t = \hat{\eta}_{0,t}$, $\hat{\mathbf{X}}_{k,t} = \hat{\mathbf{m}}_{k,t}$, and $u_{t,m}$ are the OLS detrended observations of $\eta_{0,t}$, $\mathbf{m}_{k,t} = \boldsymbol{\eta}_{k,t} + \mathbf{A}_{k,q}\boldsymbol{\tau}_{q,t}$, and u_t, respectively, obtained by correcting for the trend polynomial $\boldsymbol{\tau}_{m,t}$, with $\hat{\mathbf{m}}_{k,t} = \boldsymbol{\Gamma}_{kk,n}^{-1}\hat{\mathbf{m}}_{k,nt}$ and scaling matrix $\boldsymbol{\Gamma}_{kk,n} = \mathbf{W}_{kk,n}^{-1}\mathbf{C}_{kk}'$. In the case $\mathbf{A}_{k,q} = \mathbf{0}_{k,q}$, where $\boldsymbol{\Gamma}_{kk,n} = n^{-1/2}\mathbf{I}_{k,k}$, we have $\hat{\mathbf{m}}_{k,nt} = \hat{\boldsymbol{\eta}}_{k,nt} = \boldsymbol{\eta}_{k,nt} - \mathbf{Q}_{km,n}\mathbf{Q}_{mm,n}^{-1}\boldsymbol{\tau}_{m,nt}$, while that in the general case $\mathbf{A}_{k,q} \neq \mathbf{0}_{k,q}$ we have

$$\hat{\mathbf{m}}_{k,nt} = \begin{pmatrix} \mathbf{d}_{q,nt} + n^{1/2}\boldsymbol{\Gamma}_{q,n}\mathbf{C}_{k,q}'\hat{\boldsymbol{\eta}}_{k,nt} \\ \mathbf{C}_{k,k-q}'\hat{\boldsymbol{\eta}}_{k,nt} \end{pmatrix}$$

with $\mathbf{d}_{q,nt} = \boldsymbol{\tau}_{q,nt} - \mathbf{Q}_{qm,n}\mathbf{Q}_{mm,n}^{-1}\boldsymbol{\tau}_{m,nt}$. Next, we assume the situation where the scaled k-vector of stochastic trend components $\boldsymbol{\eta}_{k,nt}$ is given by

$$\boldsymbol{\eta}_{k,nt} = \begin{pmatrix} \boldsymbol{\eta}_{k_1,nt} \\ \boldsymbol{\eta}_{k_2,nt} \end{pmatrix} = \mathbf{B}_{k,k}\begin{pmatrix} \boldsymbol{\eta}_{k_1,nt} \\ n^{-1/2}\mathbf{u}_{k_2,t} \end{pmatrix}$$

with $\mathbf{B}_{k,k} = \mathbf{I}_{k,k}$ in the case of inclusion of k_2 stationary regressors, $\mathbf{u}_{k_2,t}$, and

$$\mathbf{B}_{k,k} = \begin{pmatrix} \mathbf{I}_{k_1,k_1} & \mathbf{0}_{k_1,k_2} \\ \mathbf{B}_{k_2,k_1} & \mathbf{I}_{k_2,k_2} \end{pmatrix}$$

in the case where there are $k_2 \geq 1$ integrated but cointegrated regressors (subcointegration), such that $\boldsymbol{\eta}_{k_2,t} = \mathbf{B}_{k_2k_1}\boldsymbol{\eta}_{k_1,t} + \mathbf{u}_{k_2,t}$. Thus, when $\mathbf{A}_{k,q} = \mathbf{0}_{k,q}$ the k-vector $\hat{\mathbf{m}}_{k,nt}$ can also be partitioned as

$$\hat{\mathbf{m}}_{k,nt} = \begin{pmatrix} \hat{\mathbf{m}}_{k_1,nt} \\ \hat{\mathbf{m}}_{k_2,nt} \end{pmatrix} = \begin{pmatrix} \hat{\boldsymbol{\eta}}_{k_1,nt} \\ \mathbf{B}_{k_2k_1}\hat{\boldsymbol{\eta}}_{k_1,nt} + n^{-1/2}\hat{\mathbf{u}}_{k_2,t} \end{pmatrix} = \mathbf{B}_{k,k}\begin{pmatrix} \hat{\boldsymbol{\eta}}_{k_1,nt} \\ n^{-1/2}\hat{\mathbf{u}}_{k_2,t} \end{pmatrix}$$

which gives

$$n^{-1}\sum_{t=1}^{n}\hat{\mathbf{m}}_{k,nt}\hat{\mathbf{m}}'_{k,nt} = \mathbf{B}_{k,k}\left\{n^{-1}\sum_{t=1}^{n}\begin{pmatrix}\hat{\boldsymbol{\eta}}_{k_1,nt}\\ n^{-1/2}\hat{\mathbf{u}}_{k_2,t}\end{pmatrix}(\hat{\boldsymbol{\eta}}'_{k_1,nt}, n^{-1/2}\hat{\mathbf{u}}'_{k_2,t})\right\}\mathbf{B}'_{k,k}$$

$$= \mathbf{B}_{k,k}\begin{pmatrix}\mathbf{Q}_{k_1k_1,n} & n^{-1}\mathbf{J}_{k_1k_2,n}\\ n^{-1}\mathbf{J}'_{k_1k_2,n} & n^{-1}\boldsymbol{\Sigma}_{k_2k_2,n}\end{pmatrix}\mathbf{B}'_{k,k}$$

and

$$n^{-1}\sum_{t=1}^{n}\hat{\mathbf{m}}_{k,nt}u_{t,m} = \mathbf{B}_{k,k}\begin{pmatrix}n^{-1}\sum_{t=1}^{n}\hat{\boldsymbol{\eta}}_{k_1,nt}u_{t,m}\\ n^{-1/2}n^{-1}\sum_{t=1}^{n}\hat{\mathbf{u}}_{k_2,t}u_{t,m}\end{pmatrix} = n^{-\kappa}\mathbf{B}_{k,k}\begin{pmatrix}n^{-(1-\kappa)}\sum_{t=1}^{n}\hat{\boldsymbol{\eta}}_{k_1,nt}u_{t,m}\\ n^{-(1/2-\kappa)}n^{-1}\sum_{t=1}^{n}\hat{\mathbf{u}}_{k_2,t}u_{t,m}\end{pmatrix}$$

$$= n^{-\kappa}\mathbf{B}_{k,k}\begin{pmatrix}\mathbf{J}_{k_1u,n}\\ n^{-(1/2-\kappa)}\mathbf{J}_{k_2u,n}\end{pmatrix}$$

with $\mathbf{Q}_{k_1k_1,n} = n^{-1}\sum_{t=1}^{n}\hat{\boldsymbol{\eta}}_{k_1,nt}\hat{\boldsymbol{\eta}}'_{k_1,nt}$, $\mathbf{J}_{k_1k_2,n} = n^{-1/2}\sum_{t=1}^{n}\hat{\boldsymbol{\eta}}_{k_1,nt}\hat{\mathbf{u}}'_{k_2,t}$, and $\boldsymbol{\Sigma}_{k_2k_2,n} = n^{-1}\sum_{t=1}^{n}\hat{\mathbf{u}}_{k_2,t}\hat{\mathbf{u}}'_{k_2,t}$ the sample covariance matrix of the stationary sequence $\hat{\mathbf{u}}_{k_2,t}$, such that $\boldsymbol{\Sigma}_{k_2k_2,n} \to^{p} \boldsymbol{\Sigma}_{k_2k_2} = E[\mathbf{u}_{k_2,t}\mathbf{u}'_{k_2,t}]$, the covariance matrix of the stationary sequence $\mathbf{u}_{k_2,t}$, by the weak law of large numbers under quite general conditions. Then, given that the scaled and normalized OLS estimation error of $\boldsymbol{\beta}_k = (\boldsymbol{\beta}'_{k_1}, \boldsymbol{\beta}'_{k_2})'$ can be represented as

$$n^{1/2+\kappa}\mathbf{B}'_{k,k}\begin{pmatrix}\hat{\boldsymbol{\beta}}_{k_1,n}-\boldsymbol{\beta}_{k_1}\\ \hat{\boldsymbol{\beta}}_{k_2,n}-\boldsymbol{\beta}_{k_2}\end{pmatrix} = n^{1/2+\kappa}\begin{pmatrix}\hat{\boldsymbol{\beta}}_{k_1,n}-\boldsymbol{\beta}_{k_1}+\mathbf{B}'_{k_2,k_1}(\hat{\boldsymbol{\beta}}_{k_2,n}-\boldsymbol{\beta}_{k_2})\\ \hat{\boldsymbol{\beta}}_{k_2,n}-\boldsymbol{\beta}_{k_2}\end{pmatrix}$$

$$= \begin{pmatrix}\mathbf{M}^{-1}_{k_1k_1,n}(\mathbf{J}_{k_1u,n}-\mathbf{J}_{k_1k_2,n}\boldsymbol{\Sigma}^{-1}_{k_2k_2,n}n^{-(1/2-\kappa)}\mathbf{J}_{k_2u,n})\\ \mathbf{M}^{-1}_{k_2k_2,n}(n^{-(1/2-\kappa)}\mathbf{J}_{k_2u,n}-n^{-1}\mathbf{J}'_{k_1k_2,n}\mathbf{Q}^{-1}_{k_1k_1,n}\mathbf{J}_{k_1u,n})\end{pmatrix}$$

with $\mathbf{M}_{k_1k_1,n} = \mathbf{Q}_{k_1k_1,n} - n^{-1}\mathbf{J}_{k_1k_2,n}\boldsymbol{\Sigma}^{-1}_{k_2k_2,n}\mathbf{J}'_{k_1k_2,n}$,and $\mathbf{M}_{k_2k_2,n} = n^{-1}[\boldsymbol{\Sigma}_{k_2k_2,n} - n^{-1}\mathbf{J}'_{k_1k_2,n}\mathbf{Q}^{-1}_{k_1k_1,n}\mathbf{J}_{k_1k_2,n}]$, we get

$$\hat{\boldsymbol{\beta}}_{k_2,n}-\boldsymbol{\beta}_{k_2} = [\boldsymbol{\Sigma}_{k_2k_2,n} - n^{-1}\mathbf{J}'_{k_1k_2,n}\mathbf{Q}^{-1}_{k_1k_1,n}\mathbf{J}_{k_1k_2,n}]^{-1}(\mathbf{J}_{k_2u,n} - n^{-(1/2+\kappa)}\mathbf{J}'_{k_1k_2,n}\mathbf{Q}^{-1}_{k_1k_1,n}\mathbf{J}_{k_1u,n})$$

and

$$\hat{\boldsymbol{\Theta}}_{k_1,n} = n^{1/2+\kappa}[\hat{\boldsymbol{\beta}}_{k_1,n}-\boldsymbol{\beta}_{k_1}+\mathbf{B}'_{k_2,k_1}(\hat{\boldsymbol{\beta}}_{k_2,n}-\boldsymbol{\beta}_{k_2})] = \mathbf{M}^{-1}_{k_1k_1,n}(\mathbf{J}_{k_1u,n}-\mathbf{J}_{k_1k_2,n}\boldsymbol{\Sigma}^{-1}_{k_2k_2,n}n^{-(1/2-\kappa)}\mathbf{J}_{k_2u,n})$$

Under the assumption of cointegration we get $\hat{\boldsymbol{\beta}}_{k_2,n}-\boldsymbol{\beta}_{k_2} \to^{p} \boldsymbol{\Sigma}^{-1}_{k_2k_2}\boldsymbol{\sigma}_{k_2u}$, with $\boldsymbol{\sigma}_{k_2u} = E[\mathbf{u}_{k_2,t}u_t]$, while that $\hat{\boldsymbol{\Theta}}_{k_1,n}$ weakly converges to a very different limiting distribution as compared to the case of inclusion of only k_1 integrated regressors such as

$$\hat{\boldsymbol{\Theta}}_{k_1,n} \Rightarrow \left(\int_0^1\mathbf{B}_{k_1,m}(s)\mathbf{B}'_{k_1,m}(s)ds\right)^{-1}$$
$$\times\left\{\int_0^1\mathbf{B}_{k_1,m}(s)dB_u(s)+\boldsymbol{\Delta}_{k_1u}-\left(\int_0^1\mathbf{B}_{k_1,m}(s)d\mathbf{B}'_{k_2}(s)+\boldsymbol{\Delta}_{k_1k_2}\right)\boldsymbol{\Sigma}^{-1}_{k_2k_2}\boldsymbol{\sigma}_{k_2u}\right\}$$

with $\boldsymbol{\Delta}_{k_1k_2} = \sum_{j=0}^{\infty}E[\boldsymbol{\varepsilon}_{k_1,t}\mathbf{u}'_{k_2,t}]$, unless $\boldsymbol{\sigma}_{k_2u} = \mathbf{0}_{k_2}$. These results imply that the sequence of OLS residuals are now given by

$$\hat{u}_t = u_{t,m} - \hat{\mathbf{X}}'_{k,t}(\hat{\boldsymbol{\theta}}_n-\boldsymbol{\theta}) = u_{t,m} - \hat{\mathbf{m}}'_{k,t}(\hat{\boldsymbol{\theta}}_n-\boldsymbol{\theta}) = u_{t,m} - n^{-\kappa}\hat{\mathbf{m}}'_{k,nt}n^{1/2+\kappa}(\hat{\boldsymbol{\theta}}_n-\boldsymbol{\theta})$$

$$= u_{t,m} - (n^{-\kappa}\hat{\boldsymbol{\eta}}'_{k_1,nt}, n^{-(1/2+\kappa)}\hat{\mathbf{u}}'_{k_2,t})n^{1/2+\kappa}\mathbf{B}'_{k,k}\begin{pmatrix}\hat{\boldsymbol{\beta}}_{k_1,n}-\boldsymbol{\beta}_{k_1}\\ \hat{\boldsymbol{\beta}}_{k_2,n}-\boldsymbol{\beta}_{k_2}\end{pmatrix}$$

$$= u_{t,m} - \hat{\mathbf{u}}'_{k_2,t}(\hat{\boldsymbol{\beta}}_{k_2,n}-\boldsymbol{\beta}_{k_2}) - n^{-\kappa}\hat{\boldsymbol{\eta}}'_{k_1,nt}\hat{\boldsymbol{\Theta}}_{k_1,n}$$

where, under cointegration, only the last term is asymptotically negligible, that is $\hat{u}_t = u_{t,m} - \hat{\mathbf{u}}'_{k_2,t}(\hat{\boldsymbol{\beta}}_{k_2,n} - \boldsymbol{\beta}_{k_2}) + O_p(n^{-1/2})$, determining very different limiting distributional results for the scaled partial sum process $n^{-1/2}\hat{U}_t$ and for the rest of elements composing the fluctuation-type statistics such as the kernel-based estimator of the long-run variance, $\hat{\omega}^2_{u,n}(q_n)$

Appendix B.

Table B.1. *Quantiles of the null distribution under cointegration of the modified CUSUM of squares test statistic,* $\hat{CS}_n$. *Number of integrated regressors, k.*

		Case of no deterministic component					Case of inclusion of a constant term (m = 1)				
Sample size, n		k = 1	2	3	4	5	k = 1	2	3	4	5
n = 100	0.01	0.3962	0.4003	0.4001	0.3983	0.3910	0.3994	0.3996	0.3973	0.3987	0.3974
	0.025	0.4334	0.4371	0.4322	0.4325	0.4340	0.4360	0.4353	0.4354	0.4370	0.4367
	0.05	0.4676	0.4713	0.4710	0.4728	0.4728	0.4731	0.4730	0.4741	0.4740	0.4751
	0.10	0.5214	0.5214	0.5227	0.5250	0.5273	0.5222	0.5244	0.5251	0.5251	0.5257
	0.25	0.6236	0.6241	0.6267	0.6303	0.6311	0.6254	0.6271	0.6284	0.6294	0.6310
	0.50	0.7757	0.7794	0.7779	0.7812	0.7827	0.7735	0.7745	0.7767	0.7779	0.7807
	0.75	0.9621	0.9643	0.9647	0.9723	0.9742	0.9631	0.9648	0.9633	0.9683	0.9715
	0.90	1.1524	1.1554	1.1601	1.1639	1.1669	1.1579	1.1583	1.1605	1.1633	1.1667
	0.95	1.2856	1.2882	1.2848	1.2883	1.2870	1.2714	1.2783	1.2783	1.2835	1.2869
	0.975	1.3803	1.3991	1.3909	1.3913	1.3970	1.3909	1.3953	1.4039	1.4017	1.4016
	0.99	1.5086	1.5092	1.5142	1.5171	1.5174	1.5239	1.5238	1.5241	1.5309	1.5253
n = 200	0.01	0.4102	0.4068	0.4074	0.4084	0.4100	0.4420	0.4009	0.4020	0.4024	0.4004
	0.025	0.4465	0.4474	0.4490	0.4495	0.4466	0.4420	0.4406	0.4430	0.4429	0.4419
	0.05	0.4886	0.4876	0.4864	0.4871	0.4860	0.4815	0.4808	0.4818	0.4815	0.4820
	0.10	0.5358	0.5353	0.5376	0.5364	0.5378	0.5331	0.5340	0.5336	0.5332	0.5353
	0.25	0.6390	0.6398	0.6396	0.6389	0.6411	0.6365	0.6369	0.6384	0.6402	0.6393
	0.50	0.7858	0.7873	0.7888	0.7883	0.7888	0.7864	0.7871	0.7874	0.7878	0.7888
	0.75	0.9731	0.9734	0.9738	0.9767	0.9775	0.9737	0.9733	0.9729	0.9738	0.9745
	0.90	1.1719	1.1738	1.1746	1.1759	1.1784	1.1711	1.1722	1.1766	1.1757	1.1764
	0.95	1.3029	1.3062	1.3072	1.3050	1.3073	1.2973	1.2996	1.3041	1.3035	1.3067
	0.975	1.4134	1.4112	1.4137	1.4168	1.4157	1.4183	1.4190	1.4224	1.4286	1.4317
	0.99	1.5523	1.5498	1.5527	1.5573	1.5568	1.5703	1.5739	1.5772	1.5808	1.5859
n = 500	0.01	0.4139	0.4146	0.4150	0.4147	0.4141	0.4167	0.4173	0.4172	0.4184	0.4191
	0.025	0.4515	0.4522	0.4536	0.4537	0.4543	0.4569	0.4565	0.4571	0.4590	0.4582
	0.05	0.4924	0.4922	0.4936	0.4938	0.4919	0.4956	0.4958	0.4952	0.4967	0.4972
	0.10	0.5454	0.5444	0.5437	0.5442	0.5447	0.5488	0.5481	0.5499	0.5470	0.5483
	0.25	0.6475	0.6479	0.6477	0.6481	0.6478	0.6497	0.6505	0.6508	0.6520	0.6520
	0.50	0.7980	0.7981	0.7977	0.7987	0.8003	0.7992	0.7993	0.8005	0.8011	0.8014
	0.75	0.9884	0.9900	0.9910	0.9899	0.9919	0.9898	0.9904	0.9899	0.9895	0.9887
	0.90	1.1896	1.1903	1.1910	1.1907	1.1894	1.1838	1.1845	1.1858	1.1876	1.1891
	0.95	1.3204	1.3193	1.3202	1.3225	1.3218	1.3118	1.3116	1.3104	1.3137	1.3131
	0.975	1.4353	1.4366	1.4371	1.4359	1.4368	1.4243	1.4254	1.4272	1.4264	1.4278
	0.99	1.5778	1.5811	1.5787	1.5912	1.5888	1.5749	1.5767	1.5779	1.5738	1.5788
n = 1000	0.01	0.4230	0.4238	0.4208	0.4225	0.4223	0.4229	0.4239	0.4235	0.4223	0.4235
	0.025	0.4600	0.4606	0.4597	0.4610	0.4602	0.4609	0.4609	0.4614	0.4620	0.4626
	0.05	0.4980	0.4973	0.4982	0.4987	0.4985	0.5012	0.5021	0.5021	0.5036	0.5035
	0.10	0.5494	0.5501	0.5494	0.5499	0.5489	0.5534	0.5533	0.5533	0.5520	0.5531
	0.25	0.6510	0.6516	0.6522	0.6523	0.6529	0.6551	0.6567	0.6570	0.6566	0.6570
	0.50	0.8038	0.8042	0.8037	0.8044	0.8042	0.8077	0.8078	0.8071	0.8079	0.8069
	0.75	0.9955	0.9950	0.9958	0.9961	0.9965	0.9952	0.9958	0.9962	0.9964	0.9966
	0.90	1.2007	1.2003	1.1981	1.2011	1.2012	1.2017	1.2033	1.2027	1.2030	1.2042
	0.95	1.3327	1.3328	1.3302	1.3306	1.3312	1.3364	1.3375	1.3385	1.3386	1.3382
	0.975	1.4501	1.4474	1.4462	1.4461	1.4459	1.4549	1.4556	1.4554	1.4575	1.4569
	0.99	1.6045	1.5999	1.5976	1.6040	1.6102	1.6009	1.6026	1.5996	1.6023	1.5992
n = 2000	0.01	0.4260	0.4263	0.4263	0.4255	0.4250	0.4258	0.4289	0.4286	0.4285	0.4296
	0.025	0.4633	0.4632	0.4615	0.4626	0.4621	0.4679	0.4684	0.4682	0.4681	0.4683
	0.05	0.5029	0.5034	0.5028	0.5027	0.5027	0.5055	0.5058	0.5055	0.5056	0.5067
	0.10	0.5566	0.5555	0.5551	0.5562	0.5551	0.5548	0.5543	0.5542	0.5550	0.5546
	0.25	0.6621	0.6624	0.6617	0.6625	0.6620	0.6605	0.6603	0.6599	0.6601	0.6606
	0.50	0.8144	0.8150	0.8151	0.8146	0.8153	0.8133	0.8133	0.8134	0.8141	0.8138
	0.75	1.0040	1.0036	1.0038	1.0037	1.0031	1.0076	1.0069	1.0067	1.0068	1.0067
	0.90	1.2070	1.2067	1.2071	1.2054	1.2061	1.2158	1.2154	1.2159	1.2144	1.2145
	0.95	1.3440	1.3458	1.3467	1.3448	1.3445	1.3498	1.3496	1.3510	1.3526	1.3508
	0.975	1.4695	1.4653	1.4631	1.4659	1.4669	1.4710	1.4719	1.4728	1.4724	1.4722
	0.99	1.6053	1.6048	1.6024	1.6095	1.6108	1.6142	1.6157	1.6184	1.6174	1.6173

Note. Percentiles computed by generating 20000 draws from the discrete time approximation (direct simulation) to the limiting random variables based on n steps.

Table B.2. *Quantiles of the null distribution under cointegration of the modified CUSUM of squares test statistic,* $\hat{\mathrm{CS}}_n$*. Number of integrated regressors, k.*

		Case of inclusion of a constant term and a linear trend (m = 2)				
Sample size, n		k = 1	2	3	4	5
n = 100	0.01	0.4003	0.4035	0.4029	0.4020	0.4014
	0.025	0.4358	0.4399	0.4418	0.4411	0.4422
	0.05	0.4757	0.4771	0.4793	0.4781	0.4788
	0.10	0.5259	0.5273	0.5272	0.5285	0.5303
	0.25	0.6287	0.6305	0.6311	0.6317	0.6336
	0.50	0.7757	0.7748	0.7777	0.7789	0.7808
	0.75	0.9587	0.9614	0.9636	0.9661	0.9661
	0.90	1.1470	1.1493	1.1506	1.1489	1.1536
	0.95	1.2720	1.2735	1.2725	1.2761	1.2752
	0.975	1.3747	1.3715	1.3751	1.3791	1.3856
	0.99	1.5084	1.5075	1.5096	1.5137	1.5205
n = 200	0.01	0.4056	0.4054	0.4053	0.4073	0.4105
	0.025	0.4440	0.4424	0.4428	0.4456	0.4453
	0.05	0.4837	0.4828	0.4841	0.4847	0.4842
	0.10	0.5361	0.5362	0.5368	0.5376	0.5385
	0.25	0.6382	0.6382	0.6384	0.6401	0.6421
	0.50	0.7857	0.7854	0.7877	0.7884	0.7884
	0.75	0.9734	0.9743	0.9748	0.9776	0.9797
	0.90	1.1714	1.1712	1.1720	1.1752	1.1761
	0.95	1.3025	1.3009	1.3009	1.3025	1.3011
	0.975	1.4158	1.4158	1.4148	1.4161	1.4194
	0.99	1.5325	1.5414	1.5529	1.5558	1.5548
n = 500	0.01	0.4158	0.4178	0.4187	0.4163	0.4161
	0.025	0.4554	0.4558	0.4579	0.4554	0.4561
	0.05	0.4934	0.4938	0.4939	0.4941	0.4957
	0.10	0.5434	0.5441	0.5448	0.5446	0.5447
	0.25	0.6496	0.6498	0.6503	0.6502	0.6497
	0.50	0.8009	0.8015	0.8017	0.8016	0.8027
	0.75	0.9886	0.9889	0.9881	0.9899	0.9913
	0.90	1.1947	1.1946	1.1948	1.1952	1.1957
	0.95	1.3259	1.3273	1.3279	1.3274	1.3278
	0.975	1.4469	1.4402	1.4439	1.4446	1.4451
	0.99	1.6004	1.6008	1.5983	1.5987	1.6025
n = 1000	0.01	0.4190	0.4195	0.4184	0.4199	0.4189
	0.025	0.4580	0.4586	0.4582	0.4587	0.4584
	0.05	0.4973	0.4968	0.4971	0.4982	0.4987
	0.10	0.5499	0.5505	0.5491	0.5488	0.5495
	0.25	0.6559	0.6559	0.6549	0.6555	0.6561
	0.50	0.8060	0.8061	0.8057	0.8067	0.8059
	0.75	0.9991	0.9998	1.0005	1.0007	1.0001
	0.90	1.2034	1.2045	1.2047	1.2035	1.2048
	0.95	1.3385	1.3395	1.3387	1.3416	1.3411
	0.975	1.4592	1.4592	1.4580	1.4599	1.4615
	0.99	1.5894	1.5890	1.5906	1.5916	1.5888
n = 2000	0.01	0.4243	0.4241	0.4238	0.4253	0.4251
	0.025	0.4654	0.4646	0.4644	0.4642	0.4636
	0.05	0.5034	0.5036	0.5033	0.5034	0.5046
	0.10	0.5527	0.5534	0.5529	0.5533	0.5531
	0.25	0.6588	0.6590	0.6592	0.6602	0.6595
	0.50	0.8129	0.8128	0.8129	0.8129	0.8128
	0.75	1.0047	1.0046	1.0048	1.0041	1.0044
	0.90	1.2106	1.2120	1.2114	1.2117	1.2103
	0.95	1.3433	1.3449	1.3432	1.3427	1.3429
	0.975	1.4718	1.4732	1.4716	1.4689	1.4706
	0.99	1.6245	1.6227	1.6251	1.6189	1.6218

Note. Percentiles computed by generating 20000 draws from the discrete time approximation (direct simulation) to the limiting random variables based on n steps.

Table B.3. *Quantiles of the null distribution under cointegration of the modified CUSUM of squares test statistic,* $\hat{\mathrm{CS}}_n(m+q)$

Sample size, n		Order of the polynomial trend function $\boldsymbol{\tau}_t = (\boldsymbol{\tau}'_{m,t}, \boldsymbol{\tau}'_{q,t})'$				
		$m+q=1$	2	3	4	5
$n = 100$	0.01	0.3497	0.3306	0.2958	0.2734	0.2558
	0.025	0.3787	0.3597	0.3189	0.2922	0.2736
	0.05	0.4074	0.3856	0.3398	0.3112	0.2895
	0.10	0.4436	0.4177	0.3674	0.3329	0.3096
	0.25	0.5146	0.4832	0.4179	0.3774	0.3464
	0.50	0.6118	0.5729	0.4874	0.4323	0.3944
	0.75	0.7232	0.6826	0.5678	0.4983	0.4503
	0.90	0.8325	0.7982	0.6518	0.5647	0.5074
	0.95	0.9040	0.8761	0.7064	0.6116	0.5445
	0.975	0.9654	0.9469	0.7601	0.6490	0.5796
	0.99	1.0384	1.0358	0.8229	0.7003	0.6172
$n = 200$	0.01	0.3642	0.3416	0.3085	0.2867	0.2701
	0.025	0.3944	0.3730	0.3328	0.3055	0.2875
	0.05	0.4222	0.4000	0.3544	0.3248	0.3025
	0.10	0.4601	0.4338	0.3817	0.3482	0.3233
	0.25	0.5325	0.4994	0.4346	0.3926	0.3610
	0.50	0.6275	0.5905	0.5025	0.4495	0.4109
	0.75	0.7392	0.7007	0.5866	0.5169	0.4691
	0.90	0.8495	0.8200	0.6733	0.5876	0.5292
	0.95	0.9191	0.8949	0.7305	0.6324	0.5675
	0.975	0.9845	0.9677	0.7841	0.6756	0.6022
	0.99	1.0630	1.0613	0.8519	0.7326	0.6469
$n = 500$	0.01	0.3728	0.3554	0.3220	0.2995	0.2835
	0.025	0.4047	0.3849	0.3452	0.3189	0.3003
	0.05	0.4344	0.4115	0.3690	0.3387	0.3168
	0.10	0.4723	0.4454	0.3961	0.3624	0.3370
	0.25	0.5431	0.5139	0.4479	0.4069	0.3759
	0.50	0.6413	0.6050	0.5183	0.4643	0.4278
	0.75	0.7532	0.7164	0.6021	0.5334	0.4866
	0.90	0.8645	0.8330	0.6892	0.6034	0.5477
	0.95	0.9363	0.9092	0.7495	0.6516	0.5893
	0.975	0.9989	0.9782	0.8041	0.6970	0.6250
	0.99	1.0803	1.0739	0.8686	0.7497	0.6743
$n = 1000$	0.01	0.3813	0.3643	0.3302	0.3057	0.2899
	0.025	0.4122	0.3935	0.3539	0.3261	0.3080
	0.05	0.4394	0.4203	0.3752	0.3458	0.3242
	0.10	0.4776	0.4542	0.4035	0.3688	0.3444
	0.25	0.5514	0.5208	0.4556	0.4128	0.3831
	0.50	0.6491	0.6150	0.5277	0.4713	0.4350
	0.75	0.7648	0.7279	0.6120	0.5421	0.4951
	0.90	0.8758	0.8493	0.6996	0.6133	0.5583
	0.95	0.9446	0.9253	0.7614	0.6649	0.5990
	0.975	1.0113	1.0022	0.8191	0.7089	0.6372
	0.99	1.0843	1.0926	0.8834	0.7652	0.6877
$n = 2000$	0.01	0.3864	0.3688	0.3383	0.3135	0.2959
	0.025	0.4175	0.3976	0.3598	0.3324	0.3132
	0.05	0.4477	0.4247	0.3803	0.3517	0.3299
	0.10	0.4851	0.4582	0.4085	0.3751	0.3501
	0.25	0.5584	0.5273	0.4615	0.4209	0.3884
	0.50	0.6546	0.6197	0.5329	0.4798	0.4409
	0.75	0.7674	0.7315	0.6182	0.5484	0.4998
	0.90	0.8790	0.8526	0.7053	0.6206	0.5641
	0.95	0.9520	0.9308	0.7653	0.6688	0.6046
	0.975	1.0171	1.0111	0.8215	0.7119	0.6446
	0.99	1.0974	1.1057	0.8895	0.7644	0.6909

Note. Percentiles computed by generating 20000 draws from the discrete time approximation (direct simulation) to the limiting random variables based on n steps.

Appendix C. *Finite sample empirical size.*

Table C.1. *Finite sample-adjusted empirical size at 5% nominal level. Case of no deterministic component,* $C\hat{S}_n(k)$*, Bartlett kernel and bandwidth* $q_n(d) = [d(n/100)^{1/4}]$.

$\phi = 0.50$, $\sigma_{k\upsilon} = 0.75$

Sample size, n		d	Number of integrated regressors, k				
			k = 1	2	3	4	5
n = 100	α = 0.00	0	0.0516	0.0506	0.0414	0.0412	0.0378
		2	0.0380	0.0388	0.0354	0.0298	0.0276
		4	0.0352	0.0264	0.0230	0.0212	0.0168
		8	0.0278	0.0230	0.0158	0.0184	0.0100
		12	0.0210	0.0200	0.0172	0.0146	0.0114
	α = 0.25	0	0.0658	0.0728	0.0616	0.0512	0.0454
		2	0.0472	0.0390	0.0338	0.0290	0.0262
		4	0.0426	0.0360	0.0278	0.0258	0.0214
		8	0.0310	0.0228	0.0166	0.0132	0.0092
		12	0.0158	0.0104	0.0126	0.0076	0.0086
	α = 0.50	0	0.1420	0.1462	0.1218	0.1166	0.1068
		2	0.0574	0.0564	0.0524	0.0400	0.0394
		4	0.0422	0.0334	0.0316	0.0290	0.0248
		8	0.0304	0.0262	0.0230	0.0162	0.0126
		12	0.0212	0.0150	0.0124	0.0106	0.0124
	α = 0.75	0	0.4394	0.3914	0.3476	0.3106	0.2846
		2	0.1098	0.1034	0.0928	0.0870	0.0696
		4	0.0586	0.0616	0.0538	0.0398	0.0314
		8	0.0292	0.0264	0.0208	0.0144	0.0148
		12	0.0154	0.0102	0.0090	0.0104	0.0074
n = 200	α = 0.00	0	0.0508	0.0446	0.0452	0.0456	0.0452
		2	0.0434	0.0420	0.0366	0.0360	0.0330
		4	0.0466	0.0440	0.0346	0.0300	0.0298
		8	0.0318	0.0340	0.0290	0.0278	0.0222
		12	0.0328	0.0242	0.0248	0.0200	0.0174
	α = 0.25	0	0.0698	0.0666	0.0642	0.0670	0.0632
		2	0.0460	0.0464	0.0410	0.0412	0.0388
		4	0.0408	0.0400	0.0384	0.0334	0.0276
		8	0.0420	0.0306	0.0326	0.0276	0.0208
		12	0.0300	0.0264	0.0250	0.0204	0.0158
	α = 0.50	0	0.1676	0.1726	0.1620	0.1582	0.1442
		2	0.0798	0.0790	0.0668	0.0702	0.0648
		4	0.0648	0.0546	0.0516	0.0462	0.0442
		8	0.0370	0.0356	0.0376	0.0342	0.0268
		12	0.0376	0.0294	0.0206	0.0178	0.0180
	α = 0.75	0	0.4766	0.4636	0.4388	0.4130	0.3880
		2	0.1484	0.1506	0.1358	0.1264	0.1172
		4	0.0902	0.0770	0.0764	0.0736	0.0680
		8	0.0412	0.0414	0.0366	0.0330	0.0296
		12	0.0256	0.0268	0.0276	0.0200	0.0140

Table C.1. *Finite sample-adjusted empirical size at 5% nominal level. Case of no deterministic component,* $\mathrm{C}\hat{\mathrm{S}}_n(k)$*, Bartlett kernel and bandwidth* $q_n(d) = [d(n/100)^{1/4}]$.
ϕ = 0.50, $\sigma_{k\upsilon}$ = 0.75 (continuation)

Sample size, n		d	Number of integrated regressors, k k = 1	2	3	4	5
n = 500	α = 0.00	0	0.0434	0.0426	0.0424	0.0450	0.0486
		2	0.0448	0.0462	0.0432	0.0420	0.0366
		4	0.0466	0.0392	0.0394	0.0386	0.0382
		8	0.0482	0.0388	0.0388	0.0400	0.0362
		12	0.0402	0.0366	0.0318	0.0292	0.0250
	α = 0.25	0	0.0730	0.0714	0.0740	0.0750	0.0660
		2	0.0518	0.0530	0.0514	0.0488	0.0478
		4	0.0478	0.0490	0.0450	0.0456	0.0482
		8	0.0370	0.0368	0.0378	0.0354	0.0338
		12	0.0372	0.0346	0.0352	0.0302	0.0294
	α = 0.50	0	0.1828	0.1878	0.1814	0.1794	0.1818
		2	0.0848	0.0882	0.0846	0.0826	0.0818
		4	0.0612	0.0628	0.0662	0.0592	0.0596
		8	0.0408	0.0422	0.0394	0.0340	0.0330
		12	0.0380	0.0360	0.0328	0.0296	0.0256
	α = 0.75	0	0.5758	0.5450	0.5546	0.5444	0.5282
		2	0.1924	0.2010	0.1840	0.1862	0.1780
		4	0.0862	0.0814	0.0874	0.0794	0.0752
		8	0.0596	0.0598	0.0508	0.0522	0.0488
		12	0.0418	0.0436	0.0394	0.0358	0.0382
n = 1000	α = 0.00	0	0.0572	0.0540	0.0468	0.0436	0.0500
		2	0.0488	0.0546	0.0522	0.0516	0.0498
		4	0.0500	0.0512	0.0466	0.0498	0.0428
		8	0.0422	0.0422	0.0400	0.0452	0.0404
		12	0.0408	0.0432	0.0356	0.0368	0.0322
	α = 0.25	0	0.0754	0.0756	0.0712	0.0708	0.0696
		2	0.0612	0.0618	0.0628	0.0602	0.0538
		4	0.0418	0.0414	0.0442	0.0404	0.0434
		8	0.0436	0.0462	0.0416	0.0404	0.0394
		12	0.0432	0.0320	0.0390	0.0328	0.0298
	α = 0.50	0	0.1968	0.1848	0.1980	0.1832	0.1782
		2	0.0690	0.0652	0.0646	0.0678	0.0678
		4	0.0382	0.0422	0.0350	0.0330	0.0296
		8	0.0480	0.0490	0.0530	0.0480	0.0472
		12	0.0416	0.0384	0.0404	0.0356	0.0350
	α = 0.75	0	0.5872	0.5898	0.5938	0.5784	0.5820
		2	0.1640	0.1582	0.1544	0.1458	0.1452
		4	0.0812	0.0742	0.0742	0.0748	0.0712
		8	0.0588	0.0506	0.0482	0.0456	0.0426
		12	0.0522	0.0532	0.0472	0.0402	0.0390

Notes. (a) This design corresponds to the situation of contemporaneous endogeneity between integrated regressors and error correction terms when $\sigma_{k\upsilon}$ > 0 and α = 0, irrespective of the value of ϕ, that is $(1/n)\sum_{t=1}^{n} E[\boldsymbol{\eta}_{k,t}u_t] = (1/n)\sum_{t=1}^{n} E[\boldsymbol{\eta}_{k,t}\upsilon_t] = E[\mathbf{w}_{k,t}\upsilon_t] = \boldsymbol{\sigma}_{k,\upsilon}$. **(b)** All the results were computed by generating 5000 draws from the discrete time approximation (direct simulation) to the limiting random variables based on n steps.