Capitulo 2
Integrated Modified OLS estimation of cointegrating regressions with deterministically trending integrated regressors and residual-based tests for cointegration

Julio Afonso

J. Afonso
Department of Institutional Economics, Economic Statistics and Econometrics
Faculty of Economics and Business Administration. University of La Laguna, Camino La Hornera, s/n. Campus de Guajara 38071, San Cristóbal de La Laguna. Tenerife. Canary Islands
email: jafonsor@ull.es

M.Ramos, M.Miranda (eds.) *Estudios en Finanzas y Contabilidad: España y América Latina*. *Estado del arte y las nuevas metodologías aplicadas*, Temas Selectos de Finanzas-©ECORFAN-Madrid, España, 2013

Abstract

In this paper we discuss the asymptotically almost efficient estimation of a univariate static cointegrating regression relationship when we take into account the deterministic structure of the integrated regressors, in a slightly more general framework that considered by Hansen (1992). After reviewing the properties of OLS and Fully Modified OLS (FM-OLS) estimation in this framework, we consider the analysis of the recently proposed Integrated Modified OLS (IM-OLS) estimator by Vogelsang and Wagner (2011) of the cointegrating vector and propose a new proper specification of the integrated modified cointegrating regression equation. This alternative method of bias removal has the advantage over the existing methods that does not require any tuning parameters, such as kernel functions and bandwidths, or lags. Also, based on the sequence of IM-OLS residuals, we propose some new test statistics based on different measures of excessive fluctuation for testing the null hypothesis of cointegration against the alternative of no cointegration. For these test statistics we derive their asymptotic null and alternative distributions, provide the relevant quantiles of the null distribution, and study their finite sample power performance under no cointegration through a simulation experiment.

Keywords: cointegration, asymptotically efficient estimation, OLS, FM-OLS, IM-OLS, trending integrated regressors

2 Introduction

Cointegration analysis is widely used in empirical macroeconomics and finance, and includes both the estimation of cointegrating relationships and hypothesis testing, and also testing the hypothesis of cointegration among nonstationary variables. In the econometric literature there are many contributions in these two topics, some of which deals with these two questions simultaneously. Given the usual linear specification of a potentially cointegrating regression, a first candidate for estimation is the method of ordinary least squares (OLS), that determines superconsistent estimates of the regression parameters under cointegration. However, with endogenous regressors the limiting distribution of the OLS estimator is contaminated by a number of nuisance parameters, also known as second order bias terms, which renders inference problematic. Consequently, there has been proposed several modifications to OLS to makes standard asymptotic inference feasible but at the cost of introducing the choice of several tuning parameters and functions. These methods include the fully modified OLS (FM-OLS) approach of Phillips and Hansen (1990), the canonical cointegrating regression (CCR) by Park (1992), and the dynamic OLS (DOLS) approach of Phillips and Loretan (1991), Saikkonen (1991) and Stock and Watson (1993). This paper deals with the analysis of a new asymptotically almost efficient estimation method of a linear cointegrating regression recently proposed by Vogelsang and Wagner (2011) (henceforth VW) that does not require any additional choice more than the initial standard assumptions on the model specification, making it a very appealing alternative.

This new estimation method, called the integrated modified OLS (IM-OLS) estimator, only requires a very simple transformation, which is free of tuning parameters or any other previous computation, of the model variables that asymptotically produces the same correction effect as the commonly used estimation methods cited above. This simplicity open the possibility to a more straightforward treatment of more complex models incorporating some additional effects and components.

Despite these advantages, the main criticism comes from the fact that, asymptotically, this estimation method does not produce fully efficient estimates as compared with the other existing methods. However, simulation results obtained by the authors seems to indicate that, even in very small samples, the efficiency loss is not severe and hence the estimation results are reliable enough.

An important issue, which is often is not taken into account and that can substantially affect the performance and properties of these estimation procedures, is the nature and structure of the deterministic component, if any, of the generating mechanism of the model variables and its relation with the deterministic component, if is considered, in the specification of the cointegrating regression. Following the work by Hansen (1992), we generalize its formulation by allowing for deterministically trending integrated regressors with a possibly different structure for their deterministic components and propose a simple rule for a proper specification of the deterministic trend function in the cointegrating regression that simultaneously correct for their effects.

Given the particular transformation of the model variables required for performing the asymptotically efficient IM-OLS estimation, we show that a proper accommodation of these components must be based on a previous transformation of the model variables, in particular the OLS detrending. With these corrected observations we perform the IM-OLS estimation of the cointegrating regression and derive the limiting distributions of the resulting estimates and residuals both under the assumption of cointegration and no cointegration.

Based on these new asymptotically efficient estimators of the vector parameters in the cointegrating regression, we consider the building of some simple statistics for testing the null hypothesis of cointegration by using different measures of excessive fluctuation in the IM-OLS residual sequence that cannot be compatible with the stationarity assumption of the error sequence. These new testing procedures are based on the statistics proposed by Shin (1994), Xiao and Phillips (2002) and Wu and Xiao (2008) with the same objective as ours, and make use of two basic measures of excessive fluctuations, the Cramér-von Mises (CvM) and Kolmogorov-Smirnov (KS) metrics. We derive their limiting null and alternative distributions and evaluate their power behavior in finite-samples through a simulation experiment.

The structure of the paper is as follows.

In Section 2.1 we formalize the specification of the data generating process and the econometric model relating the variables, a linear static cointegrating regression, together with the set of assumptions needed to obtain the more relevant distributional results of the usual estimators of the model parameters and residuals. We review the estimation results from some commonly existing estimation methods in this setup, as well as the specification of a set of related semiparametric statistics that we use as a reference for later development of new testing procedures for the null hypothesis of cointegration in Section 4 based on the results of the new estimation method that we analyze.

This section also introduce a very simple to compute testing procedure based on OLS residuals for testing the null hypothesis of no cointegration that generalize a previous one in the univariate analysis of an individual time series.

Section 3 reviews the main characteristics and introduce some new results for the new recently proposed estimation method for an univariate cointegrating regression equation model by Vogelsang and Wagner (2011), which is called the IM-OLS estimator. As for the existing estimation methods reviewed in Section 2.1, we analyze the more appropriate treatment of the underlying deterministic component characterizing the observations of the integrated regressors and also obtain new asymptotic results concerning the behavior of the estimates under the assumption of no cointegration. With these results, Section 4 propose a new set of semiparametric statistics for testing the null hypothesis of cointegration based on the behavior of some simple functionals of the IM-OLS residuals and characterize their limiting distributions, both under cointegration and no cointegration. Finally, all the mathematical proofs are collected in Appendix A while that Appendix B presents some numerical results, including critical values for the proposed testing procedures and the illustration of their behavior through its power performance in finite samples.

2.1 The model, OLS and FM-OLS estimation of the linear cointegrating regression with trending regressors

We assume that the variables of interest, the scalar Y_t and the k-dimensional vector $\mathbf{X}_{k,t} = (X_{1,t},...,X_{k,t})'$, come from the following data generating process (DGP)

$$\begin{pmatrix} Y_t \\ \mathbf{X}_{k,t} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\alpha}'_{0,p}\boldsymbol{\tau}_{p,t} \\ \mathbf{A}_{k,p}\boldsymbol{\tau}_{p,t} \end{pmatrix} + \begin{pmatrix} \eta_{0,t} \\ \boldsymbol{\eta}_{k,t} \end{pmatrix} = \mathbf{A}_p\boldsymbol{\tau}_{p,t} + \boldsymbol{\eta}_t \quad t = 1,...,n \tag{2.1}$$

Where $\boldsymbol{\eta}_t = (\eta_{0,t},\eta'_{k,t})'$ is the stochastic trend component that satisfy the first order recurrence relation $\boldsymbol{\eta}_t = \boldsymbol{\eta}_{t-1} + \boldsymbol{\varepsilon}_t$

With $\boldsymbol{\varepsilon}_t = (\varepsilon_{0,t},\boldsymbol{\varepsilon}'_{k,t})'$ a k+1 vector zero mean sequence of error processes. Also, we consider the general case where both Y_t and each element of the k vector $\mathbf{X}_{k,t} = (X_{1,t},...,X_{k,t})'$ contains a deterministic trend component given by a polynomial trend function of an arbitrary order $p_i \geq 0$, $i = 0, 1, \ldots, k$, that is $d_{i,t} = \boldsymbol{\alpha}'_{i,p_i}\boldsymbol{\tau}_{p_i,t}$, with $\boldsymbol{\alpha}_{i,p_i} = (\boldsymbol{\alpha}_{i,0},\boldsymbol{\alpha}_{i,1},...,\boldsymbol{\alpha}_{i,p_i})'$, and $\boldsymbol{\tau}_{p_i,t} = (1,t,...,t^{p_i})'$. To make this assumption compatible with the standard formulation in (2.1) where all the deterministic trend components appears as if it were of the same type and order, we have to write.

$$\boldsymbol{\alpha}'_{i,p_i}\boldsymbol{\tau}_{p_i,t} = (\boldsymbol{\alpha}'_{i,p_i} : \mathbf{0}'_{p-p_i})\begin{pmatrix} \boldsymbol{\tau}_{p_i,t} \\ \boldsymbol{\tau}_{p-p_i,t} \end{pmatrix} = \boldsymbol{\alpha}'_{i,p}\boldsymbol{\tau}_{p,t}, \quad i = 0, 1, \ldots, k \tag{2.2}$$

With $p = \max(p_0, p_1, ..., p_k)$ and $\mathbf{0}_{p-p_i}$ a $(p-p_i)\times 1$ vector of zeroes, so that.

$$\mathbf{A}_{k,p}\boldsymbol{\tau}_{p,t} = \begin{pmatrix} \boldsymbol{\alpha}'_{1,p_1}\boldsymbol{\tau}_{p_1,t} \\ \vdots \\ \boldsymbol{\alpha}'_{k,p_k}\boldsymbol{\tau}_{p_k,t} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\alpha}'_{1,p} \\ \vdots \\ \boldsymbol{\alpha}'_{k,p} \end{pmatrix}\boldsymbol{\tau}_{p,t} \tag{2.3}$$

With this formulation, we introduce the static potentially cointegrating regression equation between the unobserved stochastic trend components of the elements in $\mathbf{Z}_t$ as

$$\eta_{0,t} = \boldsymbol{\eta}'_{k,t}\boldsymbol{\beta}_k + u_t \tag{2.4}$$

Which gives

$$Y_t = \boldsymbol{\alpha}'_p\boldsymbol{\tau}_{p,t} + \boldsymbol{\beta}'_k\mathbf{X}_{k,t} + u_t \quad t = 1,\ldots,n \tag{2.5}$$

With $\boldsymbol{\alpha}_p = \boldsymbol{\alpha}_{0,p} - \mathbf{A}'_{k,p}\boldsymbol{\beta}_k$. Associated to the deterministic component we introduce the polynomial order trend and sample size dependent scaling matrix $\boldsymbol{\Gamma}_{p,n}$, given by $\boldsymbol{\Gamma}_{p,n} = diag(1,n^{-1},\ldots,n^{-p})$, which gives $\boldsymbol{\tau}_{p,tn} = \boldsymbol{\Gamma}_{p,n}\boldsymbol{\tau}_{p,t} \to \boldsymbol{\tau}_p(r) = (1,r,\ldots,r^p)'$ uniformly over $r \in [0,1]$ as $n\to\infty$. Also we have that $n^{-1}\sum_{t=1}^{[nr]}\boldsymbol{\tau}_{p,tn} \to \int_0^r \boldsymbol{\tau}_p(s)ds$, and $n^{-1}\sum_{t=1}^{n}\boldsymbol{\tau}_{p,tn}\boldsymbol{\tau}'_{p,tn} = n^{-1}\mathbf{Q}_{pp,n} = \bar{\mathbf{Q}}_{pp,n} \to \mathbf{Q}_{pp}$ as $n\to\infty$, with $\mathbf{Q}_{pp}$ be defined as $\mathbf{Q}_{pp} = \int_0^1 \boldsymbol{\tau}_p(s)\boldsymbol{\tau}_p(s)'ds < \infty$,. In order to complete the specification of our data generating process we next introduce a quite general and common assumption on the error terms involved in (2.5).

Assumption 2.1.1

We assume that the error term in the cointegrating regression u_t satisfy the first-order recurrence relation $u_t = \alpha u_{t-1} + \upsilon_t$, with $|\alpha| \leq 1$, where the zero mean (k+1)-dimensional error sequence $\boldsymbol{\zeta}_t = (\upsilon_t, \boldsymbol{\varepsilon}'_{k,t})'$ verify any of the existing conditions that guarantee the validity of the functional central limit theorem (FCLT) approximation of the form

$$\begin{pmatrix} B_{\upsilon,n}(r) \\ \mathbf{B}_{k,n}(r) \end{pmatrix} = n^{-1/2}\sum_{t=1}^{[nr]}\begin{pmatrix} \upsilon_t \\ \boldsymbol{\varepsilon}_{k,t} \end{pmatrix} \Rightarrow \mathbf{B}(r) = \begin{pmatrix} B_\upsilon(r) \\ \mathbf{B}_k(r) \end{pmatrix} = \mathbf{BM}(\boldsymbol{\Omega}_0) = \boldsymbol{\Omega}_0^{1/2}\mathbf{W}(r) \quad 0 \leq r \leq 1 \tag{2.6}$$

With $\mathbf{W}(r) = (W_\upsilon(r), \mathbf{W}'_k(r))'$ a k+1-dimensional standard Brownian motion, and $\boldsymbol{\Omega}_0$ the covariance matrix of $\mathbf{B}(r)$, which is assumed to be positive definite and that can also be interpreted as the long-run covariance matrix of the vector error sequence $\boldsymbol{\zeta}_t$, that is $\boldsymbol{\Omega}_0 = E[\boldsymbol{\zeta}_t\boldsymbol{\zeta}'_t] + \sum_{j=1}^{\infty}(E[\boldsymbol{\zeta}_{t-j}\boldsymbol{\zeta}'_t] + E[\boldsymbol{\zeta}_t\boldsymbol{\zeta}'_{t-j}])$, which can be decomposed as $\boldsymbol{\Omega}_0 = \boldsymbol{\Delta}_0 + \boldsymbol{\Lambda}'_0$, with $\boldsymbol{\Delta}_0 = \boldsymbol{\Sigma}_0 + \boldsymbol{\Lambda}_0 = \sum_{j=0}^{\infty}E[\boldsymbol{\zeta}_{t-j}\boldsymbol{\zeta}'_t]$ the one-sided long-run covariance matrix, where $\boldsymbol{\Sigma}_0 = E[\boldsymbol{\zeta}_t\boldsymbol{\zeta}'_t]$, and $\boldsymbol{\Lambda}_0 = \sum_{j=1}^{\infty}E[\boldsymbol{\zeta}_{t-j}\boldsymbol{\zeta}'_t]$. This covariance matrix is partitioned according to the components of $\boldsymbol{\zeta}_t$ as ω_υ^2, $\boldsymbol{\omega}_{k\upsilon} = \boldsymbol{\omega}'_{\upsilon k}$, and $\boldsymbol{\Omega}_{kk}$. The assumption of positive definiteness of $\boldsymbol{\Omega}_0$ excludes cointegration among the k integrated regressors $\mathbf{X}_{k,t}$ (subcointegration) with $\mathbf{B}_k(r) = \mathbf{BM}(\boldsymbol{\Omega}_{k,k})$, $\boldsymbol{\Omega}_{k,k} > 0$. Given the upper triangular Cholesky decomposition of the matrix $\boldsymbol{\Omega}_0$, we then have that $B_\upsilon(r) = B_{\upsilon.k}(r) + \boldsymbol{\gamma}'_{k\upsilon}\mathbf{B}_k(r)$, with $B_{\upsilon.k}(r) = \omega_{\upsilon.k}W_\upsilon(r)$, and $\mathbf{B}_k(r) = \boldsymbol{\Omega}_{k,k}^{1/2}\mathbf{W}_k(r)$, where $\boldsymbol{\gamma}_{k\upsilon} = \boldsymbol{\Omega}_{k,k}^{-1}\boldsymbol{\omega}'_{\upsilon,k}$ and $\omega_{\upsilon.k}^2 = E[B_{\upsilon.k}(r)^2] = E[B_{\upsilon.k}(r)B_\upsilon(r)] = \omega_\upsilon^2 - \boldsymbol{\omega}_{\upsilon,k}\boldsymbol{\Omega}_{k,k}^{-1}\boldsymbol{\omega}'_{\upsilon,k}$ is the conditional variance of $B_\upsilon(r)$ given $\mathbf{B}_k(r)$, which gives $E[\mathbf{B}_k(r)B_{\upsilon.k}(r)] = \mathbf{0}_k$.

For the initial values $\boldsymbol{\eta}_{k,0}$ and u_0, we introduce the very general conditions $\boldsymbol{\eta}_{k,0} = o_p(n^{1/2})$, and $u_0 = o_p(n^{1/2})$, which include the particular case of constant finite values. In the case of a stationary error term u_t, with |α| < 1, we then have that $n^{-1/2}\sum_{t=1}^{[nr]} u_t \Rightarrow B_u(r) = (1-\alpha)^{-1}B_\upsilon(r)$, with $B_u(r) = B_{u.k}(r) + \boldsymbol{\gamma}_k'\mathbf{B}_k(r)$, $\boldsymbol{\gamma}_k = \boldsymbol{\Omega}_{kk}^{-1}\boldsymbol{\omega}_{ku}$, $E[B_u(r)^2] = \omega_u^2 = (1-\alpha)^{-2}\omega_\upsilon^2$, $E[B_{u.k}(r)^2] = \omega_{u.k}^2 = \omega_u^2 - \boldsymbol{\gamma}_k'\boldsymbol{\Omega}_{kk}\boldsymbol{\gamma}_k = (1-\alpha)^{-2}\omega_{\upsilon.k}^2$, and $E[\mathbf{B}_k(r)B_u(r)] = \boldsymbol{\omega}_{ku} = (1-\alpha)^{-1}\boldsymbol{\omega}_{k\upsilon}$, while that in the case of nonstationarity, that is when α = 1, then $n^{-1/2}u_{[nr]} \Rightarrow B_u(r) = B_\upsilon(r)$, with $\omega_u^2 = \omega_\upsilon^2$. This means that, from the initial condition on $\boldsymbol{\zeta}_t = (\upsilon_t, \boldsymbol{\varepsilon}_{k,t}')'$, we have that under stationary error terms u_t the sequence $\boldsymbol{\xi}_t = (u_t, \boldsymbol{\varepsilon}_{k,t}')'$ also satisfy a multivariate invariance principle with a long-run covariance matrix $\boldsymbol{\Omega}$ given by the components ω_u^2, $\boldsymbol{\omega}_{ku} = \boldsymbol{\omega}_{uk}'$, and $\boldsymbol{\Omega}_{kk}$. Particular attention must be paid to the long-run covariance vector between $\boldsymbol{\varepsilon}_{k,t}$ and u_t, $\boldsymbol{\omega}_{ku}$, given that it controls for the endogeneity of the integrated regressors in the cointegrating regression model.

With these results then we have:

$$n^{-(1-v)}U_{[nr]} = n^{-(1-v)}\sum_{t=1}^{[nr]} u_t \Rightarrow J_u(r) = \begin{cases} B_u(r) & v = 1/2 \\ \int_0^r B_u(s)ds & v = -1/2 \end{cases} \tag{2.7}$$

With v =1/2 and v = –1/2 indicating, respectively, the stationary and nonstationary cases.

This formulation forms the base for obtaining standard limiting distributional results for the estimators of model parameters and residuals both under cointegration and no cointegration. However, there are some other useful formulations that allow for a more general and unified treatment of the different behavior of this scaled partial sum process of the error correction terms u_t under these two situations. One can cite, for example, the local-to-unity approach introduced by Phillips (1987) that considers the situation where the autoregressive parameter α depends on the sample size as $\alpha = \alpha_n = 1 + n^{-1}\lambda$, with λ ≤ 0, where λ < 0 corresponds to the stationarity case (that is, cointegration), while that λ = 0 corresponds to the nonstationary case (that is, no cointegration) irrespective to the sample size. Taking into account that we can write $\alpha_n = \exp(\lambda/n) + O(n^{-2})$ for small values of λ, then we have that:

$$J_{n,\lambda}(r) = n^{-1/2}u_{[nr]} = (u_0/\sqrt{n})e^{[nr]\lambda/n} + n^{-1/2}\sum_{t=1}^{[nr]} e^{([nr]-t)\lambda/n}\upsilon_t + o_p(1) \Rightarrow J_\lambda(r) \tag{2.8}$$

Under the above assumption on the initial value u_0, where the weak limit $J_\lambda(r)$ in the near-integration case is given by the Gaussian process $J_\lambda(r) = \int_0^r e^{(r-s)\lambda}dB_\upsilon(s) = B_\upsilon(r) + \lambda\int_0^r e^{(r-s)\lambda}B_\upsilon(s)ds$ that is called an Ornstein-Uhlenbeck process which, for fixed r > 0, has the distribution $J_\lambda(r) =^d N(0,(e^{2r\lambda}-1)/2\lambda)$.

An alternative, and more recent, formulation has been proposed by Jansson (2005a, b), in the context of developing a point optimal cointegration test and also incorporating a local-to-unity representation, which is given by $\Delta u_t = (1-\theta L)\upsilon_t$, where $\theta = \theta_n = 1 - n^{-1}\lambda$, with $\lambda \geq 0$. In this framework, the value $\lambda = 0$ corresponds to the case of stationarity, while that $\lambda > 0$ corresponds to no cointegration, as can be checked when writting $u_t = u_0 - \theta\upsilon_0 + \theta\upsilon_t + (\lambda/\sqrt{n})[(1/\sqrt{n})\sum_{i=1}^{t}\upsilon_i]$, which gives.

$$n^{-1/2}U_{[nr]} = \tfrac{[nr]}{\sqrt{n}}(u_0 - \theta\upsilon_0) + \theta n^{-1/2}\sum_{t=1}^{[nr]}\upsilon_t + \lambda(1/n)\sum_{t=1}^{[nr]}(1/\sqrt{n})\sum_{i=1}^{t}\upsilon_i \qquad (2.9)$$

With weak limit $n^{-1/2}U_{[nr]} \Rightarrow B_\upsilon(r) + \lambda\int_0^r B_\upsilon(s)ds$, only when the initial values u_0 and υ_0 are both of order $o_p(n^{-1/2})$. These two cases provide very different stochastic limits but both determine the same orders of convergence for the estimates of the model parameters $\boldsymbol{\alpha}_p$ and $\boldsymbol{\beta}_k$ (as will be seen later), given by the ones corresponding to the case of no cointegration in the standard framework.[12]

Once discussed all these questions concerning the different representations and stochastic properties of the correction error term u_t in the cointegrating regression model, in what follows we will retain the most standard formulation introduced in (2.7) to address the central issue of this study which is the study of the properties and behavior of some alternative and commonly used estimation methods and test statistics and of the new estimation method considered in Section 2..

Given the specification of the linear static cointegrating regression equation (2.5), the standard approach is to use Ordinary Least Squares (OLS) method to estimate the vector parameters $\boldsymbol{\alpha}_p$ and $\boldsymbol{\beta}_k$, which gives:

$$\begin{pmatrix} \hat{\boldsymbol{\alpha}}_{p,n} - \boldsymbol{\alpha}_p \\ \hat{\boldsymbol{\beta}}_{k,n} - \boldsymbol{\beta}_k \end{pmatrix} = \left(\sum_{t=1}^{n} \begin{pmatrix} \boldsymbol{\tau}_{p,t} \\ \mathbf{X}_{k,t} \end{pmatrix} (\boldsymbol{\tau}'_{p,t}, \mathbf{X}'_{k,t}) \right)^{-1} \sum_{t=1}^{n} \begin{pmatrix} \boldsymbol{\tau}_{p,t} \\ \mathbf{X}_{k,t} \end{pmatrix} u_t \qquad (2.10)$$

Taking into account the structure for the deterministic and stochastic trend components of the observed processes Y_t and $\mathbf{X}_{k,t}$ in (2.1), we can write:

$$\begin{pmatrix} \boldsymbol{\tau}_{p,t} \\ \mathbf{X}_{k,t} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\Gamma}_{p,n}^{-1}\boldsymbol{\tau}_{p,tn} \\ \mathbf{A}_{k,p}\boldsymbol{\Gamma}_{p,n}^{-1}\boldsymbol{\tau}_{p,tn} + \boldsymbol{\eta}_{k,t} \end{pmatrix} = \mathbf{W}_n \begin{pmatrix} \boldsymbol{\tau}_{p,tn} \\ n^{-1/2}\boldsymbol{\eta}_{k,t} \end{pmatrix} = \mathbf{W}_n \begin{pmatrix} \boldsymbol{\tau}_{p,tn} \\ \boldsymbol{\eta}_{k,tn} \end{pmatrix} \qquad (2.11)$$

With weighting matrix $\mathbf{W}_n$ given by

[12] For a recent study on the asymptotic and finite-sample properties and behavior of some commonly used estimation methods in the presence of highly persistent regression errors (or, equivalently, strongly serially correlated error terms) see, for example, Kurozumi and Hayakawa (2009). This paper makes use of a related but different approach to the above considered *n* local-to-unity system, the so called *m* local-to-unity system, which seems more appropriate when focus on the cointegrating relation. Also, for a recent study of the asymptotic and finite-sample properties of a variety of estimation methods for a single cointegrating regression model making use of a modified version of the *n* local-to-unity approach and the one proposed by Jansson (2005a, b), see Afonso-Rodríguez (2013).

$$\mathbf{W}_n = \begin{pmatrix} \mathbf{\Gamma}_{p,n}^{-1} & \mathbf{0}_{p+1,k} \\ \mathbf{A}_{k,p}\mathbf{\Gamma}_{p,n}^{-1} & \sqrt{n}\mathbf{I}_{k,k} \end{pmatrix} \tag{2.12}$$

Which allows the reweighted regressors $(\boldsymbol{\tau}'_{p,tn}, n^{-1/2}\boldsymbol{\eta}'_{k,tn})$ to converge weakly to a full-ranked process, so that the OLS estimation error of $\boldsymbol{\alpha}_p$ and $\boldsymbol{\beta}_k$ can be written as

$$\begin{aligned} \begin{pmatrix} \hat{\boldsymbol{\alpha}}_{p,n} - \boldsymbol{\alpha}_p \\ \hat{\boldsymbol{\beta}}_{k,n} - \boldsymbol{\beta}_k \end{pmatrix} &= n^{-v}(\mathbf{W}'_n)^{-1}\left((1/n)\sum_{t=1}^{n}\begin{pmatrix} \boldsymbol{\tau}_{p,tn} \\ n^{-1/2}\boldsymbol{\eta}_{k,t} \end{pmatrix}(\boldsymbol{\tau}'_{p,tn}, n^{-1/2}\boldsymbol{\eta}'_{k,t})\right)^{-1} \\ &\quad \times n^{-(1-v)}\sum_{t=1}^{n}\left\{\begin{pmatrix} \boldsymbol{\tau}_{p,tn} \\ n^{-1/2}\boldsymbol{\eta}_{k,t} \end{pmatrix}u_t\right\} \end{aligned} \tag{2.13}$$

With the power v taking values ±1/2 depending on the stochastic properties of the error sequence u_t, and determining the order of consistency of the OLS estimates. From this expression we have that $n^v\mathbf{W}'_n[(\hat{\boldsymbol{\alpha}}_{p,n} - \boldsymbol{\alpha}_p)', (\hat{\boldsymbol{\beta}}_{k,n} - \boldsymbol{\beta}_k)']'$ can be written as:

$$\begin{aligned} \hat{\boldsymbol{\Theta}}_n(v) = \begin{pmatrix} \hat{\boldsymbol{\Theta}}_{p,n}(v) \\ \hat{\boldsymbol{\Theta}}_{k,n}(v) \end{pmatrix} &= \begin{pmatrix} n^v\mathbf{\Gamma}_{p,n}^{-1}[(\hat{\boldsymbol{\alpha}}_{p,n} - \boldsymbol{\alpha}_p) + \mathbf{A}'_{k,p}(\hat{\boldsymbol{\beta}}_{k,n} - \boldsymbol{\beta}_k)] \\ n^{1/2+v}(\hat{\boldsymbol{\beta}}_{k,n} - \boldsymbol{\beta}_k) \end{pmatrix} \\ &= \left((1/n)\sum_{t=1}^{n}\begin{pmatrix} \boldsymbol{\tau}_{p,tn} \\ n^{-1/2}\boldsymbol{\eta}_{k,t} \end{pmatrix}(\boldsymbol{\tau}'_{p,tn}, n^{-1/2}\boldsymbol{\eta}'_{k,t})\right)^{-1}\left\{n^{-(1-v)}\sum_{t=1}^{n}\begin{pmatrix} \boldsymbol{\tau}_{p,tn} \\ n^{-1/2}\boldsymbol{\eta}_{k,t} \end{pmatrix}u_t\right\} \end{aligned} \tag{2.14}$$

The usual result in this context is as in (2.14) but with $\mathbf{A}_{k,p} = \mathbf{0}_{k,p+1}$, that corresponds to the case where the integrated regressors have no deterministic component which, in our formulation, implies that the deterministic term appearing in the cointegrating equation must correspond to the one contained in Y_t.

The first relevant question concerning the effect of explicitly considering the structure of the deterministic components underlying the observations of the integrated regressors is that, in general terms, the OLS estimator of the trend parameters in the cointegrating regression model contains is biased in finite samples and is asymptotically unbiased only when $p = 0$, that is, when all the integrated regressors only contain at most a constant term. To see this, the first term in (2.14) can be written as

$$\begin{aligned} &\mathbf{\Gamma}_{p,n}^{-1}[(\hat{\boldsymbol{\alpha}}_{p,n} - \boldsymbol{\alpha}_p) + \mathbf{A}'_{k,p}(\hat{\boldsymbol{\beta}}_{k,n} - \boldsymbol{\beta}_k)] \\ &\qquad = \mathbf{Q}_{pp,n}^{-1}\sum_{t=1}^{n}\boldsymbol{\tau}_{p,tn}u_t - \mathbf{Q}_{pp,n}^{-1}\sum_{t=1}^{n}\boldsymbol{\tau}_{p,tn}(n^{-1/2}\boldsymbol{\eta}'_{k,t})[\sqrt{n}(\hat{\boldsymbol{\beta}}_{k,n} - \boldsymbol{\beta}_k)] \\ &\qquad = n^{-v}\left\{\bar{\mathbf{Q}}_{pp,n}^{-1}n^{-(1-v)}\sum_{t=1}^{n}\boldsymbol{\tau}_{p,tn}u_t - \bar{\mathbf{Q}}_{pp,n}^{-1}(1/n)\sum_{t=1}^{n}\boldsymbol{\tau}_{p,tn}(n^{-1/2}\boldsymbol{\eta}'_{k,t})\hat{\boldsymbol{\Theta}}_{k,n}(v)\right\} \end{aligned} \tag{2.15}$$

Where $\hat{\boldsymbol{\Theta}}_{k,n}(v) = n^{1/2+v}(\hat{\boldsymbol{\beta}}_{k,n} - \boldsymbol{\beta}_k)$ indicates the scaled OLS estimator of the cointegrating vector. This gives that the scaled OLS estimator of $\boldsymbol{\alpha}_p$ can be written as

$$n^{v}\boldsymbol{\Gamma}_{p,n}^{-1}(\hat{\boldsymbol{\alpha}}_{p,n}-\boldsymbol{\alpha}_{p})=-n^{-1/2}\boldsymbol{\Gamma}_{p,n}^{-1}\mathbf{A}_{k,p}'\hat{\boldsymbol{\Theta}}_{k,n}(v) \\ +\left\{\bar{\mathbf{Q}}_{pp,n}^{-1}n^{-(1-v)}\sum_{t=1}^{n}\boldsymbol{\tau}_{p,tn}u_{t}-\bar{\mathbf{Q}}_{pp,n}^{-1}(1/n)\sum_{t=1}^{n}\boldsymbol{\tau}_{p,tn}(n^{-1/2}\boldsymbol{\eta}_{k,t}')\hat{\boldsymbol{\Theta}}_{k,n}(v)\right\} \tag{2.16}$$

Where $n^{-1/2}\boldsymbol{\Gamma}_{p,n}^{-1}=diag(1/\sqrt{n},\sqrt{n},...,n^{p-1/2})$, so that the first term contains the bias caused by the trend parameters in $\mathbf{A}_{k,p}=(\boldsymbol{\alpha}_{1,p}',...,\boldsymbol{\alpha}_{k,p}')'$, while that the estimate of $\boldsymbol{\beta}_k$ is exactly invariant to the presence of deterministic trends in the regressors, i.e., to the values of $\mathbf{A}_{k,p}$. Hansen (1992) has considered a similar situation, but assuming that $Y_t=\eta_{0,t}$ with $d_{0,t}=\boldsymbol{\alpha}_{0,p_0}'\boldsymbol{\tau}_{p_0,t}=0$, and $p_i = m$, $i = 1, ..., k$, with $\boldsymbol{\tau}_{m,t}=(t^{p_1},t^{p_2},...,t^{p_m})'$, $1 \leq p_1 <...< p_m$, and scaling matrix $\boldsymbol{\Gamma}_{m,n}=diag(n^{-p_1},n^{-p_2},...,n^{-p_m})$ (see Theorem 1(a, b), p.93).[13] The main differences with our approach are the no inclusion of a constant term and the inclusion of a rank condition on the coefficient matrix $\mathbf{A}_{k,m}$, particularly, rank($\mathbf{A}_{k,m}$) = $m \leq k$. Then, from (2.1) we have that

$$\boldsymbol{\Gamma}_{m,n}[(\mathbf{A}_{k,m}'\mathbf{A}_{k,m})^{-1}\mathbf{A}_{k,m}']\mathbf{X}_{k,[nr]}=\boldsymbol{\tau}_{m,[nr]n}+\sqrt{n}\boldsymbol{\Gamma}_{m,n}[(\mathbf{A}_{k,m}'\mathbf{A}_{k,m})^{-1}\mathbf{A}_{k,m}'](n^{-1/2}\boldsymbol{\eta}_{k,[nr]}) \\ =\boldsymbol{\tau}_{m,[nr]n}+O_p(n^{-(p_1-1/2)})\Rightarrow\boldsymbol{\tau}_m(r) \tag{2.17}$$

Which allows the possibility to develop a sequence of weights which yield a nondegenerate design limiting matrix when estimating (2.3) by OLS under the restriction $\boldsymbol{\alpha}_m=\mathbf{0}_m$. However, as can see from the previous result, this only yields consistent results when $p_1 \geq 1$, and there is no constant term in the regression neither in the polynomial trend function.[14] Under the assumption of cointegration (v = 1/2), then the limiting distribution of the last term in (2.6) is given by

$$n^{-(1-v)}\sum_{t=1}^{[nr]}\begin{pmatrix}\boldsymbol{\tau}_{p,tn}\\ n^{-1/2}\boldsymbol{\eta}_{k,t}\end{pmatrix}u_t\Rightarrow\int_0^r\begin{pmatrix}\boldsymbol{\tau}_p(s)\\ \mathbf{B}_k(s)\end{pmatrix}dB_u(s)+\begin{pmatrix}\mathbf{0}_{p+1}\\ r\boldsymbol{\Delta}_{k,u}\end{pmatrix} \\ =\left\{\int_0^r\begin{pmatrix}\boldsymbol{\tau}_p(s)\\ \mathbf{B}_k(s)\end{pmatrix}dB_{u.k}(s)+\int_0^r\begin{pmatrix}\boldsymbol{\tau}_p(s)\\ \mathbf{B}_k(s)\end{pmatrix}d\mathbf{B}_k(s)'\gamma_k\right\}+\begin{pmatrix}\mathbf{0}_{p+1}\\ r\boldsymbol{\Delta}_{k,u}\end{pmatrix} \tag{2.18}$$

With $\boldsymbol{\Delta}_{k,u}=\sum_{j=0}^{\infty}E[\boldsymbol{\varepsilon}_{k,t-j}u_t]$ given by the probability limit of

[13] In the context of evaluating the effects of detrending in the estimation of a cointegrating regresión, it is also worth to mention the work by Xiao and Phillips (1999), where the authors compare the results of OLS detrending and detrending after quasi-differencing when the variables in a multivariate cointegrated VAR model contain a deterministic trend function. However, given the differences between our framework of analysis and the one used in this paper, we are not going to make use of their results.

[14] See also Hassler (2001) for a related study in the case where the specification of the cointegrating regression equation does not include any deterministic term but the integrated regressors $\mathbf{X}_{k,t}$ do contain a constant term.

$\Delta_{n,ku} = n^{-1}\sum_{t=1}^{n} E[\eta_{k,t}u_t]$.[15]This limiting distribution contains the second-order bias due to the correlation, both contemporaneous and over time, between the error term u_t and $\varepsilon_{k,t}$ (endogeneity of the stochastic trend components of the regressors), and the non-centrality bias that comes from the fact that the regression errors are serially correlated through the parameter $\Delta_{k,u}$. This second-order bias determine a miscentering, an asymmetry, and a nonscale nuisance parameter dependency to the limit distribution of $\hat{\beta}_{k,n}$. For the first term above we have that $\int_0^1 \mathbf{B}_k(s)dB_{u.k}(s) = \omega_{u.k}\Omega_{kk}^{1/2}\int_0^1 \mathbf{W}_k(s)dW_{u.k}(s)$ where, given the independence between $\mathbf{B}_k(r)$ and $B_{u.k}(r)$, conditioning on $\mathbf{B}_k(r)$ (or $\mathbf{W}_k(r)$) can be used to show that this term is a zero mean Gaussian mixture of the form

$$\int_0^1 \mathbf{W}_k(s)dW_{u.k}(s) = \int_{\mathbf{G}_{k,k}>0} N(\mathbf{0}_k, \mathbf{G}_{k,k}^{-1})dP(\mathbf{G}_{k,k}), \quad \mathbf{G}_{k,k} = \left(\int_0^1 \mathbf{W}_k(s)\mathbf{W}_k(s)'\right)^{-1} \tag{2.19}$$

The second term in the expression between brackets is a matrix unit root distribution, arising from the *k* stochastic trends in $\mathbf{X}_{k,t}$, which is cancelled under strict exogeneity of the regressors, that is when $\omega_{ku} = \mathbf{0}_k$. Finally, from the decomposition of $\mathbf{X}_{k,t}$ in (2.1) and the results in (2.14) and (2.15), the sequence of OLS residuals given by

$$\begin{aligned} \hat{u}_{t,p}(k) &= u_t - \tau_{p,t}'(\hat{\alpha}_{p,n} - \alpha_p) - \mathbf{X}_{k,t}'(\hat{\beta}_{k,n} - \beta_k) \\ &= u_t - \tau_{p,tn}'(\Gamma_{p,n}^{-1}[(\hat{\alpha}_{p,n} - \alpha_p) + \mathbf{A}_{k,p}'(\hat{\beta}_{k,n} - \beta_k)]) - (n^{-1/2}\eta_{k,t}')[\sqrt{n}(\hat{\beta}_{k,n} - \beta_k)] \end{aligned} \tag{2.20}$$

That can also be written as

$$\begin{aligned} \hat{u}_{t,p}(k) &= u_t - \tau_{p,tn}'\mathbf{Q}_{pp,n}^{-1}\sum_{j=1}^{n}\tau_{p,jn}u_j \\ &\quad - n^{-1/2}\left\{\eta_{k,t}' - \tau_{p,tn}'\mathbf{Q}_{pp,n}^{-1}\sum_{j=1}^{n}\tau_{p,jn}\eta_{k,j}'\right\}[\sqrt{n}(\hat{\beta}_{k,n} - \beta_k)] \\ &= u_{t,p} - n^{-(1/2+v)}\eta_{kt,p}'[n^{1/2+v}(\hat{\beta}_{k,n} - \beta_k)] \end{aligned} \tag{2.21}$$

So that the OLS residuals are exactly invariant to the trend parameters, and are decomposed in terms of the detrended versions of u_t and $\eta_{k,t}$ as defined in (2.8). For later use, we define the partial sum of the detrended errors in the cointegrating regression (2.3) as $U_{t,p} = \sum_{j=1}^{t} u_{j,p}$, with:

$$n^{-(1-v)}U_{[nr],p} = n^{-(1-v)}\sum_{t=1}^{[nr]}u_{t,p} = n^{-(1-v)}\sum_{t=1}^{[nr]}u_t - n^{-1}\sum_{t=1}^{[nr]}\tau_{p,tn}'\bar{\mathbf{Q}}_{n,pp}^{-1}n^{-(1-v)}\sum_{t=1}^{n}\tau_{p,tn}u_t \tag{2.22}$$

Where, asymptotically, we get

$$n^{-(1-v)}U_{[nr],p} \Rightarrow J_{u,p}(r) = \begin{cases} V_{u,p}(r) & v = 1/2 \quad (|\alpha| < 1) \\ \int_0^r B_{u,p}(s)ds & v = -1/2 \ (\alpha = 1) \end{cases} \tag{2.23}$$

With $V_{u,p}(r)$ a generalized (*p*+1)th-level Brownian bridge process given by

---

[15] The result $r\Delta_{k,u}$ is obtained by writing $\Delta_{n,ku}(r) = n^{-1}\sum_{t=1}^{[nr]}E[\eta_{k,t}u_t] = \frac{[nr]}{n}([nr])^{-1}\sum_{t=1}^{[nr]}E[\eta_{k,t}u_t]$, so that $\Delta_{n,ku}(r) = \frac{[nr]}{n}[([nr])^{-1}\sum_{t=1}^{[nr]}E[\eta_{k,0}u_t] + \sum_{j=0}^{[nr]-1}(([nr])^{-1}\sum_{t=j+1}^{[nr]}E[\varepsilon_{k,t-j}u_t])]$ and the use of the initial condition $\eta_{k,0}$, and Assumption 2.1 on the properties of the error terms.

$$V_{u,p}(r) = B_u(r) - \int_0^r \boldsymbol{\tau}_p'(s)ds\,\mathbf{Q}_{pp}^{-1}\int_0^1 \boldsymbol{\tau}_p(s)dB_u(s) \tag{2.24}$$

With variance $E[V_{u,p}(r)^2] = \omega_u^2 \cdot b_p(r)$, where $b_p(r) = r - \int_0^r \boldsymbol{\tau}_p'(s)ds\,\mathbf{Q}_{pp}^{-1}\int_0^r \boldsymbol{\tau}_p(s)ds$, and $B_{u,p}(r)$ a (p+1)th-order detrended Brownian motion process defined as

$$B_{u,p}(r) = B_u(r) - \boldsymbol{\tau}_p'(r)\mathbf{Q}_{pp}^{-1}\int_0^1 \boldsymbol{\tau}_p(s)B_u(s)ds \tag{2.25}$$

As the stochastic limits in (2.9).[16] Finally, making use of (2.24), and the relation $B_u(r) = B_{u.k}(r) + \boldsymbol{\gamma}_k'\mathbf{B}_k(r)$ we then have that $V_{u,p}(r)$ can be decomposed as

$$\begin{aligned} V_{u,p}(r) &= B_{u.k}(r) - \int_0^r \boldsymbol{\tau}_p'(s)ds\,\mathbf{Q}_{pp}^{-1}\int_0^1 \boldsymbol{\tau}_p(s)dB_{u.k}(s) \\ &\quad + \boldsymbol{\gamma}_k'\ \mathbf{B}_k(r) - \int_0^1 d\mathbf{B}_k(s)\boldsymbol{\tau}_p(s)'\mathbf{Q}_{pp}^{-1}\int_0^r \boldsymbol{\tau}_p'(s)ds\ = V_{u.k,p}(r) + \boldsymbol{\gamma}_k'\mathbf{V}_{k,p}(r) \end{aligned} \tag{2.26}$$

Where, by construction, it is verified $E[\mathbf{V}_{k,p}(r)V_{u.k,p}(r)] = E[\mathbf{B}_{k,p}(r)V_{u.k,p}(r)] = \mathbf{0}_k$, with $\mathbf{B}_{k,p}(r)$ defined in (2.14) below, and $Var[V_{u.k,p}(r)] = \omega_{u.k}^2 \cdot b_p(r)$.

In order to complete this analysis and to establish the basis for our proposal in the next section, we consider an alternative specification to the cointegrating regression equation (2.3). By applying the partitioned OLS estimation to the regression equation (2.5) with respect to the trend parameters, we have that this model can also be written as

$$\hat{Y}_{t,p} = \boldsymbol{\beta}_k'\hat{\mathbf{X}}_{kt,p} + u_{t,p},\quad t = 1,\ldots,n \tag{2.27}$$

Where $\hat{Y}_{t,p} = Y_t - \boldsymbol{\tau}_{p,tn}'\mathbf{Q}_{pp,n}^{-1}\sum_{j=1}^n \boldsymbol{\tau}_{p,jn}Y_j$, $\hat{\mathbf{X}}_{kt,p} = \mathbf{X}_{k,t} - \sum_{j=1}^n \mathbf{X}_{k,j}\boldsymbol{\tau}_{p,jn}'\mathbf{Q}_{pp,n}^{-1}\boldsymbol{\tau}_{p,tn}$, and $u_{t,p} = u_t\ - \boldsymbol{\tau}_{p,tn}'\mathbf{Q}_{pp,n}^{-1}\sum_{j=1}^n \boldsymbol{\tau}_{p,jn}u_j$ denote the detrended observations of the model variables obtained by OLS fitting of their original observations to a pth-order polynomial trend function, where p is chosen according to the rule $p \geq \max(p_0, p_1, \ldots, p_k)$ in the case where the polynomial trend functions in Y_t and each component of $\mathbf{X}_{k,t}$ differ in their orders. The next Proposition 2.1 determines the effectiveness of this procedure to make the OLS-based estimation results invariant to the trend parameters in (2.1).

Proposition 2.1. *Given (2.1)-(2.2), when considering the OLS detrending of* Y_t *and* $\mathbf{X}_{k,t}$

[16] Explicit expressions for these two limiting processes, $V_{u,p}(r)$ and $B_{u,p}(r)$, can be obtained in the leading cases of p = 0 (constant), and p = 1 (constant and linear trend). Specifically, we have that $V_{u,0}(r) = B_u(r) - rB_u(1)$, and $V_{u,1}(r) = B_u(r) + (2-3r)rB_u(1) - 6r(1-r)\int_0^1 B_u(s)ds$ for the first and second-level Brownian bridge, while that $B_{u,0}(r) = B_u(r) - \int_0^1 B_u(s)ds$, and $B_{u,1}(r) = B_u(r) + 2(3r-2)\int_0^1 B_u(s)ds$ $-2(6r-3)\int_0^1 sB_u(s)ds$ are the particular expressions for the demeaned and demeaned and detrended Brownian processes, respectively.

by fitting a polynomial trend function of order p = max(p_0, p_1, ..., p_k) to each of these variables, then we have that

$$\hat{Y}_{t,p} = \eta_{0t,p} = \eta_{0,t} - \boldsymbol{\tau}'_{p,tn}\mathbf{Q}^{-1}_{pp,n}\sum_{j=1}^{n}\boldsymbol{\tau}_{p,jn}\eta_{0,j} \tag{2.28}$$

And

$$\hat{\mathbf{X}}_{kt,p} = \boldsymbol{\eta}_{kt,p} = \boldsymbol{\eta}_{k,t} - \sum_{j=1}^{n}\boldsymbol{\eta}_{k,j}\boldsymbol{\tau}'_{p,jn}\mathbf{Q}^{-1}_{pp,n}\boldsymbol{\tau}_{p,tn} = (\eta_{1t,p},\ldots,\eta_{kt,p})' \tag{2.29}$$

Where $\eta_{0t,p}$ *and* $\boldsymbol{\eta}_{kt,p}$ *are generalized detrended transformations of* $\eta_{0,t}$ *and* $\boldsymbol{\eta}_{k,t}$, *with*

$$n^{-1/2}\boldsymbol{\eta}_{k[nr],p} \Rightarrow \mathbf{B}_{k,p}(r) = \mathbf{B}_k(r) - \int_0^1 \mathbf{B}_k(s)\boldsymbol{\tau}_p(s)'ds\,\mathbf{Q}^{-1}_{pp}\boldsymbol{\tau}_p(r) \tag{2.30}$$

a (p+1)-order detrended transformation of $\mathbf{B}_k(r)$. *According to Lemma A.2 in Phillips and Hansen (1990),* $\mathbf{B}_{k,p}(r) = \mathbf{BM}(\boldsymbol{\Omega}_{k,k}\cdot v_p(r))$ *is a full rank Gaussian processes, with* $v_p(r)$ *a scalar function of r and* $\boldsymbol{\tau}_p(\cdot)$.

Proof. See Appendix A.1.

By OLS estimation of the cointegrating vector component $\boldsymbol{\beta}_k$ in (2.13) we have that

$$\begin{aligned} n^{(1/2+v)}(\hat{\boldsymbol{\beta}}_{k,n} - \boldsymbol{\beta}_k) &= \left(\sum_{t=1}^{n}\hat{\mathbf{X}}_{kt,p}\hat{\mathbf{X}}'_{kt,p}\right)^{-1}\sum_{t=1}^{n}\hat{\mathbf{X}}_{kt,p}u_{t,p} \\ &= \left((1/n)\sum_{t=1}^{n}(n^{-1/2}\boldsymbol{\eta}_{kt,p})(n^{-1/2}\boldsymbol{\eta}'_{kt,p})\right)^{-1} n^{-(1-v)}\sum_{t=1}^{n}(n^{-1/2}\boldsymbol{\eta}_{kt,p})u_{t,p} \end{aligned} \tag{2.31}$$

Given that $u_{t,p} = u_t - n^{-v}\boldsymbol{\tau}'_{p,tn}\bar{\mathbf{Q}}^{-1}_{pp,n}n^{-(1-v)}\sum_{j=1}^{n}\boldsymbol{\tau}_{p,jn}u_j$, then under cointegration (with v = 1/2) we have that $n^{-1/2}\sum_{j=1}^{n}\boldsymbol{\tau}_{p,jn}u_j = O_p(1)$, and thus $u_{t,p} = u_t + O_p(n^{-1/2})$, so that this expression gives the same limiting result as before. This last expression allows to obtain the limiting distribution of the OLS estimator of $\boldsymbol{\beta}_k$ under no cointegration. In this case, making use of the result in (2.9) for $v = -1/2$ when $\alpha = 1$, we get the following

$$\hat{\boldsymbol{\beta}}_{k,n} - \boldsymbol{\beta}_k \Rightarrow \left(\int_0^1 \mathbf{B}_{k,p}(s)\mathbf{B}_{k,p}(s)'ds\right)^{-1}\int_0^1 \mathbf{B}_{k,p}(s)dJ_{u,p}(s) \tag{2.32}$$

Where $dJ_{u,p}(r) = B_{u,p}(r)dr$.

The OLS residuals in (2.8) can be used as the basis for building some simple statistics for testing the null hypothesis of cointegration against the alternative of no cointegration, given that $\hat{u}_{t,p}(k) = O_p(1)$ when $v = 1/2$, and $\hat{u}_{t,p}(k) = O_p(n^{1/2})$ when $v = -1/2$.

This difference in behavior under the null and the alternative can be exploited by searching for excessive fluctuations in the sequence of scaled partial sum of residuals $\hat{B}_{[nr],p}(k)=n^{-1/2}\sum_{t=1}^{[nr]}\hat{u}_{t,p}(k)=n^{-1/2}\hat{U}_{[nr],p}(k)$ through several global measures, such as a Cramér-von Mises (CvM) measure of fluctuation as in Shin (1994), or a Kolmogorov-Smirnov (KS) measure as in Xiao (1999), Xiao and Phillips (2002), and Wu and Xiao (2008).[17] From (2.8), the scaled partial sum of OLS residuals is given by

$$n^{-(1-v)}\hat{U}_{[nr],p}(k)=n^{-1/2+v}\hat{B}_{[nr],p}(k)=n^{-(1-v)}U_{[nr],p}-(1/n)\sum_{t=1}^{[nr]}(n^{-1/2}\boldsymbol{\eta}_{kt,p}')\hat{\boldsymbol{\Theta}}_{k,n}(v) \qquad (2.33)$$

Where $\hat{B}_{[nr],p}(k)=n^{-1/2}\hat{U}_{[nr],p}(k)$ under cointegration with v = 1/2. Given that, apart of the asymptotic behavior of $n^{-(1-v)}U_{[nr],p}$ stated in (2.9)-(2.11), the limit distribution of $n^{-(1-v)}\hat{U}_{[nr],p}(k)$ is mainly determined by that of $\hat{\Theta}_{k,n}(v)=n^{1/2+v}(\hat{\boldsymbol{\beta}}_{k,n}-\boldsymbol{\beta}_k)$, revealing its dependence on the same nuisance parameters that before. Thus, under cointegration we get

$$\hat{B}_{[nr],p}(k)\Rightarrow V_{u,p}(r)-\int_0^r\mathbf{B}_{k,p}(s)'ds\left(\int_0^r\mathbf{B}_{k,p}(s)\mathbf{B}_{k,p}(s)'ds\right)^{-1}\left(\int_0^1\mathbf{B}_{k,p}(s)dV_{u,p}(s)+\boldsymbol{\Delta}_{ku}\right) \qquad (2.34)$$

Which can also be written as

$$\hat{B}_{[nr],p}(k)\Rightarrow\omega_u w_{uk,p}(r) \qquad (2.35)$$

With

$$w_{uk,p}(r)=w_{u,p}(r)-\int_0^r\mathbf{W}_{k,p}'(s)ds\left(\int_0^1\mathbf{W}_{k,p}(s)\mathbf{W}_{k,p}'(s)ds\right)^{-1}\left(\int_0^1\mathbf{W}_{k,p}(s)dw_{u,p}(s)+\boldsymbol{\Omega}_{kk}^{-1/2}\boldsymbol{\Delta}_{ku}\right) \qquad (2.36)$$

Where $Var[w_{u,p}(r)]=b_p(r)$, and $Var[\mathbf{W}_{k,p}(r)]=v_p(r)\mathbf{I}_{k,k}$, with $b_p(r)$ and $v_p(r)$ defined above. In any case, this limit null distribution depends on ω_u^2, p and k, and only in the case of strictly exogeneous regressors and serially uncorrelated error correction terms in the cointegrating equation it is free of any remaining nuisance parameter. Thus, the main difficulty comes from the dependency on $\boldsymbol{\Omega}_{kk}$ and $\boldsymbol{\Delta}_{ku}$ that cannot be removed trough a scaling transformation. On the other hand, under no cointegration, we have that $\hat{B}_{[nr],p}(k)=O_p(n)$, which determines that

$$\begin{aligned}n^{-3/2}\hat{U}_{[nr],p}(k)&=n^{-1}\hat{B}_{[nr],p}(k)\\&\Rightarrow J_{u,p}(r)-\int_0^r\mathbf{B}_{k,p}(s)'ds\left(\int_0^1\mathbf{B}_{k,p}(s)\mathbf{B}_{k,p}(s)'ds\right)^{-1}\int_0^1\mathbf{B}_{k,p}(s)B_{u,p}(s)ds\end{aligned} \qquad (2.37)$$

[17] The test statistic proposed by Shin (1994) is the generalization of the KPSS statistic for the null of stationarity by Kwiatkowski et.al. (1992), while the test statistics considered in Xiao (1999), Xiao and Phillips (2002), and Wu and Xiao (2008) are the generalizations of the KS test statistic formulated by Xiao (2001) to the cointegrating framework, which can also be interpreted as a CUSUM-type test statistic.

With $J_{u,p}(r) = \int_0^r B_{u,p}(s)ds$ (see equation (2.9).

This random limit, using the factorization $B_{u,p}(r) = \omega_u W_{u,p}(r)$ with $B_{u,p}(r)$ defined in (2.11), can also be factorized as $\omega_u R_{p,k}(r)$, with $R_{p,k}(r)$ given by

$$R_{p,k}(r) = \int_0^r W_{u,p}(s)ds - \int_0^r \mathbf{W}_{k,p}(s)'ds \left(\int_0^1 \mathbf{W}_{k,p}(s)\mathbf{W}_{k,p}(s)'ds\right)^{-1} \int_0^1 \mathbf{W}_{k,p}(s)W_{u,p}(s)ds \quad (2.38)$$

And where $Var[W_{u,p}(r)] = v_p(r)$. Existing statistics for testing the null hypothesis of cointegration against the alternative of no cointegration that make use of the stochastic properties of $\hat{B}_{[nr],p}(k)$ consider different global measures of what can be considered excessive fluctuation not compatible with the assumption of a stable long-run relationship among the model variables. Thus, the CvM-type test by Shin (1994) is based on a global measure of fluctuation given by $S_{n,p}(k) = (1/n)\sum_{t=1}^{n}(\hat{B}_{t,p}(k))^2$, while that the KS-type test statistic proposed by Wu and Xiao (2008) is based on the recursive centered measure of maximum fluctuation $R_{n,p}(k) = \max_{t=1,\ldots,n} |\hat{B}_{t,p}(k) - (t/n)\hat{B}_{n,p}(k)|$. Xiao (1999), and Xiao and Phillips (2002) consider a no centered version of this test statistic given by $CS_{n,p}(k) = \max_{t=1,\ldots,n} |\hat{B}_{t,p}(k)|$, which is the same as $R_{n,p}(k)$ when based on OLS residuals and the deterministic component contains a constant term. The main problem with this approach is that, unless corrected, the null distribution of all these test statistics are plagued of nuisance parameters due to endogeneity of regressors and the serial correlation in the error terms that cannot be removed by simple scaling methods. There exist some different methods, which are known as asymptotically efficient estimation methods, to remove these parameters and that differ in the treatment of each source of bias. Among the existing estimation methods, the three most commonly used are the Dynamic OLS (DOLS) estimator proposed by Phillips and Loretan (1991), Saikkonen (1991) and Stock and Watson (1993), the Canonical Cointegrating Regression (CCR) estimator by Park (1992), and the Fully-Modified OLS (FM-OLS) estimator by Phillips and Hansen (1990). These three estimators are asymptotically equivalent and, as was proved by Saikkonen (1991), efficient. The corrected test statistic proposed by Shin (1994) makes use of the DOLS residuals, while that the test statistics considered in Xiao (1999), Xiao and Phillips (2002) and Wu and Xiao (2008) are based on FM-OLS residuals. To our knowledge, there is no similar test statistics based on the residuals from the CCR estimation method. For a recent review and comparison of these three alternative estimation methods see, e.g., Kurozumi and Hayakawa (2009), and the references therein. Phillips (1995) and Phillips and Chang (1995) have considered the usefulness of the FM-OLS estimation method in a wide variety of situations relating the stochastic trend component of the set of regressors. This method was originally designed to estimate cointegrating relations directly by modifying standard OLS estimator with semi-nonparametric corrections that take account of endogeneity and serial correlation, with the main appeal that one can use the FM corrections to determine how important these effects are in an empirical application.

As indicated in Phillips (1995), in cases where there are major differences with OLS, the sources of such differences could be easily located and this could also helps to provide additional information about important features of the data.

Under the assumption that the long-run covariance matrix of $\xi_t = (u_t, \varepsilon'_{k,t})'$, Ω, is known, the FM-OLS estimator of the cointegrating regression model (2.11) is given by

$$\begin{pmatrix} \hat{\alpha}^{+}_{p,n} \\ \hat{\beta}^{+}_{k,n} \end{pmatrix} = \left[\sum_{t=1}^{n} \begin{pmatrix} \tau_{p,t} \\ \mathbf{X}_{k,t} \end{pmatrix} (\tau'_{p,t}, \mathbf{X}'_{k,t}) \right]^{-1} \left\{ \sum_{t=1}^{n} \begin{pmatrix} \tau_{p,t} \\ \mathbf{X}_{k,t} \end{pmatrix} Y_t^{+} - \begin{pmatrix} \mathbf{0}_{p+1} \\ n\Delta^{+}_{ku} \end{pmatrix} \right\} \tag{2.39}$$

With $\Delta^{+}_{ku} = \Delta_{ku} - \Delta_{kk}\gamma_k$, and the transformed observations of the dependent variable Y_t^{+} are usually given by $Y_t^{+} = Y_t - \Delta\mathbf{X}'_{k,t}\gamma_k$, where $\gamma_k = \Omega_{kk}^{-1}\omega_{ku}$ and the first difference of the regressors can be decomposed as $\mathbf{Z}_{k,t} = \Delta\mathbf{X}_{k,t} = \mathbf{A}_{k,p}\Delta\tau_{p,t} + \varepsilon_{k,t}$.[18]It is evident that $\mathbf{Z}_{k,t} = \varepsilon_{k,t}$ when $p = 0$, but in any other case we obtain the following decomposition

$$\mathbf{Z}_{k,t} = \Phi_{k,p-1}\tau_{p-1,t} + \varepsilon_{k,t} = (\Phi_{k,p-1} : \mathbf{0}_k)\begin{pmatrix} \tau_{p-1,t} \\ \tau_{p,t} \end{pmatrix} + \varepsilon_{k,t} = \Phi_{k,p}\Gamma^{-1}_{p,n}\tau_{p,tn} + \varepsilon_{k,t} \tag{2.40}$$

Where the matrix of trend coefficients $\Phi_{k,p-1}$ is given by a linear combination of the elements in $\mathbf{A}_{k,p}$. The following result establish the relation between this unfeasible version of the FM-OLS estimator and the OLS estimator of α_p and β_k in (2.11).

Proposition 2.2. *Given (2.1)-(2.25), and the FM-OLS estimator of (2.5) in (2.39), then we have that*

$$(a)\begin{pmatrix} \Gamma^{-1}_{p,n}[\hat{\alpha}^{+}_{p,n} + \mathbf{A}'_{k,p}\hat{\beta}^{+}_{k,n}] \\ \sqrt{n}\hat{\beta}^{+}_{k,n} \end{pmatrix} = \begin{pmatrix} \Gamma^{-1}_{p,n}[\hat{\alpha}_{p,n} + \mathbf{A}'_{k,p}\hat{\beta}_{k,n} - \Phi'_{k,p}\gamma_k] \\ \sqrt{n}\hat{\beta}_{k,n} \end{pmatrix}$$
$$- \begin{pmatrix} \mathbf{M}^{-1}_{pp,n}\left\{ \sum_{t=1}^{n} \tau_{p,tn}\varepsilon'_{k,t} - \mathbf{Q}_{pk,n}\mathbf{Q}^{-1}_{kk,n}\sum_{t=1}^{n}\eta_{k,tn}\varepsilon'_{k,t} \right\}\gamma_k - \sqrt{n}\mathbf{M}^{-1}_{pp,n}\mathbf{Q}_{pk,n}\mathbf{Q}^{-1}_{kk,n}\Delta^{+}_{ku} \\ \mathbf{M}^{-1}_{kk,n}\left\{ \sum_{t=1}^{n}\eta_{k,tn}\varepsilon'_{k,t} - \mathbf{Q}'_{pk,n}\mathbf{Q}^{-1}_{pp,n}\sum_{t=1}^{n}\tau_{p,tn}\varepsilon'_{k,t} \right\}\gamma_k + \sqrt{n}\mathbf{M}^{-1}_{kk,n}\Delta^{+}_{ku} \end{pmatrix} \tag{2.41}$$

With FM-OLS residuals, such that

$$(b)\hat{u}^{+}_{t,p}(k) = \hat{u}_{t,p}(k) - \varepsilon'_{kt,p}\gamma_k + (1/\sqrt{n})(n^{-1/2}\eta'_{kt,p})\bar{\mathbf{M}}^{-1}_{kk,n}\left((1/\sqrt{n})\sum_{t=1}^{n}(n^{-1/2}\eta_{kt,p})\varepsilon'_{kt,p}\cdot\gamma_k + \Delta^{+}_{ku} \right) \tag{2.42}$$

Where $\mathbf{M}_{pp,n} = \mathbf{Q}_{pp,n} - \mathbf{Q}_{pk,n}\mathbf{Q}^{-1}_{kk,n}\mathbf{Q}'_{pk,n}$, $\mathbf{Q}_{pk,n} = \sum_{t=1}^{n}\tau_{p,tn}\eta'_{k,tn}$, $\mathbf{Q}_{kk,n} = \sum_{t=1}^{n}\eta_{k,tn}\eta'_{k,tn}$, *and* $\mathbf{M}_{kk,n} = \sum_{t=1}^{n}(n^{-1/2}\eta_{kt,p})(n^{-1/2}\eta'_{kt,p})$ *in (2.24), and where* $\bar{\mathbf{M}}_{kk,n} = (1/n)\mathbf{M}_{kk,n}$ *with* $\varepsilon_{kt,p} = \varepsilon_{k,t} - (1/\sqrt{n})[(1/\sqrt{n})\sum_{j=1}^{n}\varepsilon_{k,j}\tau'_{p,jn}]\bar{\mathbf{Q}}^{-1}_{pp,n}\tau_{p,tn}$ *in (2.42).*

[18] It can be shown that the correction term for Y_t is associated with the correction for the endogeneity bias while Δ^{+}_{ku} eliminates the non-centrality bias.

Proof. See Appendix A.2.,Remark 2.1 The results in (2.41)-(2.42) clearly show the additional bias arising in the estimation of the trend parameters, and also allows to check how the FM-OLS estimates acts to correct for the OLS estimation results. The expression in (2.41) can also be rewritten as

$$\mathbf{W}_n'\begin{pmatrix}(\hat{\boldsymbol{\alpha}}_{p,n}^{+}-\boldsymbol{\alpha}_p)+\boldsymbol{\Phi}_{k,p}'\boldsymbol{\gamma}_k\\ \hat{\boldsymbol{\beta}}_{k,n}^{+}-\boldsymbol{\beta}_k\end{pmatrix}=\mathbf{W}_n'\begin{pmatrix}\hat{\boldsymbol{\alpha}}_{p,n}-\boldsymbol{\alpha}_p\\ \hat{\boldsymbol{\beta}}_{k,n}-\boldsymbol{\beta}_k\end{pmatrix}$$
$$-(1/\sqrt{n})\begin{pmatrix}\bar{\mathbf{M}}_{pp,n}^{-1}\left\{\bar{\mathbf{D}}_{pk,n}-\bar{\mathbf{Q}}_{pk,n}\bar{\mathbf{Q}}_{kk,n}^{-1}(1/\sqrt{n})\sum_{t=1}^{n}\boldsymbol{\eta}_{k,tn}\boldsymbol{\varepsilon}_{k,t}'\right\}\boldsymbol{\gamma}_k-\bar{\mathbf{M}}_{pp,n}^{-1}\bar{\mathbf{Q}}_{pk,n}\bar{\mathbf{Q}}_{kk,n}^{-1}\boldsymbol{\Delta}_{ku}^{+}\\ \bar{\mathbf{M}}_{kk,n}^{-1}\left\{(1/\sqrt{n})\sum_{t=1}^{n}\boldsymbol{\eta}_{k,tn}\boldsymbol{\varepsilon}_{k,t}'-\bar{\mathbf{Q}}_{pk,n}'\bar{\mathbf{Q}}_{pp,n}^{-1}\bar{\mathbf{D}}_{pk,n}\right\}\boldsymbol{\gamma}_k+\bar{\mathbf{M}}_{kk,n}^{-1}\boldsymbol{\Delta}_{ku}^{+}\end{pmatrix} \quad (2.43)$$

Where the scaled matrices appearing in the last term are given by $\bar{\mathbf{M}}_{pp,n}=(1/n)\mathbf{M}_{pp,n}$, $\bar{\mathbf{D}}_{pk,n}=(1/\sqrt{n})\mathbf{D}_{pk,n}$, with $\mathbf{D}_{pk,n}=\sum_{t=1}^{n}\boldsymbol{\tau}_{p,tn}\boldsymbol{\varepsilon}_{k,t}'$, $\bar{\mathbf{Q}}_{pk,n}=(1/n)\mathbf{Q}_{pk,n}$, and $\bar{\mathbf{Q}}_{kk,n}=(1/n)\mathbf{Q}_{kk,n}$.

All of these matrices have finite stochastic limits, determining the way as FM-OLS estimation corrects for the second-order bias arising in the OLS estimation. An alternative representation to these two is obtained when writing

$$Y_t^{+}=u_t-\boldsymbol{\varepsilon}_{k,t}'\boldsymbol{\gamma}_k+(\boldsymbol{\tau}_{p,tn}',n^{-1/2}\boldsymbol{\eta}_{k,t}')\mathbf{W}_n'\begin{pmatrix}\boldsymbol{\alpha}_p-\boldsymbol{\Phi}_{k,p}'\boldsymbol{\gamma}_k\\ \boldsymbol{\beta}_k\end{pmatrix} \quad (2.44)$$

So that (2.39) can be alternatively expressed as

$$n^{v}\mathbf{W}_n'\begin{pmatrix}\hat{\boldsymbol{\alpha}}_{p,n}^{+}\\ \hat{\boldsymbol{\beta}}_{k,n}^{+}\end{pmatrix}=n^{v}\mathbf{W}_n'\begin{pmatrix}\boldsymbol{\alpha}_p-\boldsymbol{\Phi}_{k,p}'\boldsymbol{\gamma}_k\\ \boldsymbol{\beta}_k\end{pmatrix}+\left((1/n)\sum_{t=1}^{n}\begin{pmatrix}\boldsymbol{\tau}_{p,tn}\\ n^{-1/2}\boldsymbol{\eta}_{k,t}\end{pmatrix}(\boldsymbol{\tau}_{p,tn}',n^{-1/2}\boldsymbol{\eta}_{k,t}')\right)^{-1}$$
$$\times\left\{n^{-(1-v)}\sum_{t=1}^{n}\begin{pmatrix}\boldsymbol{\tau}_{p,tn}\\ n^{-1/2}\boldsymbol{\eta}_{k,t}\end{pmatrix}(u_t-\boldsymbol{\varepsilon}_{k,t}'\boldsymbol{\gamma}_k)-n^{v}\mathbf{W}_n^{-1}\begin{pmatrix}\mathbf{0}_{p+1}\\ \boldsymbol{\Delta}_{ku}^{+}\end{pmatrix}\right\} \quad (2.45)$$

So that

$$n^{v}\mathbf{W}_n'\begin{pmatrix}(\hat{\boldsymbol{\alpha}}_{p,n}^{+}-\boldsymbol{\alpha}_p)+\boldsymbol{\Phi}_{k,p}'\boldsymbol{\gamma}_k\\ \hat{\boldsymbol{\beta}}_{k,n}^{+}-\boldsymbol{\beta}_k\end{pmatrix}=\left((1/n)\sum_{t=1}^{n}\begin{pmatrix}\boldsymbol{\tau}_{p,tn}\\ n^{-1/2}\boldsymbol{\eta}_{k,t}\end{pmatrix}(\boldsymbol{\tau}_{p,tn}',n^{-1/2}\boldsymbol{\eta}_{k,t}')\right)^{-1}$$
$$\times\left\{n^{-(1-v)}\sum_{t=1}^{n}\begin{pmatrix}\boldsymbol{\tau}_{p,tn}\\ n^{-1/2}\boldsymbol{\eta}_{k,t}\end{pmatrix}(u_t-\boldsymbol{\varepsilon}_{k,t}'\boldsymbol{\gamma}_k)-n^{-1/2+v}\begin{pmatrix}\mathbf{0}_{p+1}\\ \boldsymbol{\Delta}_{ku}^{+}\end{pmatrix}\right\} \quad (2.46)$$

Where $n^{-(1-v)}\sum_{t=1}^{n}(u_t-\boldsymbol{\varepsilon}_{k,t}'\boldsymbol{\gamma}_k)\Rightarrow B_{u.k}(r)$, $n^{-(1-v)}\sum_{t=1}^{n}\boldsymbol{\tau}_{p,tn}(u_t-\boldsymbol{\varepsilon}_{k,t}'\boldsymbol{\gamma}_k)\Rightarrow\int_0^1\boldsymbol{\tau}_p(r)dB_{u.k}(r)$, and also $n^{-(1-v)}\sum_{t=1}^{n}(n^{-1/2}\boldsymbol{\eta}_{k,t})(u_t-\boldsymbol{\varepsilon}_{k,t}'\boldsymbol{\gamma}_k)\Rightarrow\int_0^1\mathbf{B}_k(r)dB_{u.k}(r)+(\boldsymbol{\Delta}_{ku}-\boldsymbol{\Delta}_{kk}\boldsymbol{\gamma}_k)$ under the assumption of cointegration (with v = 1/2), which determines the desired result of a limiting distribution free of nuisance parameters other than the conditional long-run variance $\omega_{u.k}^2$ trough the detrended Brownian process $B_{u.k}(r)$ given in (2.25).

Also, given the partial sum process of FM-OLS residuals, $\hat{U}^{+}_{[nr],p}(k)=\sum_{t=1}^{[nr]}\hat{u}^{+}_{t,p}(k)$, the result in (2.42) allows to obtain the following decomposition of the scaled partial sum process of these residuals

$$\begin{aligned} n^{-(1-v)}\hat{U}^{+}_{[nr],p}(k) &= n^{-(1-v)}\hat{U}_{[nr],p}(k)-n^{-1/2+v}\boldsymbol{\gamma}_k' n^{-1/2}\sum_{t=1}^{[nr]}\boldsymbol{\varepsilon}_{kt,p} \\ &+n^{-1/2+v}(1/n)\sum_{t=1}^{[nr]}(n^{-1/2}\boldsymbol{\eta}_{kt,p}')\bar{\mathbf{M}}_{kk,n}^{-1}\left((1/\sqrt{n})\sum_{t=1}^{n}(n^{-1/2}\boldsymbol{\eta}_{kt,p})\boldsymbol{\varepsilon}_{kt,p}'\cdot\boldsymbol{\gamma}_k+\boldsymbol{\Delta}_{ku}^{+}\right) \end{aligned} \quad (2.47)$$

Where $n^{-1/2}\sum_{t=1}^{[nr]}\boldsymbol{\varepsilon}_{kt,p}\Rightarrow\mathbf{V}_{k,p}(r)$, $(1/\sqrt{n})\sum_{t=1}^{n}(n^{-1/2}\boldsymbol{\eta}_{kt,p})\boldsymbol{\varepsilon}_{kt,p}'\Rightarrow\int_0^1\mathbf{B}_{k,p}(s)d\mathbf{V}_{k,p}(s)+\boldsymbol{\Delta}_{kk}$, with $\mathbf{V}_{k,p}(r)$ defined in (2.43). Then, taking all these results together with the weak limit of $\hat{B}^{+}_{[nr],p}(k)=n^{-1/2}\hat{U}^{+}_{[nr],p}(k)$ under cointegration, we get

$$n^{-1/2}\hat{U}^{+}_{[nr],p}(k)\Rightarrow V_{u.k,p}(r)-\int_0^r\mathbf{B}_{k,p}(s)'ds\left(\int_0^1\mathbf{B}_{k,p}(s)\mathbf{B}_{k,p}(s)'ds\right)^{-1}\int_0^1\mathbf{B}_{k,p}(s)dV_{u.k,p}(s) \quad (2.48)$$

With $V_{u.k,p}(s)$ defined in (2.43), that allows for valid (no standard) inference given that these residuals are exactly invariant to the deterministic trend components of the integrated regressors and also to the nuisance parameters arising when using OLS estimates. Under no cointegration, when $v = -1/2$, it is immediate verify that the limiting distribution of the FM-OLS estimates and residuals coincide with that of OLS estimates and residuals, as in (2.32) and (2.37).Remark 2.2. Using (2.40), with $\mathbf{Z}_{k,t}=\Delta\mathbf{X}_{k,t}=\boldsymbol{\Phi}_{k,p}\boldsymbol{\tau}_{p,t}+\boldsymbol{\varepsilon}_{k,t}$, then by OLS detrending of $\Delta\mathbf{X}_{k,t}$ we have $\hat{\mathbf{Z}}_{kt,p}=\boldsymbol{\varepsilon}_{kt,p}$ with $\boldsymbol{\varepsilon}_{kt,p}$ defined in Proposition 2.2(b) above. If we define now Y_t^{+} as $Y_t^{+}=Y_t-\hat{\mathbf{Z}}_{kt,p}'\boldsymbol{\gamma}_k=Y_t-\boldsymbol{\varepsilon}_{kt,p}'\boldsymbol{\gamma}_k$, as indicated by Hansen (1992) (page 93), then the FM-OLS estimator of $\boldsymbol{\alpha}_p$ and $\boldsymbol{\beta}_k$ is now given by

$$\begin{pmatrix}\hat{\boldsymbol{\alpha}}^{+}_{p,n}\\ \hat{\boldsymbol{\beta}}^{+}_{k,n}\end{pmatrix}=\begin{pmatrix}\hat{\boldsymbol{\alpha}}_{p,n}\\ \hat{\boldsymbol{\beta}}_{k,n}\end{pmatrix}-(\mathbf{W}_n')^{-1}\begin{pmatrix}\mathbf{Q}_{pp,n} & \mathbf{Q}_{pk,n}\\ \mathbf{Q}_{pk,n}' & \mathbf{Q}_{kk,n}\end{pmatrix}^{-1}\left\{\sum_{t=1}^{n}\begin{pmatrix}\boldsymbol{\tau}_{p,tn}\\ \boldsymbol{\eta}_{k,tn}\end{pmatrix}\boldsymbol{\varepsilon}_{kt,p}'\cdot\boldsymbol{\gamma}_k+\mathbf{W}_n^{-1}\begin{pmatrix}\mathbf{0}_{p+1}\\ n\boldsymbol{\Delta}_{ku}^{+}\end{pmatrix}\right\} \quad (2.49)$$

Which gives exactly the same expressions as before for $\hat{\boldsymbol{\beta}}^{+}_{k,n}$, but not for the estimator of the trend parameters in $\hat{\boldsymbol{\alpha}}^{+}_{p,n}$, as will be stated in the next result. The FM-OLS estimator in (2.39), as well as all the results in (2.41) and (2.42), is not feasible since it is defined in terms of the unknown quantities $\boldsymbol{\gamma}_k$ and $\boldsymbol{\Delta}_{ku}^{+}=\boldsymbol{\Delta}_{ku}-\boldsymbol{\Delta}_{kk}\boldsymbol{\gamma}_k$. The feasible version is obtained by replacing these elements by nonparametric kernel estimates of the components of the long-run covariance matrix $\boldsymbol{\Omega}$ based on the OLS residuals in (2.8) and the stationary stochastic component of the regressors, that must be consistent under the assumption of cointegration, and requires a proper choice of the bandwidth to ensure the asymptotic correction for serial correlation and endogeneity.

Although in the work by Phillips and Hansen (1990) they consider the effects of deterministically trending integrated regressors on the results of the estimation of the cointegrating equation, making use of previous results by Phillips and Durlauf (1986) and Park and Phillips (1988), the estimation of $\mathbf{\Omega}$ is based on the sample serial covariance matrices of the sequence $\boldsymbol{\xi}_{t,p}(k)=(\hat{u}_{t,p}(k),\mathbf{Z}'_{k,t})'$, without further considerations about the consequences of any remaining deterministic component in the series of first differences for the observations of the regressors. Also, the papers by Xiao (1999) and Xiao and Phillips (2002) make use of this sequence, both to implement feasible version of the FM-OLS estimation method and for the estimation of the conditional long-run variance $\omega^2_{u.k}=\omega^2_u-\boldsymbol{\omega}_{u,k}\mathbf{\Omega}^{-1}_{kk}\boldsymbol{\omega}_{k,u}$. Next proposition establish the more relevant results concerning the effects on the feasible FM-OLS estimates when we explicitly take into account these circumstances.

Proposition 2.3. *Given (2.1)-(2.5), then we have that*

(a) *When using the sequence* $\boldsymbol{\xi}_{t,p}(k)=(\hat{u}_{t,p}(k),\mathbf{Z}'_{k,t})'$ *for computation of the nonparametric kernel estimator of the long-run variance* $\mathbf{\Omega}$*, we have that*

$$\hat{\mathbf{\Omega}}_n(m_n)=\mathbf{\Omega}_n(m_n)+m_n(\mathbf{C}_n+\mathbf{C}'_n+\mathbf{F}_n)\bar{w}_n(m_n)+O_p(m_n/n) \tag{2.50}$$

Where $\bar{w}_n(m_n)=m_n^{-1}\sum_{h=-(n-1)}^{n-1}w(h/m_n)\to\lambda$*, and* $\mathbf{\Omega}_n(m_n)\to^p\mathbf{\Omega}$*, which is the kernel estimator based on the set of sample serial covariance matrices of the sequence* $\boldsymbol{\upsilon}_{t,p}(k)=(\hat{u}_{t,p}(k),\boldsymbol{\varepsilon}'_{k,t})'$*, that is, it is given by* $\mathbf{\Omega}_n(m_n)=\sum_{h=-(n-1)}^{n-1}w(h/m_n)\mathbf{\Sigma}_n(h)$*, with* $\mathbf{\Sigma}_n(h)=(1/n)\sum_{t=h+1}^{n}\boldsymbol{\upsilon}_{t,p}(k)\boldsymbol{\upsilon}_{t-h,p}(k)$*. Also, matrices* $\mathbf{C}_n$ *and* $\mathbf{F}_n$ *are given by*

$$\mathbf{C}_n=\begin{pmatrix}0 & \mathbf{0}'_k\\ \mathbf{0}_k & \mathbf{C}_{kk,n}\end{pmatrix},\ and\ \mathbf{F}_n=\begin{pmatrix}0 & \mathbf{0}'_k\\ \mathbf{0}_k & \mathbf{F}_{kk,n}\end{pmatrix} \tag{2.51}$$

With $\mathbf{C}_{kk,n}=\bar{\mathbf{D}}'_{pk,n}(n^{-1/2}\mathbf{\Gamma}^{-1}_{p,n})\mathbf{\Phi}'_{kp}$, $\mathbf{F}_{kk,n}=\mathbf{\Phi}_{kp}\mathbf{\Gamma}^{-1}_{p,n}\bar{\mathbf{Q}}_{pp,n}\mathbf{\Gamma}^{-1}_{p,n}\mathbf{\Phi}'_{kp}$*. Then, the FM-OLS estimator is given by*

$$\begin{aligned}\hat{\mathbf{\Theta}}^+_n(v)&=\begin{pmatrix}\hat{\mathbf{\Theta}}^+_{p,n}(v)\\ \hat{\mathbf{\Theta}}^+_{k,n}(v)\end{pmatrix}=n^v\mathbf{W}'_n\begin{pmatrix}(\hat{\boldsymbol{\alpha}}^+_{p,n}-\boldsymbol{\alpha}_p)+\mathbf{\Phi}'_{k,p}\hat{\boldsymbol{\gamma}}_{k,n}(m_n)\\ \hat{\boldsymbol{\beta}}^+_{k,n}-\boldsymbol{\beta}_k\end{pmatrix}\\ &=\left(n^{-1}\sum_{t=1}^{n}\begin{pmatrix}\boldsymbol{\tau}_{p,tn}\\ n^{-1/2}\boldsymbol{\eta}_{k,t}\end{pmatrix}(\boldsymbol{\tau}'_{p,tn},n^{-1/2}\boldsymbol{\eta}'_{k,t})\right)^{-1}\left\{n^{-(1-v)}\sum_{t=1}^{n}\begin{pmatrix}\boldsymbol{\tau}_{p,tn}\\ n^{-1/2}\boldsymbol{\eta}_{k,t}\end{pmatrix}\hat{z}_t-n^{-1/2+v}\begin{pmatrix}\mathbf{0}_{p+1}\\ \hat{\mathbf{\Delta}}^+_{ku,n}(m_n)\end{pmatrix}\right\}\end{aligned} \tag{2.52}$$

Where $\hat{z}_t=u_t-\hat{\boldsymbol{\gamma}}'_{k,n}(m_n)\boldsymbol{\varepsilon}_{k,t}$*, with* $\hat{\boldsymbol{\gamma}}_{k,n}(m_n)=\hat{\mathbf{\Omega}}^{-1}_{kk,n}(m_n)\hat{\boldsymbol{\omega}}_{ku,n}(m_n)$.

(b) When using the sequence $\hat{\xi}_{t,p}(k)=(\hat{u}_{t,p}(k),\hat{\mathbf{Z}}'_{kt,p})'$ *for computation of the nonparametric kernel estimator of the long-run variance* $\mathbf{\Omega}$*, we have that* $\hat{\mathbf{\Omega}}_n(m_n)=\mathbf{\Omega}_n(m_n)+O_p(m_n/n)\to^p \mathbf{\Omega}$*, where the limiting distribution of the FM-OLS estimator under cointegration is now given by*

$$\hat{\mathbf{\Theta}}_n^+(v)=\begin{pmatrix}\hat{\mathbf{\Theta}}_{p,n}^+(v)\\ \hat{\mathbf{\Theta}}_{k,n}^+(v)\end{pmatrix}=n^v\mathbf{W}'_n\begin{pmatrix}\hat{\boldsymbol{\alpha}}_{p,n}^+-\boldsymbol{\alpha}_p\\ \hat{\boldsymbol{\beta}}_{k,n}^+-\boldsymbol{\beta}_k\end{pmatrix}$$
$$\Rightarrow\begin{pmatrix}\mathbf{Q}_{pp}^{-1}\mathbf{D}_{pk}\boldsymbol{\gamma}_k\\ \mathbf{0}_{k,p+1}\end{pmatrix}+\left(\int_0^1\begin{pmatrix}\boldsymbol{\tau}_p(s)\\ \mathbf{B}_k(s)\end{pmatrix}(\boldsymbol{\tau}_p(s)',\mathbf{B}_k(s)')ds\right)^{-1}\int_0^1\begin{pmatrix}\boldsymbol{\tau}_p(s)\\ \mathbf{B}_k(s)\end{pmatrix}dB_{u.k}(s) \qquad (2.53)$$

With $\mathbf{D}_{pk}=\int_0^1\boldsymbol{\tau}_p(s)d\mathbf{B}_k(s)'$ *the weak limit of* $\bar{\mathbf{D}}_{pk,n}=n^{-1/2}\sum_{t=1}^n\boldsymbol{\tau}_{p,tn}\boldsymbol{\varepsilon}'_{k,t}$.

Proof. See Appendix A.3.

Remark 2.3 The results in part (a) only applies for $p \geq 1$, with $\mathbf{C}_n=\mathbf{F}_n=\mathbf{0}_{k+1,k+1}$ when $p = 0$.[19] Also, it is assumed that the limit λ of $\bar{w}_n(m_n)$ as $n\to\infty$ is finite. For example, for the Bartlett kernel, which is symmetric with weighting function $w(h/m_n)=1-hm_n^{-1}$ for $|h| \leq m_n-1$ and zero otherwise, we have that $\bar{w}_n(m_n)=1$. Also, given the matrices $\mathbf{C}_n$ and $\mathbf{F}_n$ in (2.51) it is evident that in this case we have $\hat{\omega}_{u,n}^2(m_n)=\omega_{u,n}^2(m_n)+o_p(n^{-1/2})$, and $\hat{\boldsymbol{\omega}}_{ku,n}(m_n)=\boldsymbol{\omega}_{ku,n}(m_n)+o_p(n^{-1/2})$ under suitable choice of the bandwidth parameter, particularly when it is imposed the usual condition $m_n=o_p(n^{1/2})$.

The remaining term, the estimator of the long-run variance of the stochastic stationary component of the integrated regressors, may contain a serious bias component as it admits the following decomposition:

$$\hat{\mathbf{\Omega}}_{kk,n}(m_n)=\mathbf{\Omega}_{kk,n}(m_n)+\bar{w}_n(m_n)[m_n(\mathbf{C}_{kk,n}+\mathbf{C}'_{kk,n})+m_n\mathbf{F}_{kk,n})]+O_p(m_n/n) \qquad (2.54)$$

Also affecting the estimators of $\boldsymbol{\gamma}_k$ and $\boldsymbol{\Delta}_{kk}$, and hence of $\boldsymbol{\Delta}_{ku}^+$. When $p = 1$, we then have

$$\mathbf{C}_n=(1/\sqrt{n})\begin{pmatrix}0 & \mathbf{0}'_k\\ \mathbf{0}_k & \mathbf{B}_{k,n}(1)\boldsymbol{\alpha}'_{k0}\end{pmatrix} \qquad (2.55)$$

[19] Similarly, in the case of a known long-run covariance matrix $\mathbf{\Omega}$, the CCR estimator proposed by Park (1992) is defined as the OLS estimator between the modified dependent variable $Y_t^*=Y_t-(\hat{\boldsymbol{\beta}}'_{k,n}\boldsymbol{\Delta}_k\mathbf{\Sigma}^{-1}+(0,\boldsymbol{\gamma}'_k))\xi_{t,p}(k)$ and $(\boldsymbol{\tau}'_{p,t},\mathbf{X}'^*_{k,t})'$, with $\mathbf{X}^*_{k,t}=\mathbf{X}_{k,t}-\boldsymbol{\Delta}_k\mathbf{\Sigma}^{-1}\xi_{t,p}(k)$, $\xi_{t,p}(k)=(\hat{u}_{t,p}(k),\mathbf{Z}'_{k,t})'$, and $\boldsymbol{\Delta}_k=\sum_{j=0}^{\infty}E[\boldsymbol{\varepsilon}_{k,t}\xi'_{t-j}]=\sum_{j=0}^{\infty}E[\boldsymbol{\varepsilon}_{k,t}(u_{t-j},\boldsymbol{\varepsilon}'_{k,t-j})]$. This method uses the same principle as the FM-OLS method to eliminate the endogeneity bias, while it deals with the non-centrality parameter in a different manner, but also relies on consistent estimates of the quantities $\boldsymbol{\Delta}_k$, $\mathbf{\Sigma}$ and $\boldsymbol{\gamma}_k$ which depend on some tuning parameters. The feasible CCR estimator when $\mathbf{\Omega}$ is unknown makes use of a nonparametric kernel estimator of all the quantities involved in these transformation factors that, when it is based on the sample autocovariances of the sequence $\xi_{t,p}(k)$, might be biased or even inconsistently estimated when there is some remaining deterministic component in the observations of the first differences of the regressors.

And

$$\mathbf{F}_n = \begin{pmatrix} 0 & \mathbf{0}_k' \\ \mathbf{0}_k & \mathbf{\Phi}_{kp}\mathbf{\Phi}_{kp}' \end{pmatrix} = \mathbf{F} \tag{2.56}$$

Where $m_n(\mathbf{C}_{kk,n} + \mathbf{C}_{kk,n}') = n^{-1/2} m_n(\mathbf{B}_{k,n}(1)\boldsymbol{\alpha}_{k0}' + \boldsymbol{\alpha}_{k0}\mathbf{B}_{k,n}(1)') = o_p(1)$, while that the last term in (2.34) is $\mathbf{F}_{kk} = \mathbf{\Phi}_{kp}\mathbf{\Phi}_{kp}' = \boldsymbol{\alpha}_{k0}\boldsymbol{\alpha}_{k0}'$, so that $\hat{\mathbf{\Omega}}_{kk,n}(m_n) = \mathbf{\Omega}_{kk,n}(m_n) + O(m_n)$ which can seriously distort the estimation results, specially in relation with the estimation of the factor correcting the non-centrality bias term, $\mathbf{\Delta}_{ku}^{+}$.[20] Although rare in practical applications, the case $p > 1$ will determine that both biasing terms in (2.54) have a non-negligible effect on the estimation results. A final comment on these results has to be with the expression in (2.52) regarding the limit distribution of the estimator of the trend parameters, which that closely resembles the result in (2.46) but with $\boldsymbol{\gamma}_k$ replaced by its estimator. Finally, it must also be commented the change in the limiting distribution of the FM-OLS estimator of the trend parameters in (2.53), which represents an asymptotic bias for the estimation of these parameters.

Given the transformed values of the dependent variable and the resulting estimates of the model parameter, the FM-OLS residuals, given by $\hat{u}_{t,p}^{+}(k) = Y_t^{+} - \boldsymbol{\tau}_{p,t}'\hat{\boldsymbol{\alpha}}_{p,n}^{+} - \mathbf{X}_{k,t}'\hat{\boldsymbol{\beta}}_{k,n}^{+}$, can be written as

$$\begin{aligned} \hat{u}_{t,p}^{+}(k) &= u_t - \mathbf{v}_{k,t}'\hat{\boldsymbol{\gamma}}_{k,n}(m_n) - n^{-v}(\boldsymbol{\tau}_{p,tn}'\hat{\mathbf{\Theta}}_{p,n}^{+}(v) + n^{-1/2}\boldsymbol{\eta}_{k,t}'\hat{\mathbf{\Theta}}_{k,n}^{+}(v)) \\ &= z_t - (\mathbf{v}_{k,t}'\hat{\boldsymbol{\gamma}}_{k,n}(m_n) - \boldsymbol{\varepsilon}_{k,t}'\boldsymbol{\gamma}_k) - n^{-v}(\boldsymbol{\tau}_{p,tn}'\hat{\mathbf{\Theta}}_{p,n}^{+}(v) + n^{-1/2}\boldsymbol{\eta}_{k,t}'\hat{\mathbf{\Theta}}_{k,n}^{+}(v)) \end{aligned} \tag{2.57}$$

Where $\mathbf{v}_{k,t} = \boldsymbol{\varepsilon}_{k,t}$ when using the sequence $\boldsymbol{\xi}_{t,p}(k) = (\hat{u}_{t,p}(k), \mathbf{Z}_{k,t}')'$ for performing the FM estimation, and $\mathbf{v}_{k,t} = \hat{\mathbf{Z}}_{kt,p} = \boldsymbol{\varepsilon}_{kt,p} = \boldsymbol{\varepsilon}_{k,t} + O_p(n^{-1/2})$ when the transformation factors are based on the sequence $\hat{\boldsymbol{\xi}}_{t,p}(k) = (\hat{u}_{t,p}(k), \hat{\mathbf{Z}}_{kt,p}')'$, with $\hat{\mathbf{\Theta}}_{p,n}^{+}(v)$ and $\hat{\mathbf{\Theta}}_{k,n}^{+}(v)$ given either by (2.32) or (2.33) in each case. In both cases, under cointegration we then have $\hat{u}_{t,p}^{+}(k) = u_t - \mathbf{v}_{k,t}'\hat{\boldsymbol{\gamma}}_{k,n}(m_n) - O_p(n^{-1/2})$. With the results stated in part (b) of Proposition 2.3, this FM-OLS residuals consistently estimate the sequence $z_t = u_t - \boldsymbol{\varepsilon}_{k,t}'\boldsymbol{\gamma}_k$, which could be used to estimate the conditional long-run variance $\omega_{u.k}^2 = \omega_u^2 - \boldsymbol{\omega}_{u,k}\mathbf{\Omega}_{kk}^{-1}\boldsymbol{\omega}_{k,u}$. With the results obtained in part (a), where $\hat{u}_{t,p}^{+}(k) = z_t - \boldsymbol{\varepsilon}_{k,t}'(\hat{\boldsymbol{\gamma}}_{k,n}(m_n) - \boldsymbol{\gamma}_k) - O_p(n^{-1/2})$, then only in the case where the regressors does not contain any deterministic component, or when they contain no more than a constant term, it could be assured the consistent estimation of $\mathbf{\Omega}$ and hence of $\boldsymbol{\gamma}_k$.

---

[20] With this results, we then have that when $p = 1$, $\hat{\mathbf{\Omega}}_{kk,n}(m_n) = \mathbf{\Omega}_{kk,n}(m_n) + m_n\bar{w}_n(m_n)\boldsymbol{\alpha}_{k0}\boldsymbol{\alpha}_{k0}' + o_p(1)$, with inverse given by $\hat{\mathbf{\Omega}}_{kk,n}^{-1}(m_n) = \mathbf{\Omega}_{kk,n}^{-1}(m_n) - \dfrac{1}{m_n\bar{w}_n(m_n) + \boldsymbol{\alpha}_{k0}'\mathbf{\Omega}_{kk,n}^{-1}(m_n)\boldsymbol{\alpha}_{k0}}\mathbf{\Omega}_{kk,n}^{-1}(m_n)\boldsymbol{\alpha}_{k0}\boldsymbol{\alpha}_{k0}'\mathbf{\Omega}_{kk,n}^{-1}(m_n) + o_p(1)$ so that $\hat{\mathbf{\Omega}}_{kk,n}^{-1}(m_n) = \mathbf{\Omega}_{kk,n}^{-1}(m_n) + O_p(m_n^{-1})$, and the resulting estimator of $\boldsymbol{\gamma}_k$ is consistent.

The standard practice consists on the use of an estimator of $\omega_{u.k}^2$ based on the estimates of ω_u^2, $\mathbf{\Omega}_{kk}$, and $\mathbf{\omega}_{k,u}$ obtained from $\hat{\mathbf{\Omega}}_h(m_n)=\sum_{h=-(n-1)}^{n-1} w(h/m_n)\hat{\mathbf{\Sigma}}_n(h)$, with $\hat{\mathbf{\Sigma}}_n(h)$ the sample serial autocovariance matrix of order *h* based on either $\boldsymbol{\xi}_{t,p}(k)$ or $\hat{\boldsymbol{\xi}}_{t,p}(k)$, which make use of the OLS residuals in (2.8). Under cointegration and consistent estimation of $\boldsymbol{\gamma}_k$, then $\hat{u}_{t,p}^{+}(k)=z_t+O_p(n^{-1/2})$ which gives

$$\hat{\kappa}_n^{+}(h)=(1/n)\sum_{t=h+1}^{n}\hat{u}_{t,p}^{+}(k)\hat{u}_{t-h,p}^{+}(k)=(1/n)\sum_{t=h+1}^{n}z_t z_{t-h}+O_p(n^{-1/2})=\kappa_n^{+}(h)+O_p(n^{-1/2}) \quad (2.58)$$

Given that we can write $z_t=(1,-\boldsymbol{\gamma}_k')\boldsymbol{\xi}_t$, then we have that

$$\kappa_n^{+}(h)=(1/n)\sum_{t=h+1}^{n}z_t z_{t-h}=(1,-\boldsymbol{\gamma}_k')\left\{(1/n)\sum_{t=h+1}^{n}\boldsymbol{\xi}_t\boldsymbol{\xi}_{t-h}\right\}\begin{pmatrix}1\\-\boldsymbol{\gamma}_k\end{pmatrix}=(1,-\boldsymbol{\gamma}_k')\mathbf{G}_n(h)\begin{pmatrix}1\\-\boldsymbol{\gamma}_k\end{pmatrix} \quad (2.59)$$

And thus we have that

$$\begin{aligned}\hat{\omega}_{u.k,n}^{+2}(m_n)&=\sum_{h=-(n-1)}^{n-1}w(h/m_n)\hat{\kappa}_n^{+}(h)=\sum_{h=-(n-1)}^{n-1}w(h/m_n)\kappa_n^{+}(h)+O_p(m_n n^{-1/2})\\&=(1,-\boldsymbol{\gamma}_k')\sum_{h=-(n-1)}^{n-1}w(h/m_n)\mathbf{G}_n(h)\begin{pmatrix}1\\-\boldsymbol{\gamma}_k\end{pmatrix}+O_p(m_n n^{-1/2})\end{aligned} \quad (2.60)$$

With probability given by

$$\hat{\omega}_{u.k,n}^{+2}(m_n)\to^{p}(1,-\boldsymbol{\gamma}_k')\mathbf{\Omega}\begin{pmatrix}1\\-\boldsymbol{\gamma}_k\end{pmatrix}=\omega_{u.k}^2 \quad (2.61)$$

Under the usual requirements relative to the permitted kernel functions and bandwidth order, $m_n=o_p(n^{1/2})$, that ensure consistent estimation for this class of nonparametric long-run variance estimators.

Once established all these relevant results, one major application is to build pivotal test statistics that consistently discriminate between cointegration and no cointegration in this setup. Making use of any of the two possible consistent estimates of the conditional long-run variance $\omega_{u.k}^2$, $\hat{\omega}_{u.k,n}^2(m_n)=\hat{\omega}_{u,n}^2-\hat{\boldsymbol{\omega}}_{uk,n}\hat{\mathbf{\Omega}}_{kk,n}^{-1}\hat{\boldsymbol{\omega}}_{uk,n}'$, or $\hat{\omega}_{u.k,n}^{+2}(m_n)$ as in (2.36), then the fluctuation-based pivotal test statistics are given by

$$\hat{S}_{n,p}(k)=(n\hat{\omega}_{u.k,n}^2)^{-1}\sum_{t=1}^{n}(\hat{B}_{t,p}^{+}(k))^2 \quad (2.62)$$

$$C\hat{S}_{n,p}(k)=\max_{t=1,\dots,n}\hat{\omega}_{u.k,n}^{-1}\,|\hat{B}_{t,p}^{+}(k)| \quad (2.63)$$

And

$$\hat{R}_{n,p}(k)=\max_{t=1,\dots,n}\hat{\omega}_{u.k,n}^{-1}\,|\hat{B}_{t,p}^{+}(k)-(t/n)\hat{B}_{n,p}^{+}(k)| \quad (2.64)$$

With $\hat{R}_{n,p}(k) = C\hat{S}_{n,p}(k)$ when the specification of the deterministic component in the cointegrating equation contains at least a constant term, $p \geq 0$, so that $\hat{U}^{+}_{n,p}(k) = 0$, where $C\hat{S}_{n,p}(k)$ and $\hat{R}_{n,p}(k)$ have been proposed by Xiao (1999) and Xiao and Phillips (2002), and Wu and Xiao (2008), respectively, while that $\hat{S}_{n,p}(k)$ is the test statistic proposed by Shin (1994), but based on FM-OLS residuals. Making use of the result in (2.29), and similarly to equation (2.18), then under cointegration we have that

$$\hat{B}^{+}_{[nr],p}(k) \Rightarrow \omega_{u.k}\left\{W_{u.k,p}(r) - \int_0^r \mathbf{W}'_{k,p}(s)ds \left(\int_0^1 \mathbf{W}_{k,p}(s)\mathbf{W}'_{k,p}(s)ds\right)^{-1} \int_0^1 \mathbf{W}_{k,p}(s)dW_{u.k,p}(s)\right\} \quad (2.65)$$

So that these three test statistics has limit null distributions that are free of nuisance parameters and are consistent against the alternative hypothesis of no cointegration. Their limiting null distributions are the corresponding functional transformations of the fundamental random limit giving between brackets in (2.65). According to several simulation experiments, these test statistics seem to perform well in finite samples in terms of empirical size and power. These authors also provide the relevant critical values for performing the tests. In all the cases, the tests are right-sided thus rejecting the null hypothesis of cointegration for high values of each of these test statistics. Asymptotic critical values can be founded in the respective papers by these authors.

Both finite sample size and power of all these test statistics crucially depends on the quality of the estimation of the long-run variance $\mathbf{\Omega}$ trough the choice of the bandwidth value and also on the kernel function. Also, as has been proved above, there could be some situations where some of the components of this matrix could be estimated with bias, which could have serious effects on the resulting FM-OLS estimates of the model parameters and residuals, and hence on the properties of the test statistics based on these results.

Finally, there exist the possibility to obtain a pivotal test statistic only based on simple functionals of the OLS residuals, with limit distribution free of nuisance parameters, but which serves to test the opposite hypothesis to the other test statistics discussed above, that is the null of no cointegration against the alternative of cointegration. In this sense, it could serve as a complement, both in the case of confirmation or conflict, to the testing procedures considered. The proposed test statistic is a generalization of the variance ratio statistic, proposed by Breitung (2002) and Breitung and Taylor (2003) to testing for a fixed unit root against stationarity, in the cointegration framework[21], and it is given by

$$V\hat{R}_{n,p}(k) = \frac{(1/n)\sum_{t=1}^{n}\left((1/n)\sum_{j=1}^{t}(n^{-1/2}\hat{u}_{j,p}(k))\right)^2}{(1/n)\sum_{t=1}^{n}(n^{-1/2}\hat{u}_{t,p}(k))^2} \quad (2.66)$$

[21] Breitung (2002) developed a generalized version of the variance ratio statistic to multivariate processes to test hypothesis on the cointegration rank among a set of $m > 1$ series. In this sense, the test statistic in (2.41) differs from the one proposed by this author and it is designed to distinguish between no cointegration against cointegration with just one cointegration relationship.

That can also be written as $\hat{VR}_{n,p}(k)=(1/n)\hat{\rho}_{n,p}(k)$, in terms of the ratio of the residual variance of the partial sum of residuals to the residual variance of the residuals, with $\hat{\rho}_{n,p}(k)=(1/n)\sum_{t=1}^{n}\hat{U}_{t,p}^{2}(k)/\sum_{t=1}^{n}\hat{u}_{t,p}^{2}(k)$, where $\hat{u}_{t,p}(k)$ are the OLS residuals from the estimation of the cointegrating regression model (2.3), and $\hat{U}_{t,p}(k)=\sum_{j=1}^{t}\hat{u}_{j,p}(k)$. [22]

Next result establish the limiting distribution of the variance ratio test statistic (2.41) under no cointegration and also its behavior under cointegration, determining that the testing procedure is left-sided, rejecting the hypothesis of no cointegration for low values of the test statistic. Appendix B.1 provides tables with the quantiles of the limiting distribution under no cointegration when the cointegrating regression model does not contain any deterministic component, and when there is a constant term or a constant and a linear trend component.

Proposition 2.4. *Given (2.1)-(2.11), Assumption 2.1 and the OLS estimator of (2.5) in (2.14), then under no cointegration we have that*

$$(a)\ \hat{VR}_{n,p}(k)\Rightarrow\frac{\int_0^1\left(\int_0^r R_{p,k}(s)ds\right)^2 dr}{\int_0^1 R_{p,k}(s)^2 ds} \qquad (2.67)$$

With $R_{p,k}(r)$ *defined in (2.21), while that under cointegration we have that*

$$(b)\ \hat{\rho}_{n,p}(k)\Rightarrow\frac{\omega_u^2}{\sigma_u^2}\int_0^1 w_{uk,p}(r)^2 dr \qquad (2.68)$$

With $w_{uk,p}(r)$ *given in (2.36), so that the test is consistent, with diverging rate* $\hat{VR}_{n,p}(k)=O_p(n^{-1})$.

Proof. The result in (2.7) follows directly from the continuous mapping theorem making use of the result in (2.37), which determines that the limiting distribution is invariant to the serial correlation in the errors from the cointegrating equation and the endogeneity of the regressors. On the other hand, under cointegration, taking the OLS residuals as in (2.21), we get that $(1/\sqrt{n})\hat{U}_{[nr],p}(k)$ weakly converges to $\omega_u w_{uk,p}(r)$, as in (2.35), and $(1/n)\sum_{t=1}^{n}\hat{u}_{t,p}^{2}(k)=(1/n)\sum_{t=1}^{n}u_t^2+O_p(n^{-1/2})\to^{p}\sigma_u^2$, so that the final result in (2.68) follows again by application of the continuous mapping theorem.

The main advantage of the use of this last testing procedure is that while not requiring the choice of any tuning parameter, it is very simple to compute and the limiting distribution under no cointegration in (2.67) is free of nuisance parameters. Appendix B.1 presents some numerical results relating the use and properties of this test statistic. Particularly, Table B.1.1 presents the relevant quantiles of the null distribution under no cointegration for k = 1, ..., 5 and a variety of choices of the deterministic component, particularly when there is no deterministic component, and also in the more usual cases of a constant term (p = 0), and a constant term and a linear trend component (p = 1).

---

[22] This proposal follows the same idea as the extension of the KPSS statistic for testing stationarity against a fixed unit root to the cointegration framework made by Shin (1994). Breitung (2002) proposed a multivariate generalization of the variance ratio statistic and a semi-nonparametric test statistic for testing for the number of cointegrating relations among the components of a m-dimensional vector, with $m \geq 1$.

Also, Table B.1.2 presents the results of a simulation experiment to evaluate the finite-sample power of the testing procedure based on the test statistic $V\hat{R}_{n,p}(k)$ against near cointegration using the local-to-unity approach by Phillips (1987). These numerical results are based on samples sizes *n* = 100, 250, 500 and 750 and reflects the behavior of the limiting distribution given in (2.42) where the component $W_{u,p}(r)$ of the random element $R_{p,k}(r)$ is replaced by $W_{c,p}(r)=W_c(r)-\boldsymbol{\tau}_p'(r)\mathbf{Q}_{pp}^{-1}\int_0^1\boldsymbol{\tau}_p(s)W_c(s)ds$, with $W_c(r)$ a standard Ornstein-Uhlenbeck process. The results indicate an acceptable power, and increasing with the sample size, that confirms the consistency property stated above.

2.2 IM-OLS estimation with trending regressors

In this section we consider the new estimator of a static cointegrating regression model like (2.3) recently proposed by VW. For implementing this new esitmation method, these authors show that a simple transformation of the variables in the cointegrating regression model allows to obtain an asymptotically unbiased estimator of $\boldsymbol{\beta}_k$ with a zero mean Gaussian mixture limiting distribution, but their analysis is limited to the case where the assumed DGP is as in (2.1) with $\mathbf{A}_{k,p}=\mathbf{0}_{k,p+1}$, that is with integrated regressors without any deterministic component. Like FM-OLS, the transformation has two steps but neither one requires the estimation of any of the components of $\boldsymbol{\Omega}$, and so the choice of bandwidth and kernel is completely avoided. Thus, the first step consists in computing the (cumulate) partial sums of the variables in both sides of (2.3) which gives the so-called integrated cointegrating regression model as

$$S_t=\boldsymbol{\alpha}_p'\mathbf{S}_{p,t}+\boldsymbol{\beta}_k'\mathbf{S}_{k,t}+U_t \quad t=1,\ldots,n \tag{2.69}$$

With

$$U_t=\textstyle\sum_{j=1}^t u_j,\ S_t=\sum_{j=1}^t Y_j=\boldsymbol{\alpha}_{0,p}'\sum_{j=1}^t\boldsymbol{\tau}_{p,j}+\sum_{j=1}^t\eta_{0,j}=\boldsymbol{\alpha}_{0,p}'\mathbf{S}_{p,t}+h_{0,t},$$
$$\mathbf{S}_{p,t}=\textstyle\sum_{j=1}^t\boldsymbol{\tau}_{p,j}=\boldsymbol{\Gamma}_{p,n}^{-1}\sum_{j=1}^t\boldsymbol{\tau}_{p,jn}=\boldsymbol{\Gamma}_{p,n}^{-1}\mathbf{S}_{p,tn} \tag{2.70}$$

And

$$\mathbf{S}_{k,t}=\textstyle\sum_{j=1}^t\mathbf{X}_{k,j}=\mathbf{A}_{k,p}\mathbf{S}_{p,t}+\sum_{j=1}^t\boldsymbol{\eta}_{k,j}=\mathbf{A}_{k,p}\boldsymbol{\Gamma}_{p,n}^{-1}\mathbf{S}_{p,tn}+\mathbf{H}_{k,t} \tag{2.71}$$

Taking together (2.70) and (2.71) we then have

$$\begin{pmatrix}\mathbf{S}_{p,t}\\ \mathbf{S}_{k,t}\end{pmatrix}=\begin{pmatrix}n\boldsymbol{\Gamma}_{p,n}^{-1} & \mathbf{0}_{p+1,k}\\ n\mathbf{A}_{k,p}\boldsymbol{\Gamma}_{p,n}^{-1} & n\sqrt{n}\mathbf{I}_{k,k}\end{pmatrix}\begin{pmatrix}n^{-1}\mathbf{S}_{p,tn}\\ n^{-3/2}\mathbf{H}_{k,t}\end{pmatrix}=\mathbf{W}_{0,n}\begin{pmatrix}n^{-1}\mathbf{S}_{p,tn}\\ n^{-3/2}\mathbf{H}_{k,t}\end{pmatrix}=\mathbf{W}_{0,n}\mathbf{g}_{tn} \tag{2.72}$$

Where, for $t=[nr]$, we get

$$\mathbf{g}_{[nr]n} = \begin{pmatrix} n^{-1}\mathbf{S}_{p,[nr]n} \\ n^{-3/2}\mathbf{H}_{k,[nr]} \end{pmatrix} \Rightarrow \mathbf{g}(r) = \begin{pmatrix} \mathbf{g}_p(r) \\ \mathbf{g}_k(r) \end{pmatrix} = \begin{pmatrix} \int_0^r \boldsymbol{\tau}_p(s)ds \\ \int_0^r \mathbf{B}_k(s)ds \end{pmatrix} \tag{2.73}$$

As $n \to \infty$. Then, application of the OLS estimation method to equation (2.69) gives the so-called Integrated-OLS (I-OLS) estimator of $\boldsymbol{\alpha}_p$ and $\boldsymbol{\beta}_k$, which provides the following transformation of the estimates

$$\begin{aligned} n^{-(1-v)}\mathbf{W}_{0,n}' \begin{pmatrix} \tilde{\boldsymbol{\alpha}}_{p,n} - \boldsymbol{\alpha}_p \\ \tilde{\boldsymbol{\beta}}_{k,n} - \boldsymbol{\beta}_k \end{pmatrix} &= \begin{pmatrix} n^{v}\boldsymbol{\Gamma}_{p,n}^{-1}[(\tilde{\boldsymbol{\alpha}}_{p,n} - \boldsymbol{\alpha}_p) + \mathbf{A}_{k,p}'(\tilde{\boldsymbol{\beta}}_{k,n} - \boldsymbol{\beta}_k)] \\ n^{1/2+v}(\tilde{\boldsymbol{\beta}}_{k,n} - \boldsymbol{\beta}_k) \end{pmatrix} \\ &= \left((1/n)\sum_{t=1}^{n}\mathbf{g}_{tn}\mathbf{g}_{tn}'\right)^{-1}(1/n)\sum_{t=1}^{n}\mathbf{g}_{tn}n^{-(1-v)}U_t \end{aligned} \tag{2.74}$$

Where

$$(1/n)\sum_{t=1}^{n}\mathbf{g}_{tn}n^{-(1-v)}U_t \Rightarrow \int_0^1 \mathbf{g}(r)J_u(r)dr \tag{2.75}$$

With $\mathbf{g}(r)$ given in (3.4), $J_u(r) = B_u(r)$ under the assumption of cointegration, $v = 1/2$, and $J_u(r) = \int_0^r B_u(s)ds$ under no cointegration, that is when $v = -1/2$. Given that the limiting result in (3.6) does not contain the additive term $\boldsymbol{\Delta}_{ku}$, as appears in the limiting distribution of the OLS estimator of the cointegrating vector under cointegration, partial summing before estimating the model thus performs the same role for I-OLS that $(\mathbf{0}_{p+1}', n\boldsymbol{\Delta}_{ku}'^{+})'$ plays for FM-OLS, but this still leaves the problem that the correlation between u_t and $\boldsymbol{\varepsilon}_{k,t}$ rules out the possibility of conditioning on $\mathbf{B}_k(r)$ to obtain a conditional asymptotic normality result.[23] Irrespective of this result concerning the presence of a second-order bias, under cointegration these new estimators of the model parameters are consistent with the same rate of convergence as with the existent estimation methods previously discussed. Given that the practical utility of these results is limited only to the case of exogenous regressores, an additional correction must be performed to achieve the desired distributional results.

The solution to this problem proposed by these authors only requires that $\mathbf{X}_{k,t}$ be added as additional regressors to the partial sum regression (2.69) as

$$S_t = \boldsymbol{\alpha}_p'\mathbf{S}_{p,t} + \boldsymbol{\beta}_k'\mathbf{S}_{k,t} + \boldsymbol{\gamma}_k'\mathbf{X}_{k,t} + \zeta_t \quad t = 1,\dots,n \tag{2.76}$$

[23] Using, as before, the structure of $B_u(r)$ from Assumption 2.1, we have that the stochastic integral in (3.6) is decomposed as $\int_0^1 \mathbf{g}(r)B_u(r)dr = \int_0^1 \mathbf{g}(r)B_{u.k}(r)dr + (\int_0^1 \mathbf{g}(r)\mathbf{B}_k(r)'dr)\boldsymbol{\gamma}_k$. By defining $\mathbf{G}(r) = \int_0^r \mathbf{g}(s)ds$, then using integration by parts we have that this term can also be expressed as $\int_0^1 \mathbf{g}(r)B_u(r)dr = \mathbf{G}(1)B_u(1) - \int_0^1 \mathbf{G}(r)dB_u(r) = \int_0^1 [\mathbf{G}(1) - \mathbf{G}(r)]dB_u(r)$ Similarly we have $\int_0^1 \mathbf{g}(r)B_{u.k}(r)dr = \int_0^1 [\mathbf{G}(1) - \mathbf{G}(r)]dB_{u.k}(r)$, which is mixed Gaussian. For the second term above we get a similar representation with the same behaviour, but it still depends on the degree of endogeneity of the regressors as measured by $\boldsymbol{\gamma}_k = \boldsymbol{\Omega}_{kk}^{-1}\boldsymbol{\omega}_{ku}$.

Where $\zeta_t = U_t - \gamma_k' \mathbf{X}_{k,t}$, and which can be now called the integrated modified (IM) cointegrating regression. Following with the same structure of analysis as in previous section, particularly in relation to the structure of the underlying deterministic component of the integrated regressors, we next consider the properties of the OLS estimator of the model parameters $(\alpha_p', \beta_k', \gamma_k')'$ in (2.76) depending on this characteristic.

2.2.1 The case of regressors without deterministic trends

By OLS estimation of $(\alpha_p', \beta_k', \gamma_k')'$ in (3.7), we have that this estimator can be written as

$$\begin{pmatrix} \tilde{\alpha}_{p,n} - \alpha_p \\ \tilde{\beta}_{k,n} - \beta_k \\ \tilde{\gamma}_{k,n} - \gamma_k \end{pmatrix} = \left(\sum_{t=1}^{n} \mathbf{g}_{tn} \mathbf{g}_{tn}' \right)^{-1} \sum_{t=1}^{n} \mathbf{g}_{tn} \zeta_t = (\mathbf{W}_{0,n}')^{-1} \left((1/n) \sum_{t=1}^{n} \mathbf{g}_{tn} \mathbf{g}_{tn}' \right)^{-1} (1/n) \sum_{t=1}^{n} \mathbf{g}_{tn} \zeta_t$$

$$= n^{1-v} (\mathbf{W}_{0,n}')^{-1} \left((1/n) \sum_{t=1}^{n} \mathbf{g}_{tn} \mathbf{g}_{tn}' \right)^{-1} (1/n) \sum_{t=1}^{n} \mathbf{g}_{tn} (n^{-(1-v)} \zeta_t) \quad (2.77)$$

Where it has been used the following representation for the set of regressors

$$\begin{pmatrix} \mathbf{S}_{p,t} \\ \mathbf{S}_{k,t} \\ \mathbf{X}_{k,t} \end{pmatrix} = \begin{pmatrix} n \Gamma_{p,n}^{-1} (n^{-1} \mathbf{S}_{p,tn}) \\ n\sqrt{n} (n^{-3/2} \mathbf{H}_{k,t}) \\ \sqrt{n} (n^{-1/2} \eta_{k,t}) \end{pmatrix} = \mathbf{W}_{1,n} \begin{pmatrix} n^{-1} \mathbf{S}_{p,tn} \\ n^{-3/2} \mathbf{H}_{k,t} \\ n^{-1/2} \eta_{k,t} \end{pmatrix} = \mathbf{W}_{1,n} \mathbf{g}_{tn} \quad (2.78)$$

Given that in this case, and from (2.1), we impose the restricction $\mathbf{A}_{k,p} = \mathbf{0}_{k,p+1}$, so that $\mathbf{X}_{k,t} = \eta_{k,t}$, where the weighting matrix is given by $\mathbf{W}_{1,n} = diag(n\Gamma_{p,n}^{-1}, n\sqrt{n}\mathbf{I}_{k,k}, \sqrt{n}\mathbf{I}_{k,k})$. Thus, we get

$$\mathbf{W}_{1,n}^{-1} \begin{pmatrix} \mathbf{S}_{p,t} \\ \mathbf{S}_{k,t} \\ \mathbf{X}_{k,t} \end{pmatrix} = \mathbf{g}_{tn} \Rightarrow \mathbf{g}(r) = \begin{pmatrix} \mathbf{g}_p(r) \\ \mathbf{g}_k(r) \\ \mathbf{B}_k(r) \end{pmatrix} = \begin{pmatrix} \int_0^r \tau_p(s) ds \\ \int_0^r \mathbf{B}_k(s) ds \\ \mathbf{B}_k(r) \end{pmatrix} = \Pi \tilde{\mathbf{g}}(r) \quad (2.79)$$

With $\Pi = diag(\mathbf{I}_{p+1}, \Omega_{kk}^{1/2}, \Omega_{kk}^{1/2})$, $\tilde{\mathbf{g}}(r) = (\mathbf{g}_p'(r), \tilde{\mathbf{g}}_k'(r), \mathbf{W}_k'(r))'$, and $\tilde{\mathbf{g}}_k(r) = \int_0^r \mathbf{W}_k'(s) ds$, so that $\mathbf{g}(r)$ and $\tilde{\mathbf{g}}(r)$ are full-ranked processes, in the sense that $\int_0^1 \tilde{\mathbf{g}}(r) \tilde{\mathbf{g}}(r)' dr > \mathbf{0}$ a.s. Rewriting (2.77) as:

$$\tilde{\Theta}_n(v) = \begin{pmatrix} \tilde{\Theta}_{p,n}(v) \\ \tilde{\Theta}_{\beta k,n}(v) \\ \tilde{\Theta}_{\gamma k,n}(v) \end{pmatrix} = n^{-(1-v)} \mathbf{W}_{1,n}' \begin{pmatrix} \tilde{\alpha}_{p,n} - \alpha_p \\ \tilde{\beta}_{k,n} - \beta_k \\ \tilde{\gamma}_{k,n} - \gamma_k \end{pmatrix} = \begin{pmatrix} n^{v} \Gamma_{p,n}^{-1} (\tilde{\alpha}_{p,n} - \alpha_p) \\ n^{1/2+v} (\tilde{\beta}_{k,n} - \beta_k) \\ n^{-1/2+v} (\tilde{\gamma}_{k,n} - \gamma_k) \end{pmatrix}$$

$$= \left((1/n) \sum_{t=1}^{n} \mathbf{g}_{tn} \mathbf{g}_{tn}' \right)^{-1} (1/n) \sum_{t=1}^{n} \mathbf{g}_{tn} (n^{-(1-v)} \zeta_t) \quad (2.80)$$

Then, the limiting distribution of the IM-OLS estimates of the model parameters only depends on the stochastic limit of the term $n^{-(1-v)}\zeta_t = n^{-(1-v)}U_t - n^{-(1/2-v)}\boldsymbol{\gamma}_k'(n^{-1/2}\boldsymbol{\eta}_{k,t})$, both under cointegration and no cointegration. From the results in Section 2, we have that under cointegration, taking v = 1/2, then

$$n^{-(1-v)}\zeta_t = n^{-1/2}\zeta_t \Rightarrow B_u(r) - \boldsymbol{\gamma}_k'\mathbf{B}_k(r) = B_{u.k}(r) + (\boldsymbol{\omega}_{ku}'\boldsymbol{\Omega}_{kk}^{-1} - \boldsymbol{\gamma}_k')\mathbf{B}_k(r) = B_{u.k}(r) \tag{2.81}$$

Where the last equality comes from the assumption that $\boldsymbol{\gamma}_k = \boldsymbol{\Omega}_{kk}^{-1}\boldsymbol{\omega}_{ku}$, while that, under no cointegration when taking v = −1/2, then we get $n^{-3/2}\zeta_t \Rightarrow J_u(r) = \int_0^r B_u(s)ds$, that do not depends on the integrated regressors. These results imply that this estimator provides a well defined limiting distribution which, apart of some scale factors, will allow to develop valid inferential procedures. Theorem 2 in VW considers this case when $\mathbf{A}_{k,p} = \mathbf{0}_{k,p+1}$, which corresponds to the case where the integrated regressors do not contain any deterministic trend component, so that the trending parameters in the specification of the cointegrating regression must be associated to the deterministic component of the dependent variable. Also, taking (2.76) and (2.77), the sequence of IM-OLS residuals is given by

$$\begin{aligned}\tilde{\zeta}_{t,p}(k) &= \zeta_t - \mathbf{S}_{p,tn}'\boldsymbol{\Gamma}_{p,n}^{-1}(\tilde{\boldsymbol{\alpha}}_{p,n} - \boldsymbol{\alpha}_p) - \mathbf{S}_{k,t}'(\tilde{\boldsymbol{\beta}}_{k,n} - \boldsymbol{\beta}_k) - \mathbf{X}_{k,t}'(\tilde{\boldsymbol{\gamma}}_{k,n} - \boldsymbol{\gamma}_k)\\ &= \zeta_t - n^{1-v}(n^{-1}\mathbf{S}_{p,tn}')[n^v\boldsymbol{\Gamma}_{p,n}^{-1}(\tilde{\boldsymbol{\alpha}}_{p,n} - \boldsymbol{\alpha}_p)] - n^{1-v}(n^{-3/2}\mathbf{H}_{k,t}')[n^{1/2+v}(\tilde{\boldsymbol{\beta}}_{k,n} - \boldsymbol{\beta}_k)]\\ &\quad - n^{1-v}(n^{-1/2}\boldsymbol{\eta}_{k,t}')[n^{-1/2+v}(\tilde{\boldsymbol{\gamma}}_{k,n} - \boldsymbol{\gamma}_k)] = \zeta_t - n^{1-v}\mathbf{g}_{tn}'\tilde{\boldsymbol{\Theta}}_n(v)\end{aligned} \tag{2.82}$$

Which implies that the scaled t-th IM-OLS residual, given by

$$n^{-(1-v)}\tilde{\zeta}_{t,p}(k) = n^{-(1-v)}\zeta_t - \mathbf{g}_{tn}'\tilde{\boldsymbol{\Theta}}_n(v) \tag{2.83}$$

Will have a well defined stochastic limiting distribution, both under cointegration and no cointegration. The following result provides all these limiting distributions under these two possible situations, extending the results in Theorem 2 and Lemma 2 by VW concerning the IM-OLS estimators and scaled IM-OLS residuals in (2.80) and (2.83).

Proposition 3.1. *Given (2.1) under the assumption that* $\mathbf{A}_{k,p} = \mathbf{0}_{k,p+1}$, *and under Assumption 2.1, then under cointegration we have that:*

$$(a)\tilde{\Theta}_n(1/2)=\begin{pmatrix} n^{1/2}\Gamma_{p,n}^{-1}(\tilde{\alpha}_{p,n}-\alpha_p) \\ n(\tilde{\beta}_{k,n}-\beta_k) \\ \tilde{\gamma}_{k,n}-\gamma_k \end{pmatrix} \Rightarrow \Theta^0 = \left(\int_0^1 \mathbf{g}(r)\mathbf{g}(r)'dr\right)^{-1}\int_0^1 \mathbf{g}(r)B_{u.k}(r)dr$$

$$= \left(\int_0^1 \mathbf{g}(r)\mathbf{g}(r)'dr\right)^{-1}\int_0^1 [\mathbf{G}(1)-\mathbf{G}(r)]dB_{u.k}(r) \qquad (2.84)$$

$$=\omega_{u.k}(\Pi')^{-1}\left(\int_0^1 \tilde{\mathbf{g}}(r)\tilde{\mathbf{g}}(r)'dr\right)^{-1}\int_0^1 [\tilde{\mathbf{G}}(1)-\tilde{\mathbf{G}}(r)]dW_{u.k}(r)$$

$$(b)n^{-1/2}\tilde{\zeta}_{[nr],p}(k)\Rightarrow B_{u.k}(r)-\mathbf{g}(r)'\left(\int_0^1 \mathbf{g}(s)\mathbf{g}(s)'ds\right)^{-1}\int_0^1 \mathbf{g}(s)B_{u.k}(s)ds$$

$$=\omega_{u.k}\left(W_{u.k}(r)-\tilde{\mathbf{g}}(r)'\left(\int_0^1 \tilde{\mathbf{g}}(s)\tilde{\mathbf{g}}(s)'ds\right)^{-1}\int_0^1 \tilde{\mathbf{g}}(s)W_{u.k}(s)ds\right)=\omega_{u.k}R_{uk,p}(r) \qquad (2.85)$$

While that, under no cointegration, we have that

$$(c)\tilde{\Theta}_n(-1/2)=\begin{pmatrix} n^{-1/2}\Gamma_{p,n}^{-1}(\tilde{\alpha}_{p,n}-\alpha_p) \\ \tilde{\beta}_{k,n}-\beta_k \\ n^{-1}(\tilde{\gamma}_{k,n}-\gamma_k) \end{pmatrix} \Rightarrow \Theta^1 = \left(\int_0^1 \mathbf{g}(r)\mathbf{g}(r)'dr\right)^{-1}\int_0^1 \mathbf{g}(r)J_u(r)dr \qquad (2.86)$$

$$=\omega_u(\Pi')^{-1}\left(\int_0^1 \tilde{\mathbf{g}}(r)\tilde{\mathbf{g}}(r)'dr\right)^{-1}\int_0^1 [\tilde{\mathbf{G}}(1)-\tilde{\mathbf{G}}(r)]W_u(r)dr$$

And

$$(d)n^{-3/2}\tilde{\zeta}_{[nr],p}(k)\Rightarrow J_u(r)-\mathbf{g}(r)'\left(\int_0^1 \mathbf{g}(s)\mathbf{g}(s)'ds\right)^{-1}\int_0^1 \mathbf{g}(s)J_u(s)ds$$

$$=\omega_u\left(\int_0^r W_u(s)ds-\tilde{\mathbf{g}}(r)'\left(\int_0^1 \tilde{\mathbf{g}}(s)\tilde{\mathbf{g}}(s)'ds\right)^{-1}\int_0^1 [\tilde{\mathbf{G}}(1)-\tilde{\mathbf{G}}(r)]W_u(r)dr\right) \qquad (2.87)$$

Where $\mathbf{G}(r)=\int_0^r \mathbf{g}(s)ds=\Pi\int_0^r \tilde{\mathbf{g}}(s)ds$, *with* $\Pi=diag(\mathbf{I}_{p+1},\Omega_{kk}^{1/2},\Omega_{kk}^{1/2})$, *and* $\tilde{\mathbf{g}}(r)=(\mathbf{g}_p'(r),\tilde{\mathbf{g}}_k'(r),\mathbf{W}_k'(r))'$, *with* $J_u(r)=\int_0^r B_u(s)ds=\omega_u\int_0^r W_u(s)ds$, *and where we can write* $\Theta^0=(\Theta_p'^0,\Theta_{\beta,k}'^0,\Theta_{\gamma,k}'^0)'=\omega_{u.k}(\Pi')^{-1}\theta^0$, *and* $\Theta^1=(\Theta_p'^1,\Theta_{\beta,k}'^1,\Theta_{\gamma,k}'^1)'=\omega_u(\Pi')^{-1}\theta^1$, *with* $\theta^j=(\theta_p'^j,\theta_{\beta,k}'^j,\theta_{\gamma,k}'^j)'$, *j* = 0, 1.

Proof. The proof of parts (a) and (b) is given, respectively, in Theorem 2 and Lemma 2 in VW. The proof of parts (c) and (d) follows the same steps and only requires make use of the weak convergence of the sequence $n^{-3/2}\zeta_{[nr]}$ to $J_u(r)=\int_0^r B_u(s)ds$ under no cointegration.

An interesting by product of these distributional results under cointegration is that, conditional on $\mathbf{W}_k(r)$, it holds that $\Theta^0\sim N(\mathbf{0},\mathbf{V}_{IM})$, where the conditional asymptotic covariance matrix $\mathbf{V}_{IM}$ is given by

$$\mathbf{V}_{IM} = \omega_{u.k}^2 (\mathbf{\Pi}')^{-1} \left(\int_0^1 \tilde{\mathbf{g}}(r)\tilde{\mathbf{g}}(r)' dr \right)^{-1} \int_0^1 [\tilde{\mathbf{G}}(1) - \tilde{\mathbf{G}}(r)][\tilde{\mathbf{G}}(1) - \tilde{\mathbf{G}}(r)]' dr \\ \times \left(\int_0^1 \tilde{\mathbf{g}}(r)\tilde{\mathbf{g}}(r)' dr \right)^{-1} \mathbf{\Pi}^{-1} \tag{2.88}$$

That differs from the conditional asymptotic covariance matrix of the FM-OLS estimators of $\boldsymbol{\alpha}_p$ and $\boldsymbol{\beta}_k$ in the case where there is no any deterministic component in the structure of the observed integrated regressors.

Denoting by $\mathbf{m}(r) = (\boldsymbol{\tau}_p(r)', \mathbf{B}_k(r)')' = \mathbf{\Pi}_{FM}\tilde{\mathbf{m}}(r)$, with $\mathbf{\Pi}_{FM} = \mathrm{diag}(\mathbf{I}_{p+1}, \mathbf{\Omega}_{kk}^{1/2})$ and $\tilde{\mathbf{m}}(r) = (\boldsymbol{\tau}_p(r)', \mathbf{W}_k(r)')'$, then the conditional asymptotic covariance matrix of the FM-OLS estimator of $\boldsymbol{\alpha}_p$ and $\boldsymbol{\beta}_k$ is given by

$$\mathbf{V}_{FM} = \omega_{u.k}^2 (\mathbf{\Pi}_{FM}')^{-1} \left(\int_0^1 \tilde{\mathbf{m}}(r)\tilde{\mathbf{m}}(r)' dr \right)^{-1} \mathbf{\Pi}_{FM}^{-1} \tag{2.89}$$

There remains the question of how to compare these two covariance matrices,[24] but the results from a simulation study in VW (2011) generally indicate that the RMSE of the IM-OLS estimates tend to be larger than the RMSE of OLS and FM-OLS, except for highly endogenous regressors where the IM-OLS estimators performs better than the FM-OLS, specially when a large bandwidth is used.

These authors also suggest the use of the first difference of the IM-OLS residuals in (2.82) to define a semiparametric kernel estimator of the conditional long-run variance $\omega_{u.k}^2$, characterizing the limiting distribution of the IM-OLS parameter estimator and residuals under cointegration. Thus, by denoting $\tilde{z}_{t,p}(k)$ as $\tilde{z}_{t,p}(k) = \Delta\tilde{\zeta}_{t,p}(k)$, then from (2.82) the sequence of first differences of the IM-OLS residuals can be decomposed as

$$\begin{aligned} \tilde{z}_{t,p}(k) &= \Delta\zeta_t - n^{1-v}(n^{-1}\Delta\mathbf{S}_{p,tn}')[n^v\mathbf{\Gamma}_{p,n}^{-1}(\tilde{\boldsymbol{\alpha}}_{p,n} - \boldsymbol{\alpha}_p)] \\ &\quad - n^{1-v}(n^{-3/2}\Delta\mathbf{H}_{k,t}')[n^{1/2+v}(\tilde{\boldsymbol{\beta}}_{k,n} - \boldsymbol{\beta}_k)] \\ &\quad - n^{1-v}(n^{-1/2}\Delta\boldsymbol{\eta}_{k,t}')[n^{-1/2+v}(\tilde{\boldsymbol{\gamma}}_{k,n} - \boldsymbol{\gamma}_k)] \\ &= z_t - n^{-v}(\boldsymbol{\tau}_{p,tn}'[n^v\mathbf{\Gamma}_{p,n}^{-1}(\tilde{\boldsymbol{\alpha}}_{p,n} - \boldsymbol{\alpha}_p)] + n^{-1/2}\boldsymbol{\eta}_{k,t}'[n^{1/2+v}(\tilde{\boldsymbol{\beta}}_{k,n} - \boldsymbol{\beta}_k)]) \\ &\quad - n^{1/2-v}\boldsymbol{\varepsilon}_{k,t}'[n^{-1/2+v}(\tilde{\boldsymbol{\gamma}}_{k,n} - \boldsymbol{\gamma}_k)] \end{aligned} \tag{2.90}$$

Where $z_t = u_t - \boldsymbol{\gamma}_k'\boldsymbol{\varepsilon}_{k,t}$ is the fully modified error term from the cointegrating equation in levels. Then, the proposal consists in computing the estimator

$$\tilde{\omega}_n^2(m_n) = \sum_{h=-(n-1)}^{n-1} w(h/m_n)\tilde{\kappa}_n(h) \tag{2.91}$$

[24] In a recent revised version of the paper by VW, available at https://www.msu.edu/~tjv/IMpaper.pdf, these authors establish the inequality $\mathbf{V}_{FM} \leq \mathbf{V}_{IM}$, conditional on $\mathbf{W}_k(s)$ for the $\boldsymbol{\alpha}_p$ and $\boldsymbol{\beta}_k$ components, by using a continuous time version of the Gauss-Markov Theorem applied in this limiting framework.

For a given choice of kernel and bandwidth, with $\tilde{\kappa}_n(h) = (1/n)\sum_{t=h+1}^{n} \tilde{z}_{t,p}(k)\tilde{z}_{t-h,p}(k)$ the *h*th-order sample serial covariance of these residuals. The following proposition characterizes the asymptotic behavior of this estimator, both under cointegration and no cointegration.

Proposition 3.2. *Given the conditions under which the limiting distributions of the IM-OLS parameter estimators and residuals are valid, and under standard assumptions on the magnitude of the bandwidth parameter, then, under cointegration we have:*

(a) $\tilde{\omega}_n^2(m_n) \Rightarrow \omega_{u.k}^2(1 + \theta_{\gamma,k}'^0 \theta_{\gamma,k}^0)$ (2.92)

While that under no cointegration we have that $\tilde{\omega}_n^2(m_n) = O_p(n^2)$*, with*

(b) $(1/n^2)\tilde{\omega}_n^2(m_n) \Rightarrow \omega_u^2 \theta_{\gamma,k}'^1 \theta_{\gamma,k}^1$ (2.93)

With $\theta_{\gamma,k}^j$*, j = 0, 1, the scale-free random component of the limiting distribution of the IM-OLS estimator of* γ_k*, both under cointegration (when j = 0) and under no cointegration (when j = 1).*

Proof**.** See Appendix A.4

Remark 3.1. From the result in (2.92) we have that, under cointegration, the kernel estimator of the conditional long-run variance converges to a multiple of $\omega_{u.k}^2$, and the convergence is in distribution instead of the usual convergence in probability, so that this estimator is not consistent in the usual sense but still proves useful to built test statistics that are free of nuisance parameters. On the other hand, under no cointegration, the estimator diverges at the rate n^2 and its limit distribution depends again on a quadratic form defined in terms of the random vector determining the limit distribution of the IM-OLS estimator of γ_k under no cointegration. The result in (2.92) was proved earlier by VM, but the one in (2.93) is new here. In both cases, the limiting results are valid under a variety of permitted kernel functions and only require to impose some upper limit on the magnitude of the bandwidth parameter, m_n, which may be determined both as a deterministic or a stochastic function of the sample size. A very general and sufficient condition can be stated as $m_n = O_p(n^{1/2-\delta})$, with $0 < \delta \leq 1/2$, which also covers the deterministic case.[25] In any case, there remains to choose some practical rule for the effective determination of the number of sample autocovariances to enter in the computation of this estimator. A final comment must be made on the limiting behavior of the estimator $\tilde{\omega}_n^2(m_n)$ under the assumption of no cointegration.

[25] See, for example, Jansson (2002, 2005b) where the sample-dependent bandwidth parameter is formulated as $\hat{m}_n = \hat{a}_n b_n$, where $\hat{a}_n$ and b_n are both positive, with $\hat{a}_n + \hat{a}_n^{-1} = O_p(1)$ and b_n is nonrandom with $b_n^{-1} + n^{-1/2} b_n = o(1)$, so that $\hat{a}_n = O_p(1)$ and $b_n = o(n^{1/2})$, which gives $\hat{m}_n = o_p(n^{1/2})$.

Particularly, it has been shown that the most commonly used nonparametric kernel estimators, such as the one that makes use of the OLS residuals through the plug-in estimator $\hat{\omega}^2_{u.k,n}(m_n) = \hat{\omega}^2_n - \hat{\boldsymbol{\omega}}'_{ku,n}\hat{\boldsymbol{\Omega}}^{-1}_{kk,n}\hat{\boldsymbol{\omega}}_{uk,n}$ or, alternatively, the one based on the FM-OLS residuals $\hat{\omega}^{+2}_{u.k,n}(m_n)$ defined in equation (2.61), are $O_p(n{\cdot}m_n)$ under no cointegration when $n\to\infty$, so that the divergence rate is dependent on the bandwidth expansion rate and it is lower than for the estimator based on the first difference of the IM-OLS residuals. As will be considered in the next section, this result has a very important impact on the behavior and properties of the test statistics that will be introduced later.Alternatively, and following the idea developed by Kiefer and Vogelsang (2005), and further analized by Sun, Phillips and Jin (2008), we could consider the so called fixed-b estimation theory of a long-run variance based on a bandwidth that is simply proportional to the sample size as $m_n = b{\cdot}n$, with $b \in (0,1]$. The results in this case were extended by VW to models with nonstationary regressors, but the asymptotics are relatively more complex and no treated here. A particular case, that can be treated without any additional development, is when $b = 1$ so that the bandwidth is set equal to the sample size, $m_n = n$. The main difference between this latter approach and the standard one, with an effective truncation of the number of autocovariances used in the computation, is that this assumption allows to obtain a limiting expression for the nonparametric long-run variance estimator that is a random variable depending on the kernel $w(\cdot)$ and the value of b. Thus, this approach represents an informative theory that allows to capture the impact of the bandwidth and kernel choices on the sampling behavior of (2.91).When $b = 1$, and by Lemma 1 in Cai and Shintani (2006) for the Bartlett kernel[26], $w(x) = 1-|x|$, for $|x| \leq 1$, we can write (2.91) as follows

$$\begin{aligned}\tilde{\omega}^2_n(n) = n^{1-2v}\Bigg(2n^{-1}\Bigg\{&\sum_{t=1}^{n}(n^{-(1-v)}\tilde{\zeta}_{t,p}(k))^2 \\ &-(n^{-(1-v)}\tilde{\zeta}_{n,p}(k))\sum_{t=1}^{n}(n^{-(1-v)}\tilde{\zeta}_{t,p}(k))\Bigg\} + (n^{-(1-v)}\tilde{\zeta}_{n,p}(k))^2\Bigg)\end{aligned} \tag{2.94}$$

Whose asymptotic distribution is proportional to $\omega^2_{u.k}$ under the cointegration assumption by making use of the result (b) in Proposition 3.1. Thus, under cointegration and by simple application of the continuous mapping theorem, we have

$$\tilde{\omega}^2_n(n) \Rightarrow \omega^2_{u.k}\left(2\int_0^1 R_{uk,p}(s)^2 ds - R_{uk,p}(1)\int_0^1 R_{uk,p}(s)ds + R_{uk,p}(1)^2\right) \tag{2.95}$$

Also, from (2.94) and making use of the result (d) in Proposition 3.2 under no cointegration, it is immediate to check that $\tilde{\omega}^2_n(n) = O_p(n^2)$, where the limiting distribution of $n^{-2}\tilde{\omega}^2_n(n)$ is proportional to the long-run variance of the error sequence in the cointegrating regression, ω^2_u, as in equation (2.93), but with a different multiplicative random component.

[26] This result is an extension of the basic one formulated by Kiefer and Vogelsang (2002) when the total sum (over the full sample size) of the variables used in the computation of the sample autocovariances is not zero, as in the case where OLS residuals from a regression with a constant term are used.

2.3 The case of regressors with deterministic trends

This last section will be dedicated to the analysis of the performance of this new estimation method for a linear cointegrating regression equation when the integrated regressors are characterized by an underlying deterministic trend component, that is, when, from (2.1) and (2.5), we write $\mathbf{X}_{k,t} = \mathbf{A}_{k,p}\mathbf{\Gamma}^{-1}_{p,n}\boldsymbol{\tau}_{p,tn} + \sqrt{n}(n^{-1/2}\boldsymbol{\eta}_{k,t})$, with $\boldsymbol{\tau}_{p,tn} = \mathbf{S}_{p,tn} - \mathbf{S}_{p,(t-1)n}$. Once analyzed the consequences of this specification on the structure of the IM-OLS estimator, we propose a simple possible solution to the problems encountered. First of all, from (2.76) we have that the set of regressors can be written as:

$$\begin{pmatrix} \mathbf{S}_{p,t} \\ \mathbf{S}_{k,t} \\ \mathbf{X}_{k,t} \end{pmatrix} = \begin{pmatrix} n\mathbf{\Gamma}^{-1}_{p,n}(n^{-1}\mathbf{S}_{p,tn}) \\ n\mathbf{A}_{k,p}\mathbf{\Gamma}^{-1}_{p,n}(n^{-1}\mathbf{S}_{p,tn}) + n\sqrt{n}(n^{-3/2}\mathbf{H}_{k,t}) \\ n\mathbf{A}_{k,p}\mathbf{\Gamma}^{-1}_{p,n}(n^{-1}\mathbf{S}_{p,tn}) + \sqrt{n}(n^{-1/2}\boldsymbol{\eta}_{k,t}) \end{pmatrix} - \begin{pmatrix} \mathbf{0}_{p+1} \\ \mathbf{0}_{k} \\ n\mathbf{A}_{k,p}\mathbf{\Gamma}^{-1}_{p,n}(n^{-1}\mathbf{S}_{p,(t-1)n}) \end{pmatrix}$$

$$= \begin{pmatrix} n\mathbf{\Gamma}^{-1}_{p,n} & \mathbf{0}_{p+1,k} & \mathbf{0}_{p+1,k} \\ n\mathbf{A}_{k,p}\mathbf{\Gamma}^{-1}_{p,n} & n\sqrt{n}\mathbf{I}_{k,k} & \mathbf{0}_{k,k} \\ n\mathbf{A}_{k,p}\mathbf{\Gamma}^{-1}_{p,n} & \mathbf{0}_{k,k} & \sqrt{n}\mathbf{I}_{k,k} \end{pmatrix} \begin{pmatrix} n^{-1}\mathbf{S}_{p,tn} \\ n^{-3/2}\mathbf{H}_{k,t} \\ n^{-1/2}\boldsymbol{\eta}_{k,t} \end{pmatrix} - \begin{pmatrix} \mathbf{0}_{p+1,p+1} & \mathbf{0}_{p+1,k} & \mathbf{0}_{p+1,k} \\ \mathbf{0}_{k,p+1} & \mathbf{0}_{k,k} & \mathbf{0}_{k,k} \\ n\mathbf{A}_{k,p}\mathbf{\Gamma}^{-1}_{p,n} & \mathbf{0}_{k,k} & \mathbf{0}_{k,k} \end{pmatrix} \begin{pmatrix} n^{-1}\mathbf{S}_{p,(t-1)n} \\ n^{-3/2}\mathbf{H}_{k,t-1} \\ n^{-1/2}\boldsymbol{\eta}_{k,t-1} \end{pmatrix} \quad (2.96)$$

Or more compactly, by using $\mathbf{g}_{tn}$ defined in (2.23), as

$$\begin{pmatrix} \mathbf{S}_{p,t} \\ \mathbf{S}_{k,t} \\ \mathbf{X}_{k,t} \end{pmatrix} = \mathbf{W}_{21,n}\mathbf{g}_{tn} - \mathbf{W}_{22,n}\mathbf{g}_{(t-1)n} = \mathbf{W}_{21,n}(\mathbf{g}_{tn} - \mathbf{W}^{-1}_{21,n}\mathbf{W}_{22,n}\mathbf{g}_{(t-1)n}) \quad (2.97)$$

Where the last term in the expression between parenthesis is given by

$$\mathbf{W}^{-1}_{21,n}\mathbf{W}_{22,n}\mathbf{g}_{(t-1)n} = \mathbf{W}^{-1}_{21,n}\mathbf{W}_{22,n}\begin{pmatrix} n^{-1}\mathbf{S}_{p,(t-1)n} \\ n^{-3/2}\mathbf{H}_{k,t-1} \\ n^{-1/2}\boldsymbol{\eta}_{k,t-1} \end{pmatrix} = \begin{pmatrix} \mathbf{0}_{p+1} \\ \mathbf{0}_{k} \\ \sqrt{n}\mathbf{A}_{k,p}\mathbf{\Gamma}^{-1}_{p,n}(n^{-1}\mathbf{S}_{p,(t-1)n}) \end{pmatrix} \quad (2.98)$$

Which diverge with the sample size, even in the simplest case of a constant term ($p = 0$). Alternatively, if we redefine the IM regression model (3.7) in terms of the IM-OLS detrended variables we have $S^{*}_{t,p} = \boldsymbol{\beta}'_{k}\mathbf{S}^{*}_{kt,p} + \boldsymbol{\gamma}'_{k}\mathbf{X}^{*}_{kt,p} + \zeta^{*}_{t,p} \quad t = 1,...,n$

Where

$$\begin{pmatrix} S^{*}_{t,p} \\ \mathbf{S}^{*}_{kt,p} \\ \mathbf{X}^{*}_{kt,p} \end{pmatrix} = \begin{pmatrix} S_{t} \\ \mathbf{S}_{k,t} \\ \mathbf{X}_{k,t} \end{pmatrix} - \sum_{j=1}^{n} \begin{pmatrix} S_{j} \\ \mathbf{S}_{k,j} \\ \mathbf{X}_{k,j} \end{pmatrix} \mathbf{S}'_{p,j} \left(\sum_{j=1}^{n} \mathbf{S}_{p,j}\mathbf{S}'_{p,j} \right)^{-1} \mathbf{S}_{p,t} \quad (2.99)$$

And similarly for $\zeta^*_{t,p}$, then given $S_t = \boldsymbol{\alpha}'_{0,p}\mathbf{S}_{p,t} + h_{0,t}$ and $\mathbf{S}_{k,t}$ as in (3.3) we then have that

$$\begin{pmatrix} S^*_{t,p} \\ \mathbf{S}^*_{kt,p} \end{pmatrix} = \begin{pmatrix} h_{0,t} \\ \mathbf{H}_{k,t} \end{pmatrix} - \sum_{j=1}^{n} \begin{pmatrix} h_{0,j} \\ \mathbf{H}_{k,j} \end{pmatrix} \mathbf{S}'_{p,j} \left(\sum_{j=1}^{n} \mathbf{S}_{p,j}\mathbf{S}'_{p,j} \right)^{-1} \mathbf{S}_{p,t} = \begin{pmatrix} h^*_{0t,p} \\ \mathbf{H}^*_{kt,p} \end{pmatrix} \tag{2.100}$$

Which are free of trend parameters, while that for $\mathbf{X}^*_{kt,p}$ we get the decomposition

$$\begin{aligned} \mathbf{X}^*_{kt,p} &= \boldsymbol{\eta}_{k,t} - \sum_{j=1}^{n} \boldsymbol{\eta}_{k,j}\mathbf{S}'_{p,j} \left(\sum_{j=1}^{n} \mathbf{S}_{p,j}\mathbf{S}'_{p,j} \right)^{-1} \mathbf{S}_{p,t} \\ &\quad + \mathbf{A}_{k,p} \left(\boldsymbol{\tau}_{p,t} - \sum_{j=1}^{n} \boldsymbol{\tau}_{p,j}\mathbf{S}'_{p,j} \left(\sum_{j=1}^{n} \mathbf{S}_{p,j}\mathbf{S}'_{p,j} \right)^{-1} \mathbf{S}_{p,t} \right) = \boldsymbol{\eta}^*_{kt,p} + \mathbf{A}_{k,p}\boldsymbol{\tau}^*_{p,t} \end{aligned} \tag{2.101}$$

With

$$\boldsymbol{\eta}^*_{kt,p} = \sqrt{n} \left\{ (n^{-1/2}\boldsymbol{\eta}_{k,t}) - (1/n^2)\sum_{j=1}^{n} (n^{-1/2}\boldsymbol{\eta}_{k,j})\mathbf{S}'_{p,jn} \left((1/n^2)\sum_{j=1}^{n} \mathbf{S}_{p,jn}\mathbf{S}'_{p,jn} \right)^{-1} (n^{-1}\mathbf{S}_{p,tn}) \right\} \tag{2.102}$$

And

$$\mathbf{A}_{k,p}\boldsymbol{\tau}^*_{p,t} = \mathbf{A}_{k,p}\boldsymbol{\Gamma}^{-1}_{p,n} \left\{ \boldsymbol{\tau}_{p,tn} - (1/n^2)\sum_{j=1}^{n} \boldsymbol{\tau}_{p,jn}\mathbf{S}'_{p,jn} \left((1/n^2)\sum_{j=1}^{n} \mathbf{S}_{p,jn}\mathbf{S}'_{p,jn} \right)^{-1} (n^{-1}\mathbf{S}_{p,tn}) \right\} \tag{2.103}$$

That is $\mathbf{A}_{k,p}\boldsymbol{\tau}^*_{p,t} = \mathbf{A}_{k,p}\boldsymbol{\Gamma}^{-1}_{p,n}\boldsymbol{\tau}^*_{p,tn}$, so that $n^{-1/2}\mathbf{X}^*_{kt,p} = n^{-1/2}\boldsymbol{\eta}^*_{kt,p} + \mathbf{A}_{k,p}(n^{-1/2}\boldsymbol{\Gamma}^{-1}_{p,n})\boldsymbol{\tau}^*_{p,tn}$. For p = 0, $n^{-1/2}\mathbf{X}^*_{kt,0} = n^{-1/2}\boldsymbol{\eta}^*_{kt,0} + \mathbf{A}_{k,0}n^{-1/2}\boldsymbol{\tau}^*_{0,tn} = n^{-1/2}\boldsymbol{\eta}^*_{kt,0} + O(n^{-1/2})$, so that the deterministic component is asymptotically irrelevant, while for $p \geq 1$ we have that $n^{-1/2}\mathbf{X}^*_{kt,p} = n^{-1/2}\boldsymbol{\eta}^*_{kt,p} + O(n^{p-1/2})$, which implies that deterministic component dominates the stochastic one yielding inconsistent results. Thus, to deal with this general case, making use of the result for the OLS detrended observations of $\mathbf{X}_{k,t}$ and $\mathbf{Z}_{k,t} = \Delta\mathbf{X}_{k,t}$, $\hat{\mathbf{Z}}_{kt,p} = \boldsymbol{\varepsilon}_{kt,p}$, in Proposition 2.1 and in Remark 2.2 respectively, then we get the following augmented version of (2.27).

$$\hat{Y}_{t,p} = \boldsymbol{\beta}'_k\hat{\mathbf{X}}_{kt,p} + \boldsymbol{\gamma}'_k\hat{\mathbf{Z}}_{kt,p} + u_{t,p} - \boldsymbol{\gamma}'_k\hat{\mathbf{Z}}_{kt,p} = \boldsymbol{\beta}'_k\hat{\mathbf{X}}_{kt,p} + \boldsymbol{\gamma}'_k\hat{\mathbf{Z}}_{kt,p} + z_{t,p}, \quad t = 1,\ldots,n \tag{2.104}$$

Which gives the following corrected version of the IM cointegrating regression equation

$$\hat{S}_{t,p} = \boldsymbol{\beta}'_k\hat{\mathbf{S}}_{kt,p} + \boldsymbol{\gamma}'_k\sum_{j=1}^{t}\hat{\mathbf{Z}}_{kj,p} + \zeta_{t,p} = \boldsymbol{\beta}'_k\hat{\mathbf{S}}_{kt,p} + \boldsymbol{\gamma}'_k\hat{\mathbf{T}}_{kt,p} + \zeta_{t,p}, \quad t = 1,\ldots,n \tag{2.105}$$

With $\hat{\mathbf{S}}_{kt,p} = \sum_{j=1}^{t}\hat{\mathbf{X}}_{kj,p} = \sum_{j=1}^{t}\boldsymbol{\eta}_{kj,p}$, and $\zeta_{t,p} = \sum_{j=1}^{t} z_{t,p} = U_{t,p} - \boldsymbol{\gamma}'_k\hat{\mathbf{T}}_{kt,p}$, where

$$\begin{pmatrix} \hat{\mathbf{S}}_{kt,p} \\ \hat{\mathbf{T}}_{kt,p} \end{pmatrix} = \begin{pmatrix} n\sqrt{n}\mathbf{I}_{k,k} & \mathbf{0}_{k,k} \\ \mathbf{0}_{k,k} & \sqrt{n}\mathbf{I}_{k,k} \end{pmatrix} \begin{pmatrix} n^{-3/2}\hat{\mathbf{S}}_{kt,p} \\ n^{-1/2}\hat{\mathbf{T}}_{kt,p} \end{pmatrix} = \mathbf{W}_n \begin{pmatrix} n^{-3/2}\hat{\mathbf{S}}_{kt,p} \\ n^{-1/2}\hat{\mathbf{T}}_{kt,p} \end{pmatrix} \tag{2.106}$$

For the partial sum of the OLS detrended observations $\hat{\mathbf{Z}}_{kt,p}$, $\hat{\mathbf{T}}_{kt,p}$, we have

$$\begin{aligned}\hat{\mathbf{T}}_{k[nr],p} &= \sum_{t=1}^{[nr]} \hat{\mathbf{Z}}_{kt,p} = \sum_{t=1}^{[nr]} \boldsymbol{\varepsilon}_{k,t} - \sum_{j=1}^{n} \boldsymbol{\varepsilon}_{k,j} \boldsymbol{\tau}'_{p,jn} \mathbf{Q}_{pp,n}^{-1} \sum_{t=1}^{[nr]} \boldsymbol{\tau}_{p,tn} \\ &= -\boldsymbol{\eta}_{k,0} + \sqrt{n}\left(n^{-1/2} \boldsymbol{\eta}_{k,[nr]} - n^{-1/2} \sum_{j=1}^{n} \boldsymbol{\varepsilon}_{k,j} \boldsymbol{\tau}'_{p,jn} \bar{\mathbf{Q}}_{pp,n}^{-1} (1/n) \sum_{t=1}^{[nr]} \boldsymbol{\tau}_{p,tn} \right)\end{aligned} \tag{2.107}$$

Which, asymptotically, converge to a k-dimensional Brownian bridge of order (p+1)

$$n^{-1/2} \hat{\mathbf{T}}_{k[nr],p} \Rightarrow \mathbf{V}_{k,p}(r) = \mathbf{B}_k(r) - \int_0^1 d\mathbf{B}_k(s) \boldsymbol{\tau}'_p(s) \mathbf{Q}_{pp}^{-1} \int_0^r \boldsymbol{\tau}_p(s) ds \tag{2.108}$$

So that the random limit for the scaled vectors in the last multiplicative term in (2.106) we have

$$\begin{pmatrix} n^{-3/2} \hat{\mathbf{S}}_{k[nr],p} \\ n^{-1/2} \hat{\mathbf{T}}_{k[nr],p} \end{pmatrix} \Rightarrow \mathbf{g}_p(r) = \begin{pmatrix} \int_0^r \mathbf{B}_{k,p}(s) ds \\ \mathbf{V}_{k,p}(r) \end{pmatrix} \tag{2.109}$$

Then, from (2.10), (2.12) and (3.22) it can be easily verified that under cointegration ($|\alpha| < 1$ in Assumption 2.1), the scaled error term in the IM cointegrating regression (3.20) behaves asymptotically as

$$n^{-1/2} \zeta_{t,p} = n^{-1/2} U_{t,p} - \boldsymbol{\gamma}'_k n^{-1/2} \hat{\mathbf{T}}_{kt,p} \Rightarrow V_{u.k,p}(r). \tag{2.110}$$

When taking $t = [nr]$, where $V_{u.k,p}(r) = B_{u.k}(r) - \int_0^r \boldsymbol{\tau}'_p(s) ds \mathbf{Q}_{pp}^{-1} \int_0^1 \boldsymbol{\tau}_p(s) dB_{u.k}(s)$ only depends on $B_{u.k}(r)$ and can also be written as $V_{u.k,p}(r) = \omega_{u.k} W_{u.k,p}(r)$, with $W_{u.k,p}(r) = W_{u.k}(r) - \int_0^r \boldsymbol{\tau}'_p(s) ds \mathbf{Q}_{pp}^{-1} \int_0^1 \boldsymbol{\tau}_p(s) dW_{u.k}(s)$. Thus, we propose to obtain the IM-OLS estimator of the coefficient vector $(\boldsymbol{\beta}'_k, \boldsymbol{\gamma}'_k)'$, based on the OLS detrended observations, which is given by

$$\begin{pmatrix} \tilde{\boldsymbol{\beta}}_{k,n} \\ \tilde{\boldsymbol{\gamma}}_{k,n} \end{pmatrix} = \left(\sum_{t=1}^{n} \begin{pmatrix} \hat{\mathbf{S}}_{kt,p} \\ \hat{\mathbf{T}}_{kt,p} \end{pmatrix} (\hat{\mathbf{S}}'_{kt,p}, \hat{\mathbf{T}}'_{kt,p}) \right)^{-1} \sum_{t=1}^{n} \begin{pmatrix} \hat{\mathbf{S}}_{kt,p} \\ \hat{\mathbf{T}}_{kt,p} \end{pmatrix} \hat{S}_{t,p} \tag{2.111}$$

With IM-OLS residual sequence given by

$$\tilde{\zeta}_{t,p}(k) = \hat{S}_{t,p} - (\hat{\mathbf{S}}'_{kt,p}, \hat{\mathbf{T}}'_{kt,p}) \begin{pmatrix} \tilde{\boldsymbol{\beta}}_{k,n} \\ \tilde{\boldsymbol{\gamma}}_{k,n} \end{pmatrix} \qquad t = 1, \ldots, n \tag{2.112}$$

Next proposition establish the main result in this section related to the weak convergence of IM-OLS estimators and residuals under the assumption of cointegration, that is when the error term sequence u_t in the original cointegrating regression equation (2.3) is nonstationary with $\alpha = 1$ in Assumption 2.1.Proposition 3.3. *Given (2.1) and (2.2), and under Assumption 2.1, the IM-OLS estimation of the cointegrating regression model in (2.5) based on the IM regression (3.20) with OLS detrended observations, then equation (2.111) determine that:*

$$(a)\begin{pmatrix} n^{1/2+v}(\tilde{\boldsymbol{\beta}}_{k,n}-\boldsymbol{\beta}_k) \\ n^{-1/2+v}\tilde{\boldsymbol{\gamma}}_{k,n} \end{pmatrix} = \left((1/n)\sum_{t=1}^{n} \begin{pmatrix} n^{-3/2}\hat{\mathbf{S}}_{kt,p} \\ n^{-1/2}\hat{\mathbf{T}}_{kt,p} \end{pmatrix} (n^{-3/2}\hat{\mathbf{S}}'_{kt,p}, n^{-1/2}\hat{\mathbf{T}}'_{kt,p}) \right)^{-1} \times (1/n)\sum_{t=1}^{n} \left\{ \begin{pmatrix} n^{-3/2}\hat{\mathbf{S}}_{kt,p} \\ n^{-1/2}\hat{\mathbf{T}}_{kt,p} \end{pmatrix} n^{-(1-v)} U_{t,p} \right\} \tag{2.113}$$

And

$$(b) n^{-(1-v)}\tilde{\zeta}_{t,p}(k) = n^{-(1-v)}\zeta_{t,p} - (n^{-3/2}\hat{\mathbf{S}}'_{kt,p}, n^{-1/2}\hat{\mathbf{T}}'_{kt,p}) \begin{pmatrix} n^{1/2+v}(\tilde{\boldsymbol{\beta}}_{k,n}-\boldsymbol{\beta}_k) \\ n^{-1/2+v}(\tilde{\boldsymbol{\gamma}}_{k,n}-\boldsymbol{\gamma}_k) \end{pmatrix} \tag{2.114}$$

Which, under cointegration, gives

$$\begin{aligned} (c)\begin{pmatrix} n(\tilde{\boldsymbol{\beta}}_{k,n}-\boldsymbol{\beta}_k) \\ \tilde{\boldsymbol{\gamma}}_{k,n}-\boldsymbol{\gamma}_k \end{pmatrix} &\Rightarrow \int_0^1 \mathbf{g}_p(r)\mathbf{g}_p(r)'dr^{-1} \int_0^1 \mathbf{g}_p(r)V_{u.k,p}(r)dr \\ &= \omega_{u.k}\boldsymbol{\Pi}^{-1} \int_0^1 \tilde{\mathbf{g}}_p(r)\tilde{\mathbf{g}}_p(r)'dr^{-1} \int_0^1 \tilde{\mathbf{g}}_p(r)W_{u.k,p}(r)dr \\ &= \omega_{u.k}\boldsymbol{\Pi}^{-1} \int_0^1 \tilde{\mathbf{g}}_p(r)\tilde{\mathbf{g}}_p(r)'dr^{-1} \int_0^1 [\tilde{\mathbf{G}}_p(1)-\tilde{\mathbf{G}}_p(r)]dW_{u.k,p}(r) \end{aligned} \tag{2.115}$$

And

$$(d) n^{-1/2}\tilde{\zeta}_{t,p}(k) \Rightarrow \omega_{u.k}\left\{ W_{u.k,p}(r) - \tilde{\mathbf{g}}_p(r)' \int_0^1 \tilde{\mathbf{g}}_p(s)\tilde{\mathbf{g}}_p(s)'ds^{-1} \times \int_0^1 [\tilde{\mathbf{G}}_p(1)-\tilde{\mathbf{G}}_p(s)]dW_{u.k,p}(s) \right\} = \omega_{u.k}\tilde{R}_{u.k,p}(r) \tag{2.116}$$

With $V_{u.k,p}(r)=\omega_{u.k}W_{u.k,p}(r)$ *as in (2.12), where* $\mathbf{g}_p(r)$ *is given in (3.22) can be written as* $\mathbf{g}_p(r)=\boldsymbol{\Pi}\cdot\tilde{\mathbf{g}}_p(r)$, *with* $\mathbf{G}_p(r)=\int_0^r \mathbf{g}_p(s)ds=\boldsymbol{\Pi}\tilde{\mathbf{G}}_p(r)$, *and* $\boldsymbol{\Pi}=diag(\boldsymbol{\Omega}_{k,k}^{1/2},\boldsymbol{\Omega}_{k,k}^{1/2})$.

Proof. See Appendix A.5.

Remark 3.1. As can be seen in (c) and (d) above, for inferential purposes related to hypothesis testing relating the model parameters $\boldsymbol{\beta}_k$ and $\boldsymbol{\gamma}_k$, these limiting results only depends on $\omega_{u.k}^2$ and $\boldsymbol{\Omega}_{k,k}$ as nuisance parameters, as in the case of the use of asymptotically fully efficient estimation methods. Specially relevant, when using the IM-OLS residuals in (3.25), is the question of possible consistent estimation of the conditional long-run variance $\omega_{u.k}^2$ based on the first differences of $\tilde{\zeta}_{t,p}(k)$, $\Delta\tilde{\zeta}_{t,p}(k)$.

With these results, and under the same assumptions on the bandwidth parameter and kernel function as before, for the nonparametric kernel estimator $\tilde{\omega}_n^2(m_n)$ defined in (3.15) we get a similar result to (3.16) under cointegration where the limiting random element $\boldsymbol{\theta}_{\gamma,k}^0$ is taken from the last *k* terms in the limit distribution given in result (c) above.

2.4 IM-OLS residual-based test for the null of cointegration

In this section we propose some new statistics based on the sequence of IM-OLS residuals, as has been defined in section 3, for testing the null hypothesis of cointegration against the alternative of no cointegration by looking for excessive fluctuations in the sample paths of this residual sequence .

These new test statistics are partially inspired by the nonparametric variance-ratio statistic proposed by Breitung (2002) for testing the unit root null hypothesis against stationarity in a univariate time series, in the sense that our statistics are totally free of tuning parameters. In our case, we look for a unit root-like behavior in the residual sequence $\tilde{\zeta}_{t,p}(k)$ which is compatible with the stationarity of the error term $z_{t,p}$ in the augmented cointegrating regression among the OLS detrended variables.

In order to obtain a complete set of test statistics, including the case where there is no deterministic component neither in the specification of the cointegrating regression nor in the underlying structure of the observed integrated regressors (that is when $\boldsymbol{\alpha}_p = \mathbf{0}_{p+1}$ and $\mathbf{A}_{k,p} = \mathbf{0}_{k,p+1}$), we also consider the case where the IM cointegrating regression is given by

$$S_t = \boldsymbol{\beta}_k' \mathbf{S}_{k,t} + \boldsymbol{\gamma}_k' \mathbf{X}_{k,t} + \zeta_t,\ t = 1, \ldots, n, \qquad (2.117)$$

Where the IM-OLS estimator of $\boldsymbol{\beta}_k$ and $\boldsymbol{\gamma}_k$ is now given by

$$\begin{pmatrix} n^{1/2+v}(\tilde{\boldsymbol{\beta}}_{k,n} - \boldsymbol{\beta}_k) \\ n^{-1/2+v}\tilde{\boldsymbol{\gamma}}_{k,n} \end{pmatrix} = \left((1/n)\sum_{t=1}^{n} \begin{pmatrix} n^{-3/2}\mathbf{S}_{k,t} \\ n^{-1/2}\mathbf{X}_{k,t} \end{pmatrix} (n^{-3/2}\mathbf{S}_{k,t}', n^{-1/2}\mathbf{X}_{k,t}') \right)^{-1} \times (1/n)\sum_{t=1}^{n} \left\{ \begin{pmatrix} n^{-3/2}\mathbf{S}_{k,t} \\ n^{-1/2}\mathbf{X}_{k,t} \end{pmatrix} n^{-(1-v)} U_t \right\} \qquad (2.118)$$

Where it is verified that $(n^{-3/2}\mathbf{S}_{k,t}', n^{-1/2}\mathbf{X}_{k,t}')' = (n^{-3/2}\mathbf{H}_{k,t}', n^{-1/2}\boldsymbol{\eta}_{k,t}')' \Rightarrow \mathbf{g}(r)$ as $n \to \infty$ for $t = [nr]$, with $\mathbf{g}(r) = (\mathbf{g}_k(r)', \mathbf{B}_k(r)')'$ that admits the same factorization as for the last two terms in (3.21). Thus, under the assumption of cointegration, the limiting distribution of these estimates is as in Proposition 3.1, with $\tilde{\mathbf{g}}(r)$ replaced by $\tilde{\mathbf{g}}(r) = (\tilde{\mathbf{g}}_k(r)', \tilde{\mathbf{W}}_k(r)')'$.Once obtained the corresponding sequence of IM-OLS residuals in the appropriate specification and estimation of the IM cointegrating regression equation, that is $\tilde{\zeta}_t(k)$ from OLS estimation of (2.117) and $\tilde{\zeta}_{t,p}(k)$ from estimation of (2.76)[27] .

[27] In what follows we use this common notation to the IM-OLS residuals from estimating (3.7), $\tilde{\zeta}_{t,p}(k)$, $t = 1, \ldots, n$, both in the case where $\mathbf{A}_{k,p} = \mathbf{0}_{k,p+1}$, see equation (3.12), and also in the case of the OLS estimation of the specification given by equation (3.20), when $\mathbf{A}_{k,p} \neq \mathbf{0}_{k,p+1}$, which is based on OLS detrended observations of the model variables.

We define the following main components of our fluctuation test statistics

$$F_{1,n}(k) = (1/n)\sum_{t=1}^{n}((1/\sqrt{n})\tilde{\zeta}_t(k))^2 \quad (2.119)$$

$$F_{2,n}(k) = \max_{t=1,\ldots,n}(1/\sqrt{n})|\tilde{\zeta}_t(k)| \quad (2.120)$$

And

$$F_{3,n}(k) = \max_{t=1,\ldots,n}(1/\sqrt{n})|\tilde{\zeta}_t(k) - (t/n)\tilde{\zeta}_n(k)| \quad (2.121)$$

When considering the case of no deterministic component, and similarly $F_{j,n}(p,k)$ j = 1, 2, and 3, when using the IM-OLS residuals $\tilde{\zeta}_{t,p}(k)$, for $p \geq 0$.

Given the simple structure of these fluctuation statistics, it is immediate to check that (2.119) and (2.120) can also be written as

$$F_{1,n}(k) = n^{1-2v}\left\{(1/n)\sum_{t=1}^{n}(n^{-(1-v)}\tilde{\zeta}_t(k))^2\right\} \quad (2.122)$$

$$F_{2,n}(k) = n^{(1-2v)/2}\max_{t=1,\ldots,n}|n^{-(1-v)}\tilde{\zeta}_t(k)| \quad (2.123)$$

And similarly, for the statistic measuring the maximum centered fluctuation (2.121),

$$F_{3,n}(k) = n^{(1-2v)/2}\max_{t=1,\ldots,n}|n^{-(1-v)}\tilde{\zeta}_t(k) - (t/n)n^{-(1-v)}\tilde{\zeta}_n(k)| \quad (2.124)$$

With the value of $v = \pm 1/2$ denoting the cases of cointegration and no cointegration, respectively. From the results in Propositions 3.1 (b),(d), and 3.2, it is easy to check that both these quantities as the nonparametric kernel estimation based on the first difference of IM-OLS residuals with a bounded bandwidth parameter are of the same order of magnitude in any of the two possible situations. Thus, under cointegration (v = 1/2), we get that $F_{j,n}(k)$ and $\tilde{\omega}_n^2(m_n)$ are $O_p(1)$, while that under no cointegration (v = −1/2), we have that $F_{1,n}(k)$ and $\tilde{\omega}_n^2(m_n)$ are $O_p(n^2)$, and $F_{j,n}(k) = O_p(n)$, for j = 2, 3.

This observation means that, when considering the building of asymptotically pivotal test statistics by combining the fluctuation measures in (2.119)-(2.121) and the estimator of the long-run conditional variance $\tilde{\omega}_n^2(m_n)$, we will not obtain the desired consistency result (that is, divergence under the alternative). Despite this undesired result, we continue to define and to explore the behavior of the following set of statistics to test the null hypothesis of cointegration:

$$\tilde{F}_{1,n}(k) = \tilde{\omega}_n^{-2}(m_n)F_{1,n}(k) = \frac{1}{n\tilde{\omega}_n^2(m_n)}\sum_{t=1}^{n}((1/\sqrt{n})\tilde{\zeta}_t(k))^2 \quad (2.125)$$

And $\tilde{F}_{j,n}(k) = \tilde{\omega}_n^{-1}(m_n)F_{j,n}(k)$, for j = 2, 3, and similarly $\tilde{F}_{j,n}(p,k)$, j = 1, 2, 3 when using the IM-OLS residuals from (2.82) and (2.105). Next, we present the limiting distributional results of the fluctuation measures in (2.119)-(2.121) in the simplest case where there is no deterministic component, both under cointegration and under no cointegration, where the extension to the case of IM-OLS estimation with a trend function is trivial from the results in Propositions 3.1 and 3.3. Proposition 4.1. *Under the null hypothesis of cointegration, that is when* $|\alpha| < 1$ *in Assumption 2.1 with v = 1/2, then:*

$$\begin{aligned} (a) F_{1,n}(k) &\Rightarrow \omega_{u.k}^2 \int_0^1 R_{u.k}(s)^2 ds \\ F_{2,n}(k) &\Rightarrow \omega_{u.k} \sup_{r\in[0,1]} |R_{u.k}(r)| \\ F_{3,n}(k) &\Rightarrow \omega_{u.k} \sup_{r\in[0,1]} |R_{u.k}(r) - r\cdot R_{u.k}(1)| \end{aligned} \tag{2.126}$$

Where

$$R_{u.k}(r) = W_{u.k}(r) - \tilde{\mathbf{g}}(r)' \left(\int_0^1 \tilde{\mathbf{g}}(s)\tilde{\mathbf{g}}(s)'ds \right)^{-1} \int_0^1 [\tilde{\mathbf{G}}(1) - \tilde{\mathbf{G}}(s)]dW_{u.k}(s) \tag{2.127}$$

With $\tilde{\mathbf{g}}(r) = (\tilde{\mathbf{g}}_k(r)', \tilde{\mathbf{W}}_k(r)')'$, $\tilde{\mathbf{g}}_k(r) = \int_0^r \mathbf{W}_k'(s)ds$ *, and similarly for* $F_{j,n}(p,k)$ *, j = 1, 2, 3, with* $R_{u.k}(r)$ *replaced by* $R_{u.k,p}(r)$ *in Proposition 3.1(b) or by* $\tilde{R}_{u.k,p}(r)$ *in Proposition 3.3(d), and the proper choice of* $\tilde{\mathbf{g}}(r)$ *depending on the assumption made about the deterministic component in the integrated regressors. Also, under the alternative hypothesis of no cointegration, that is when* $\alpha = 1$ *in Assumption 2.1 with v = –1/2, then:*

$$\begin{aligned} (b) n^{-2}F_{1,n}(k) &\Rightarrow \int_0^1 J_k(s)^2 ds \\ n^{-1}F_{2,n}(k) &\Rightarrow \sup_{r\in[0,1]} |J_k(r)| \\ n^{-1}F_{3,n}(k) &\Rightarrow \sup_{r\in[0,1]} |J_k(r) - rJ_k(1)| \end{aligned} \tag{2.128}$$

Where

$$J_k(r) = J_u(r) - \tilde{\mathbf{g}}(r)' \left(\int_0^1 \tilde{\mathbf{g}}(s)\tilde{\mathbf{g}}(s)'ds \right)^{-1} \int_0^1 \tilde{\mathbf{g}}(s)J_u(s)ds \tag{2.129}$$

With $J_u(r) = \int_0^r B_u(s)ds$ *, and similarly for* $n^{-2}F_{1,n}(p,k)$ *, and* $n^{-1}F_{j,n}(p,k)$ *, j = 2, 3, with* $J_k(r)$ *replaced by* $J_{k,p}(r)$ *defined as*

$$J_{k,p}(r) = J_{u,p}(r) - \tilde{\mathbf{g}}(r)' \left(\int_0^1 \tilde{\mathbf{g}}(s)\tilde{\mathbf{g}}(s)'ds \right)^{-1} \int_0^1 \tilde{\mathbf{g}}(s)J_{u,p}(s)ds \tag{2.130}$$

Where $J_{u,p}(r) = \int_0^r B_{u,p}(s)ds$.

Proof. See Appendix A.6.

Next, by combining these last results with the one relating the limiting behavior of the nonparametric kernel estimator of the conditional long-run variance in Proposition 3.2 based on the first difference of the IM-OLS residuals with a bounded bandwidth, we get the following result characterizing the limiting distribution of the fluctuation test statistics introduced above.

Corolary 4.2. *Under the null hypothesis of cointegration, that is when* $|\alpha| < 1$ *in Assumption 2.1 with v = 1/2, then:*

$$\tilde{F}_{1,n}(k) \Rightarrow (1+\theta_{\gamma,k}'^{0}\theta_{\gamma,k}^{0})^{-1}\int_0^1 R_{u.k}(s)^2 ds \tag{2.131}$$

$$\tilde{F}_{2,n}(k) \Rightarrow (1+\theta_{\gamma,k}'^{0}\theta_{\gamma,k}^{0})^{-1/2} \sup_{r\in[0,1]} |R_{u.k}(r)| \tag{2.132}$$

And

$$\tilde{F}_{3,n}(k) \Rightarrow (1+\theta_{\gamma,k}'^{0}\theta_{\gamma,k}^{0})^{-1/2} \sup_{r\in[0,1]} |R_{u.k}(r) - r\cdot R_{u.k}(1)| \tag{2.133}$$

While that $\tilde{F}_{j,n}(k) = O_p(1)$*, j = 1, 2, 3 under no cointegration, with*

$$\tilde{F}_{1,n}(k) \Rightarrow \omega_u^{-2}(\theta_{\gamma,k}'^{1}\theta_{\gamma,k}^{1})^{-1}\int_0^1 J_k(s)^2 ds \tag{2.134}$$

$$\tilde{F}_{2,n}(k) \Rightarrow \omega_u^{-1}(\theta_{\gamma,k}'^{1}\theta_{\gamma,k}^{1})^{-1/2} \sup_{r\in[0,1]} |J_k(r)| \tag{2.135}$$

And $\tilde{F}_{3,n}(k) \Rightarrow \omega_u^{-1}(\theta_{\gamma,k}'^{1}\theta_{\gamma,k}^{1})^{-1/2} \sup_{r\in[0,1]} |J_k(r) - rJ_k(1)|$*, where the limiting random elements* $R_{u.k}(r)$ *and* $J_k(r)$ *can be conveniently replaced by the one determining the limiting distribution of the IM-OLS residuals when including the adjustment for deterministic components in Propositions 3.1 and 3.3. The same applies to the structure of* $\theta_{\gamma,k}^{0}$ *and* $\theta_{\gamma,k}^{1}$ *characterizing the limiting distribution of the IM-OLS estimator of* γ_k *.*

Proof. It follows directly by combining the results in Proposition 3.2 and 4.1 and the application of the continuous mapping theorem.

Remark 4.1. These results clearly show that the random limits, both under the null of cointegration and under the alternative hypothesis of no cointegration, are free of nuisance parameters, and only depends on the number of integrated regressors and the structure of the deterministic component in the cointegrating regression and the assumption made on such components characterizing the generating process of the observed regressors. In the three cases, the tests are right-sided, rejecting the null of cointegration for high values of the corresponding test statistic. Also, from these results we may immediately conclude that the testing procedures for the null of cointegration based on these test statistics are inconsistent in the usual sense that the behavior under the alternative of no cointegration does not depend on the sample size, and hence there seems that they cannot correctly discriminate between these two situations. However, given that the random limits are very different in each case, we may expect to obtain certain useful results in terms of power. This last issue will be examined numerically.

Appendix B.2 (Tables B.2.1 and B.2.2) presents relevant quantiles of the null distribution of these three testing procedures, $c_{j,\alpha}(k)$, for each specification of the deterministic component in the cointegrating regression as well as for each assumption on the underlying deterministic component characterizing the observations of the k integrated regressors in the model, for k = 1, ..., 5. These quantiles are computed numerically with 20000 replications and 2000 observations, in the simplest case where $\xi_t = (u_t, \varepsilon'_{k,t})' \sim iidN(\mathbf{0}, \mathbf{I}_{k+1})$ and α = 0. For computation of the estimator of the conditional long-run variance $\tilde{\omega}_n^2(m_n)$ in this simple setup, we consider the case m_n = 0, so that $\tilde{\omega}_n^2(m_n) = \tilde{\kappa}_n(0)$. As can be seen in equation (2.90) and in the proof of Proposition 3.2, under cointegration the first difference of the IM-OLS residuals are given by

$$\tilde{z}_{t,p}(k) = z_t - n^{1/2-v}\varepsilon'_{k,t}[n^{-1/2+v}(\tilde{\gamma}_{k,n} - \gamma_k)] + O_p(n^{-1/2})$$

So that the short-run sample variance can be decomposed as

$$\begin{aligned}\tilde{\kappa}_n(0) &= (1/n)\sum_{t=1}^{n} z_t^2 + (\tilde{\gamma}_{k,n} - \gamma_k)'(1/n)\sum_{t=1}^{n}\varepsilon_{k,t}\varepsilon'_{k,t}(\tilde{\gamma}_{k,n} - \gamma_k) - 2(\tilde{\gamma}_{k,n} - \gamma_k)'(1/n)\sum_{t=1}^{n}\varepsilon_{k,t}z_t \\ &\quad + 2O_p(n^{-1/2})(1/\sqrt{n})\left\{(1/\sqrt{n})\sum_{t=1}^{n} z_t - (\tilde{\gamma}_{k,n} - \gamma_k)'(1/\sqrt{n})\sum_{t=1}^{n}\varepsilon_{k,t}\right\} + O_p(n^{-1})\end{aligned} \quad (2.136)$$

This means that the limiting null distribution of $\tilde{\kappa}_n(0)$ is given by

$$\begin{aligned}\tilde{\kappa}_n(0) \Rightarrow \tilde{\kappa}(0) &= \sigma_u^2 - \boldsymbol{\omega}'_{ku}\boldsymbol{\Omega}_{kk}^{-1}(2\boldsymbol{\sigma}_{ku} - \boldsymbol{\Sigma}_{kk}\boldsymbol{\Omega}_{kk}^{-1}\boldsymbol{\omega}_{ku}) \\ &\quad + [\sigma_u^2 - \boldsymbol{\omega}'_{ku}\boldsymbol{\Omega}_{kk}^{-1}\boldsymbol{\omega}_{ku} + (\omega_u^2 - \sigma_u^2)]\boldsymbol{\theta}'^0_{\gamma k}\boldsymbol{\Omega}_{kk}^{-1/2}\boldsymbol{\Sigma}_{kk}\boldsymbol{\Omega}_{kk}^{-1/2}\boldsymbol{\theta}^0_{\gamma k} \\ &\quad - 2\omega_{u.k}\boldsymbol{\theta}'^0_{\gamma k}\boldsymbol{\Omega}_{kk}^{-1/2}(\boldsymbol{\sigma}_{ku} - \boldsymbol{\Sigma}_{kk}\boldsymbol{\Omega}_{kk}^{-1}\boldsymbol{\omega}_{ku})\end{aligned} \quad (2.137)$$

Under serially uncorrelated error terms u_t and $\varepsilon_{k,t}$, so that $\omega_u^2 = \sigma_u^2$, and $\boldsymbol{\Omega}_{kk} = \boldsymbol{\Sigma}_{kk}$, and the additional assumption that these error terms are only contemporaneously correlated, that is $\boldsymbol{\omega}_{ku} = \boldsymbol{\sigma}_{ku}$, then the above weak limit of $\tilde{\kappa}_n(0)$, $\tilde{\kappa}(0)$, will reduce to $\tilde{\kappa}(0) = \sigma_{u.k}^2(1 + \boldsymbol{\theta}'^0_{\gamma k}\boldsymbol{\theta}^0_{\gamma k})$, with $\sigma_{u.k}^2 = \sigma_u^2 - \boldsymbol{\sigma}'_{ku}\boldsymbol{\Sigma}_{kk}^{-1}\boldsymbol{\sigma}_{ku}$ the conditional short-run variance of u_t given $\varepsilon_{k,t}$. Also, Tables B.2.3 and B.2.4 present the results of the power behavior in finite samples, when α = 1, of the testing procedure based on the test statistics $\tilde{F}_{1,n}(k)$ and $\tilde{F}_{1,n}(p,k)$ defined in (4.6) in each case considered before with a deterministic sample-dependent bandwidth given by $m_n = [d(n/100)^{1/4}]$, for d = 1, 4, and 12. Alternatively, and for comparison purposes, we also compute the power performance of this fluctuation-based statistic when using the OLS-based estimator of the conditional long-run variance, $\hat{\omega}_{u.k,n}^2(m_n) = \hat{\omega}_n^2 - \hat{\boldsymbol{\omega}}'_{ku,n}\hat{\boldsymbol{\Omega}}_{kk,n}^{-1}\hat{\boldsymbol{\omega}}_{uk,n}$, denoted as $\hat{F}_{1,n}(p,k)$.[28]

[28] The quantiles of the asymptotic null distribution of the testing procedure based on the test statistic $\hat{F}_{1,n}(p,k) = \hat{\omega}_{u.k,n}^{-2}(m_n)F_{1,n}(p,k)$ are different from those shown in Tables B.2.1 and B.2.2, are not presented here, but can be requested from the author.

These results are presented in Tables B.2.5 and B.2.6, and show the usual pattern of increasing power with the sample size that comes from the different rate of divergence of the numerator and denominator under no cointegration. However, for the test based on the statistic $\tilde{F}_{1,n}(p,k)$, the power performance is quite different, displaying an increasing power with the sample size for low dimensional models (k = 1, 2), decreasing for high dimensional models ($k \geq$ 3), but converging to a relatively acceptable common level that depends on the specification of the deterministic component (see Figures 2.1, 2.2).

2.5 Conclusions and some extensions

The present paper is devoted to the analysis of the asymptotically efficient estimation of a linear static cointegrating regression model by making use of a new recently proposed estimation method by Vogelsang and Wagner (2011), the so-called integrated modified OLS estimator (IM-OLS) that has the main advantage that does not require the choice of any tuning parameter, when we deal with deterministically trending integrated regressors. We show that this method must be modified to correctly accommodate the structure of the deterministic component of the regressors and to avoid possible inconsistencies in the estimation results. As a by product of these results, we propose the use of the IM-OLS residuals to build some new simple statistics to testing the null hypothesis of cointegration against the alternative of no cointegration.

While the main component of these new test statistics seems to work well in detecting excessive fluctuations in the residual sequence under no cointegration, it is not yet clear how to obtain pivotal test statistics free of nuisance parameters and consistent tests given the difficulties in obtaining a proper estimator of a long-run variance. This central question will be studied in future work, as well as the consideration of more complex deterministic components and their treatment in the context of the IM-OLS estimation.

2.6 References

Afonso-Rodríguez, J.A. (2013). "Integrated modified OLS estimation of cointegrating regressions with deterministically trending regressors under near cointegration". Paper presented at XXVII International Congress of Applied Economics, ASEPELT.

Breitung, J. (2002). "Nonparametric tests for unit roots and cointegration". Journal of Econometrics, 108(2), pp. 343-363.

Breitung, J., A.M.R. Taylor (2003). Corrigendum to "Nonparametric tests for unit roots and cointegration". Journal of Econometrics, 117(2), pp. 401-404.

Cai, Y., M. Shintani (2006). "On the alternative long-run variance ratio test for a unit root". Econometric Theory, 22(3), pp. 347-372.

Hansen, B.E. (1992). "Efficient estimation and testing of cointegrating vectors in the presence of deterministic trends". Journal of Econometrics, 53(1-3), pp. 87-121.

Hassler, U. (2001). “The effect of linear time trends on the KPSS test for cointegration”. Journal of Time Series Analysis, 22(3), pp. 283-292.

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

Jansson, M. (2005a). “Point optimal tests of the null hypothesis of cointegration”. Journal of Econometrics, 124(1), pp. 187-201.

Jansson, M. (2005b). “Tests of the null hypothesis of cointegration based on efficient tests for a unit MA root”. In: Andrews, D.W.K, J.H. Stock (Eds.), Identification and Inference in Econometric Models: Essays in Honour of T.J. Rothenberg, Chapter 15, pp. 357-374. Cambridge University Press, Cambridge.

Kiefer, N.M., T.J. Vogelsang (2005). “A new asymptotic theory for heteroskedasticity-autocorrelation robust tests”. Econometric Theory, 21(6), pp. 1130-1164.

Kurozumi, E., K. Hayakawa (2009). “Asymptotic properties of the efficient estimators for cointegrating regression models with serially dependent errors”. Journal of Econometrics, 149(2), pp. 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), pp. 159-178.

Park, J.Y. (1992). “Canonical cointegrating regressions”. Econometrica, 60(1), pp. 119-143.

Park, J.Y., P.C.B. Phillips (1988). “Statistical inference in regressions with integrated processes: Part 1”. Econometric Theory, 4(3), pp. 468-498.

Phillips, P.C.B. (1987). “Towards a unified asymptotic theory for autoregression”. Biometrika, 74(3), pp. 535-547.

Phillips, P.C.B. (1995). “Fully modified least squares and vector autoregression”. Econometrics, 63(5), pp. 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), pp. 99-125.

Phillips, P.C.B., M. Loretan (1991). “Estimating long-run economic equilibria”. The Review of Economic Studies, 58(3), pp. 407-436.

Phillips, P.C.B., S.N. Durlauf (1986). “Multiple time series with integrated variables”. Review of Economic Studies, 53(4), pp. 473-496.

Phillips, P.C.B., Y. Chang (1995). “Time series regression with mixtures of integrated processes”. Econometric Theory, 11(5), pp. 1033-1094.

Saikkonen, P. (1991). “Asymptotically efficient estimation of cointegration regressions”. Econometric Theory, 7(1), pp. 1-21.

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

Stock, J.H., M.W. Watson (1993). "A simple estimator of cointegrating vectors in higher-order integrated systems". Econometrica, 61(4), pp. 783-820.

Sun, Y., P.C.B. Phillips, S. Jin (2008). "Optimal bandwidth selection in heteroskedasticity-autocorrelation robust testing". Econometric, 76(1), pp. 175-194.

Vogelsang, T.J., M. Wagner (2011). "Integrated Modified OLS estimation and fixed-b inference for cointegrating regressions". Working Paper No.263, Institute for Advanced Studies (IHS), Vienna.

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

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

Xiao, Z. (2001). "Testing the null hypothesis of stationarity against an autoregressive unit root alternative". Journal of Time Series Analysis, 22(1), pp. 87-103.

Xiao, Z., P.C.B. Phillips (1999). "Efficient detrending in cointegrating regression". Econometric Theory, 15(4), pp. 519-548.

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

2.7 Appendix A. Mathematical proofs

A.1 Proof of Proposition 2.1. By OLS detrending of the observed processes Y_t and $\mathbf{X}_{k,t}$, as defined by (2.1) and (2.2), we have that

$$\begin{pmatrix} \hat{Y}_{t,p} \\ \hat{\mathbf{X}}_{kt,p} \end{pmatrix} = \begin{pmatrix} Y_t \\ \mathbf{X}_{k,t} \end{pmatrix} - \sum_{j=1}^{n} \begin{pmatrix} Y_j \\ \mathbf{X}_{k,j} \end{pmatrix} \boldsymbol{\tau}'_{p,j} \cdot \mathbf{Q}_{n,pp}^{-1} \boldsymbol{\tau}_{p,t} \quad t = 1, \ldots, n \tag{2.137}$$

Each of the components above can be decomposed as

$$\eta_{it,p} + \boldsymbol{\alpha}'_{i,p_i} (\boldsymbol{\tau}_{p_i,t} - \textstyle\sum_{j=1}^{n} \boldsymbol{\tau}_{p_i,j} \boldsymbol{\tau}'_{p,j} \mathbf{Q}_{n,pp}^{-1} \boldsymbol{\tau}_{p,t}) \qquad i = 0, 1, \ldots, k, \tag{2.138}$$

Where $\eta_{it,p} = \eta_{i,t} - \sum_{j=1}^{n} \eta_{i,j} \boldsymbol{\tau}'_{p,j} \mathbf{Q}_{n,pp}^{-1} \boldsymbol{\tau}_{p,t}$, and

$$\begin{aligned}\boldsymbol{\tau}_{p_i,t}-\sum_{j=1}^{n}\boldsymbol{\tau}_{p_i,j}\boldsymbol{\tau}'_{p,j}\mathbf{Q}_{n,pp}^{-1}\boldsymbol{\tau}_{p,t}&=\boldsymbol{\tau}_{p_i,t}-\sum_{j=1}^{n}\boldsymbol{\tau}_{p_i,j}(\boldsymbol{\tau}'_{p_i,j}:\boldsymbol{\tau}'_{p-p_i,j})\mathbf{Q}_{n,pp}^{-1}\begin{pmatrix}\boldsymbol{\tau}_{p_i,t}\\ \boldsymbol{\tau}_{p-p_i,t}\end{pmatrix}\\ &=\boldsymbol{\tau}_{p_i,t}-(\mathbf{Q}_{n,p_ip_i}:\mathbf{Q}_{n,p_i(p-p_i)})\begin{pmatrix}\mathbf{Q}_{n,p_ip_i} & \mathbf{Q}_{n,p_i(p-p_i)}\\ \mathbf{Q}'_{n,p_i(p-p_i)} & \mathbf{Q}_{n,(p-p_i)(p-p_i)}\end{pmatrix}^{-1}\begin{pmatrix}\boldsymbol{\tau}_{p_i,t}\\ \boldsymbol{\tau}_{p-p_i,t}\end{pmatrix}\\ &=\boldsymbol{\tau}_{p_i,t}-(\mathbf{I}_{p_i+1,p_i+1}:\mathbf{0}_{p_i+1,p-p_i})\begin{pmatrix}\boldsymbol{\tau}_{p_i,t}\\ \boldsymbol{\tau}_{p-p_i,t}\end{pmatrix}=\mathbf{0}_{p_i+1}\end{aligned} \tag{2.139}$$

Given the block structure for the inverse of $\mathbf{Q}_{n,pp}$, when $p_i < p$ for all $i = 0, 1, \ldots, k$. Obviously, the same result directly holds when $p_i = p$, while that if any $p_i > p$, then we have $\boldsymbol{\alpha}'_{i,p_i}(\boldsymbol{\tau}_{p_i,t}-\sum_{j=1}^{n}\boldsymbol{\tau}_{p_i,j}\boldsymbol{\tau}'_{p,j}\mathbf{Q}_{n,pp}^{-1}\boldsymbol{\tau}_{p,t})=\boldsymbol{\alpha}'_{i,p_i-p}(\boldsymbol{\tau}_{p_i-p,t}-\mathbf{Q}_{n,(p_i-p)p}\mathbf{Q}_{n,pp}^{-1}\boldsymbol{\tau}_{p,t})$, which does not vanish and it is of order $O(n^{p_i})$.

■

A.2 Proof of Proposition 2.2. First, given that we can write

$$\left(\sum_{t=1}^{n}\begin{pmatrix}\boldsymbol{\tau}_{p,tn}\\ \boldsymbol{\eta}_{k,tn}\end{pmatrix}(\boldsymbol{\tau}'_{p,tn},\boldsymbol{\eta}'_{k,tn})\right)^{-1}=\begin{pmatrix}\mathbf{Q}_{pp,n} & \mathbf{Q}_{pk,n}\\ \mathbf{Q}'_{pk,n} & \mathbf{Q}_{kk,n}\end{pmatrix}^{-1}=\begin{pmatrix}\mathbf{M}_{pp,n}^{-1} & -\mathbf{M}_{pp,n}^{-1}\mathbf{Q}_{pk,n}\mathbf{Q}_{kk,n}^{-1}\\ -\mathbf{M}_{kk,n}^{-1}\mathbf{Q}'_{pk,n}\mathbf{Q}_{pp,n}^{-1} & \mathbf{M}_{kk,n}^{-1}\end{pmatrix} \tag{2.140}$$

Then using (2.23) we have that

$$\begin{aligned}&\begin{pmatrix}\mathbf{M}_{pp,n}^{-1} & -\mathbf{M}_{pp,n}^{-1}\mathbf{Q}_{pk,n}\mathbf{Q}_{kk,n}^{-1}\\ -\mathbf{M}_{kk,n}^{-1}\mathbf{Q}'_{pk,n}\mathbf{Q}_{pp,n}^{-1} & \mathbf{M}_{kk,n}^{-1}\end{pmatrix}\sum_{t=1}^{n}\begin{pmatrix}\boldsymbol{\tau}_{p,tn}\\ \boldsymbol{\eta}_{k,tn}\end{pmatrix}\triangle\mathbf{X}'_{k,t}\\ &=\begin{pmatrix}\mathbf{M}_{pp,n}^{-1} & -\mathbf{M}_{pp,n}^{-1}\mathbf{Q}_{pk,n}\mathbf{Q}_{kk,n}^{-1}\\ -\mathbf{M}_{kk,n}^{-1}\mathbf{Q}'_{pk,n}\mathbf{Q}_{pp,n}^{-1} & \mathbf{M}_{kk,n}^{-1}\end{pmatrix}\left\{\begin{pmatrix}\mathbf{Q}_{pp,n}\boldsymbol{\Gamma}_{p,n}^{-1}\boldsymbol{\Phi}'_{k,p}\\ \mathbf{Q}'_{pk,n}\boldsymbol{\Gamma}_{p,n}^{-1}\boldsymbol{\Phi}'_{k,p}\end{pmatrix}+\sum_{t=1}^{n}\begin{pmatrix}\boldsymbol{\tau}_{p,tn}\\ \boldsymbol{\eta}_{k,tn}\end{pmatrix}\boldsymbol{\varepsilon}'_{k,t}\right\}\\ &=\begin{pmatrix}\boldsymbol{\Gamma}_{p,n}^{-1}\boldsymbol{\Phi}'_{k,p}\\ \mathbf{0}_{k,k}\end{pmatrix}+\begin{pmatrix}\mathbf{M}_{pp,n}^{-1}\left\{\sum_{t=1}^{n}\boldsymbol{\tau}_{p,tn}\boldsymbol{\varepsilon}'_{k,t}-\mathbf{Q}_{pk,n}\mathbf{Q}_{kk,n}^{-1}\sum_{t=1}^{n}\boldsymbol{\eta}_{k,tn}\boldsymbol{\varepsilon}'_{k,t}\right\}\\ \mathbf{M}_{kk,n}^{-1}\left\{\sum_{t=1}^{n}\boldsymbol{\eta}_{k,tn}\boldsymbol{\varepsilon}'_{k,t}-\mathbf{Q}'_{pk,n}\mathbf{Q}_{pp,n}^{-1}\sum_{t=1}^{n}\boldsymbol{\tau}_{p,tn}\boldsymbol{\varepsilon}'_{k,t}\right\}\end{pmatrix}\end{aligned} \tag{2.141}$$

And

$$\begin{pmatrix}\mathbf{M}_{pp,n}^{-1} & -\mathbf{M}_{pp,n}^{-1}\mathbf{Q}_{pk,n}\mathbf{Q}_{kk,n}^{-1}\\ -\mathbf{M}_{kk,n}^{-1}\mathbf{Q}'_{pk,n}\mathbf{Q}_{pp,n}^{-1} & \mathbf{M}_{kk,n}^{-1}\end{pmatrix}\mathbf{W}_{n}^{-1}\begin{pmatrix}\mathbf{0}_{p+1}\\ n\boldsymbol{\Delta}_{ku}^{+}\end{pmatrix}=\sqrt{n}\begin{pmatrix}-\mathbf{M}_{pp,n}^{-1}\mathbf{Q}_{pk,n}\mathbf{Q}_{kk,n}^{-1}\boldsymbol{\Delta}_{ku}^{+}\\ \mathbf{M}_{kk,n}^{-1}\boldsymbol{\Delta}_{ku}^{+}\end{pmatrix} \tag{2.142}$$

With $\mathbf{W}_n$ given in (2.5). Taking these results together we get (2.24). Second, given the sequence of FM-OLS residuals, defined by $\hat{u}_{t,p}^{+}(k)=Y_t^{+}-(\boldsymbol{\tau}'_{p,t},\mathbf{X}'_{k,t})(\hat{\boldsymbol{\alpha}}'^{+}_{p,n},\hat{\boldsymbol{\beta}}'^{+}_{k,n})'$, with $Y_t^{+}=Y_t-(\boldsymbol{\tau}'_{p,tn}\boldsymbol{\Gamma}_{p,n}^{-1}\boldsymbol{\Phi}'_{k,p}+\boldsymbol{\varepsilon}'_{k,t})\boldsymbol{\gamma}_k$, can be written as

$$\hat{u}^{+}_{t,p}(k) = \hat{u}_{t,p}(k) - \boldsymbol{\varepsilon}'_{k,t}\boldsymbol{\gamma}_k + \boldsymbol{\tau}'_{p,tn}\mathbf{M}^{-1}_{pp,n}\left(\left\{\sum_{t=1}^{n}\boldsymbol{\tau}_{p,tn}\boldsymbol{\varepsilon}'_{k,t} - \mathbf{Q}_{pk,n}\mathbf{Q}^{-1}_{kk,n}\sum_{t=1}^{n}\boldsymbol{\eta}_{k,tn}\boldsymbol{\varepsilon}'_{k,t}\right\}\boldsymbol{\gamma}_k - \sqrt{n}\mathbf{Q}_{pk,n}\mathbf{Q}^{-1}_{kk,n}\boldsymbol{\Delta}^{+}_{ku}\right) + \boldsymbol{\eta}'_{k,tn}\mathbf{M}^{-1}_{kk,n}\left(\left\{\sum_{t=1}^{n}\boldsymbol{\eta}_{k,tn}\boldsymbol{\varepsilon}'_{k,t} - \mathbf{Q}'_{pk,n}\mathbf{Q}^{-1}_{pp,n}\sum_{t=1}^{n}\boldsymbol{\tau}_{p,tn}\boldsymbol{\varepsilon}'_{k,t}\right\}\boldsymbol{\gamma}_k + \sqrt{n}\boldsymbol{\Delta}^{+}_{ku}\right) \quad (2.143)$$

Or, in more compact form, as in (2.25) by using the equality

$$\mathbf{M}^{-1}_{pp,n}\left\{\sum_{t=1}^{n}\boldsymbol{\tau}_{p,tn}\boldsymbol{\varepsilon}'_{k,t} - \mathbf{Q}_{pk,n}\mathbf{Q}^{-1}_{kk,n}\sum_{t=1}^{n}\boldsymbol{\eta}_{k,tn}\boldsymbol{\varepsilon}'_{k,t}\right\} = \mathbf{Q}^{-1}_{pp,n}\left\{\sum_{t=1}^{n}\boldsymbol{\tau}_{p,tn}\boldsymbol{\varepsilon}'_{k,t} - \mathbf{Q}_{pk,n}\mathbf{M}^{-1}_{kk,n}\sum_{t=1}^{n}\boldsymbol{\eta}_{k,tn}\boldsymbol{\varepsilon}'_{kt,p}\right\} \quad (2.144)$$

And

$$\textstyle\sum_{t=1}^{n}\boldsymbol{\eta}_{k,tn}\boldsymbol{\varepsilon}'_{kt,p} = \sum_{t=1}^{n}\boldsymbol{\eta}_{k,tn}(\boldsymbol{\varepsilon}'_{k,t} - \boldsymbol{\tau}'_{p,tn}\mathbf{Q}^{-1}_{pp,n}\sum_{t=1}^{n}\boldsymbol{\tau}_{p,jn}\boldsymbol{\varepsilon}'_{k,j}) = \sum_{t=1}^{n}(n^{-1/2}\boldsymbol{\eta}_{kt,p})\boldsymbol{\varepsilon}'_{kt,p}. \quad (2.145)$$

■

A.3 Proof of Proposition 2.3. For proof of part (a), we have that the sequence $\boldsymbol{\xi}_{t,p}(k) = (\hat{u}_{t,p}(k), \mathbf{Z}'_{k,t})'$ can be expressed as $\boldsymbol{\xi}_{t,p}(k) = \boldsymbol{\upsilon}_{t,p}(k) + \boldsymbol{\varphi}_{t,p}(k)$, with $\boldsymbol{\upsilon}_{t,p}(k) = (\hat{u}_{t,p}(k), \boldsymbol{\varepsilon}'_{k,t})'$, and $\boldsymbol{\varphi}_{t,p}(k) = (0, \boldsymbol{\tau}'_{p,tn}\boldsymbol{\Gamma}^{-1}_{p,n}\boldsymbol{\Phi}'_{kp})'$, where $\boldsymbol{\Phi}_{kp} = \mathbf{0}_{k,p+1}$ when $p = 0$, and $\boldsymbol{\Phi}_{kp} = (\boldsymbol{\Phi}_{k,p-1} : \mathbf{0}_k)$ when $p \geq 1$. The sample autocovariance covariance matrix of order $|h| \geq 0$, $\hat{\boldsymbol{\Sigma}}_n(h) = (1/n)\sum_{t=h+1}^{n}\boldsymbol{\xi}_{t,p}(k)\boldsymbol{\xi}'_{t-h,p}(k)$, is decomposed as

$$\hat{\boldsymbol{\Sigma}}_n(h) = \boldsymbol{\Sigma}_n(h) + (1/n)\sum_{t=h+1}^{n}\boldsymbol{\varphi}_{t,p}(k)\boldsymbol{\varphi}'_{t-h,p}(k) + (1/n)\sum_{t=h+1}^{n}[\boldsymbol{\upsilon}_{t,p}(k)\boldsymbol{\varphi}'_{t-h,p}(k) + \boldsymbol{\varphi}_{t,p}(k)\boldsymbol{\upsilon}'_{t-h,p}(k)] \quad (2.146)$$

Where $\boldsymbol{\Sigma}_n(h) = (1/n)\sum_{t=h+1}^{n}\boldsymbol{\upsilon}_{t,p}(k)\boldsymbol{\upsilon}'_{t-h,p}(k) \to^{p} E[\boldsymbol{\xi}_t\boldsymbol{\xi}'_{t-h}]$ under cointegration. For $h = 0$ we have that

$$\mathbf{C}_n = (1/n)\sum_{t=1}^{n}\boldsymbol{\upsilon}_{t,p}(k)\boldsymbol{\varphi}'_{t,p}(k) = \begin{pmatrix} 0 & \mathbf{0}'_k \\ \mathbf{0}_k & n^{-1/2}\left\{n^{-1/2}\sum_{t=1}^{n}\boldsymbol{\varepsilon}_{k,t}\boldsymbol{\tau}'_{p,tn}\right\}\boldsymbol{\Gamma}^{-1}_{p,n}\boldsymbol{\Phi}'_{kp} \end{pmatrix} = \begin{pmatrix} 0 & \mathbf{0}'_k \\ \mathbf{0}_k & \bar{\mathbf{D}}'_{pk,n}(n^{-1/2}\boldsymbol{\Gamma}^{-1}_{p,n})\boldsymbol{\Phi}'_{kp} \end{pmatrix} = \begin{pmatrix} 0 & \mathbf{0}'_k \\ \mathbf{0}_k & \mathbf{C}_{kk,n} \end{pmatrix} \quad (2.147)$$

By the orthogonality condition $\sum_{t=1}^{n}\boldsymbol{\tau}_{p,tn}\hat{u}_{t,p}(k) = \mathbf{0}_{p+1}$, with $\bar{\mathbf{D}}_{pk,n} = n^{-1/2}\sum_{t=1}^{n}\boldsymbol{\tau}_{p,tn}\boldsymbol{\varepsilon}'_{k,t}$, and $\mathbf{F}_n = (1/n)\sum_{t=h+1}^{n}\boldsymbol{\varphi}_{t,p}(k)\boldsymbol{\varphi}'_{t,p}(k) = \begin{pmatrix} 0 & \mathbf{0}'_k \\ \mathbf{0}_k & \boldsymbol{\Phi}_{kp}\boldsymbol{\Gamma}^{-1}_{p,n}\bar{\mathbf{Q}}_{pp,n}\boldsymbol{\Gamma}^{-1}_{p,n}\boldsymbol{\Phi}'_{kp} \end{pmatrix} = \begin{pmatrix} 0 & \mathbf{0}'_k \\ \mathbf{0}_k & \mathbf{F}_{kk,n} \end{pmatrix}$

So that we can write $\hat{\boldsymbol{\Sigma}}_n(0) = \boldsymbol{\Sigma}_n(0) + \mathbf{C}_n + \mathbf{C}'_n + \mathbf{F}_n$. For $|h| \geq 1$, we have that $\boldsymbol{\varphi}_{t,p}(k)$ can be written as $\boldsymbol{\varphi}_{t,p}(k) = \boldsymbol{\varphi}_{t-h,p}(k) + (h/n)\mathbf{d}_{t,p}(k)$, where the last term is given by

$$\mathbf{d}_{t,p}(k)=\begin{pmatrix}0\\ \mathbf{\Phi}_{kp}\mathbf{\Gamma}_{p,n}^{-1}\Delta_h\boldsymbol{\tau}_{p,tn}\end{pmatrix} \tag{2.148}$$

With $\Delta_h\boldsymbol{\tau}_{p,tn}=(0,\boldsymbol{\tau}'_{p-1,tn}\mathbf{C}_{p,p})'+O(h/n)$, and $\mathbf{C}_{p,p}=diag(1,2,...,p)$. With this we have

$$\begin{aligned}(1/n)\sum_{t=h+1}^{n}\boldsymbol{\upsilon}_{t,p}(k)\boldsymbol{\varphi}'_{t-h,p}(k)&=(1/n)\sum_{t=1}^{n}\boldsymbol{\upsilon}_{t,p}(k)\boldsymbol{\varphi}'_{t,p}(k)\\ &-\left\{n^{-1}\sum_{t=1}^{h}\boldsymbol{\upsilon}_{t,p}(k)\boldsymbol{\varphi}'_{t,p}(k)+(h/n)n^{-1}\sum_{t=h+1}^{n}\boldsymbol{\upsilon}_{t,p}(k)\mathbf{d}'_{t,p}(k)\right\}\\ &=(1/n)\sum_{t=1}^{n}\boldsymbol{\upsilon}_{t,p}(k)\boldsymbol{\varphi}'_{t,p}(k)+O_p(h/n)\end{aligned} \tag{2.149}$$

$$\begin{aligned}(1/n)\sum_{t=h+1}^{n}\boldsymbol{\varphi}_{t,p}(k)\boldsymbol{\upsilon}'_{t-h,p}(k)&=(1/n)\sum_{t=1}^{n}\boldsymbol{\varphi}_{t,p}(k)\boldsymbol{\upsilon}'_{t,p}(k)\\ &-\left\{n^{-1}\sum_{t=n-h+1}^{n}\boldsymbol{\varphi}_{t,p}(k)\boldsymbol{\upsilon}'_{t,p}(k)-(h/n)n^{-1}\sum_{t=h+1}^{n}\mathbf{d}_{t,p}(k)\boldsymbol{\upsilon}'_{t-h,p}(k)\right\}\\ &=(1/n)\sum_{t=1}^{n}\boldsymbol{\varphi}_{t,p}(k)\boldsymbol{\upsilon}'_{t,p}(k)+O_p(h/n)\end{aligned} \tag{2.150}$$

And

$$\begin{aligned}(1/n)\sum_{t=h+1}^{n}\boldsymbol{\varphi}_{t,p}(k)\boldsymbol{\varphi}'_{t-h,p}(k)&=(1/n)\sum_{t=1}^{n}\boldsymbol{\varphi}_{t,p}(k)\boldsymbol{\varphi}'_{t,p}(k)\\ &-\left\{n^{-1}\sum_{t=1}^{h}\boldsymbol{\varphi}_{t,p}(k)\boldsymbol{\varphi}'_{t,p}(k)+(h/n)n^{-1}\sum_{t=h+1}^{n}\boldsymbol{\varphi}_{t,p}(k)\mathbf{d}'_{t,p}(k)\right\}\\ &=(1/n)\sum_{t=1}^{n}\boldsymbol{\varphi}_{t,p}(k)\boldsymbol{\varphi}'_{t,p}(k)+O_p(h/n)\end{aligned} \tag{2.151}$$

So that $\hat{\mathbf{\Sigma}}_n(h)=\mathbf{\Sigma}_n(h)+\mathbf{C}_n+\mathbf{C}'_n+\mathbf{F}_n+O_p(h/n)$. Then, the kernel estimator of the long-run covariance matrix, $\hat{\mathbf{\Omega}}_n(m_n)=\sum_{h=-(n-1)}^{n-1}w(h/m_n)\hat{\mathbf{\Sigma}}_n(h)$, is decomposed as

$$\hat{\mathbf{\Omega}}_n(m_n)=\mathbf{\Omega}_n(m_n)+m_n\bar{w}_n(m_n)(\mathbf{C}_n+\mathbf{C}'_n+\mathbf{F}_n)+\frac{m_n}{n}\sum_{h=-(n-1)}^{n-1}w(h/m_n)O_p(h/m_n) \tag{2.152}$$

Where $\mathbf{\Omega}_n(m_n)=\sum_{h=-(n-1)}^{n-1}w(h/m_n)\mathbf{\Sigma}_n(h)$, and the last term is $O_p(m_n/n)$.

For proof of part (b), we have that under cointegration $\hat{\boldsymbol{\xi}}_{t,p}(k)=\boldsymbol{\xi}_t+O_p(n^{-1/2})$, which gives $(1/n)\sum_{t=h+1}^{n}\hat{\boldsymbol{\xi}}_{t,p}(k)\hat{\boldsymbol{\xi}}'_{t-h,p}(k)=(1/n)\sum_{t=h+1}^{n}\boldsymbol{\xi}_t\boldsymbol{\xi}'_{t-h}+o_p(n^{-1/2})$, where $\boldsymbol{\xi}_t=(u_t,\boldsymbol{\varepsilon}'_{k,t})'$. This means that, under standard and suitable assumptions on the bandwidth choice, we get $\hat{\mathbf{\Omega}}_n(m_n)\rightarrow^p\mathbf{\Omega}$. Also, taking $\hat{\mathbf{Z}}_{kt,p}=\boldsymbol{\varepsilon}_{kt,p}$ as has been defined in Proposition 2.2(b), then we have that the transformed observations of the dependent variable can be decomposed as

$$Y_t^+ = Y_t - \hat{\mathbf{Z}}_{kt,p}'\hat{\gamma}_{k,n}(m_n) = u_t - \boldsymbol{\varepsilon}_{k,t}'\hat{\gamma}_{k,n}(m_n) + (\boldsymbol{\tau}_{p,tn}', n^{-1/2}\boldsymbol{\eta}_{k,t}')\mathbf{W}_n'\begin{pmatrix}\boldsymbol{\alpha}_p\\ \boldsymbol{\beta}_k\end{pmatrix} + (\boldsymbol{\tau}_{p,tn}', n^{-1/2}\boldsymbol{\eta}_{k,t}')\begin{pmatrix}\mathbf{Q}_{pp,n}^{-1}\mathbf{D}_{pk,n}\hat{\gamma}_{k,n}(m_n)\\ \mathbf{0}_k\end{pmatrix} \quad (2.153)$$

Which gives

$$n^v\mathbf{W}_n'\begin{pmatrix}\hat{\boldsymbol{\alpha}}_{p,n}^+ - \boldsymbol{\alpha}_p\\ \hat{\boldsymbol{\beta}}_{k,n}^+ - \boldsymbol{\beta}_k\end{pmatrix} = n^{-1/2+v}\begin{pmatrix}\bar{\mathbf{Q}}_{pp,n}^{-1}\bar{\mathbf{D}}_{pk,n}\hat{\gamma}_{k,n}(m_n)\\ \mathbf{0}_k\end{pmatrix} + \left((1/n)\sum_{t=1}^{n}\begin{pmatrix}\boldsymbol{\tau}_{p,tn}\\ n^{-1/2}\boldsymbol{\eta}_{k,t}\end{pmatrix}(\boldsymbol{\tau}_{p,tn}', n^{-1/2}\boldsymbol{\eta}_{k,t}')\right)^{-1} \times\left\{(1/n)\sum_{t=1}^{n}\begin{pmatrix}\boldsymbol{\tau}_{p,tn}\\ n^{-1/2}\boldsymbol{\eta}_{k,t}\end{pmatrix}(u_t - \boldsymbol{\varepsilon}_{k,t}'\hat{\gamma}_{k,n}(m_n)) - \begin{pmatrix}\mathbf{0}_{p+1}\\ \hat{\boldsymbol{\Delta}}_{ku,n}^+(m_n)\end{pmatrix}\right\} \quad (2.154)$$

Then, under cointegration and making use of the consistent estimation of $\boldsymbol{\gamma}_k$ and $\boldsymbol{\Delta}_{ku}^+$, we get the result stated in (2.33).

■

A.4 Proof of Proposition 3.2.

Making use of the results in Proposition 3.1 and the structure of the first differences of the IM-OLS residuals in equation (3.25) we have that, under cointegration, we can write

$$\tilde{z}_{t,p}(k) = (1, -(\tilde{\gamma}_{k,n} - \boldsymbol{\gamma}_k)')\begin{pmatrix}z_t\\ \boldsymbol{\varepsilon}_{k,t}\end{pmatrix} + O_p(n^{-1/2}),\ t = 1, ..., n. \quad (2.155)$$

With this, the hth-order sample serial covariance $\tilde{\kappa}_n(h) = (1/n)\sum_{t=h+1}^{n}\tilde{z}_{t,p}(k)\tilde{z}_{t-h,p}(k)$ is given by $\tilde{\kappa}_n(h) = (1, -(\tilde{\gamma}_{k,n} - \boldsymbol{\gamma}_k)')(1/n)\sum_{t=h+1}^{n}\begin{pmatrix}z_t\\ \boldsymbol{\varepsilon}_{k,t}\end{pmatrix}(z_{t-h}, \boldsymbol{\varepsilon}_{k,t-h}')\begin{pmatrix}1\\ -(\tilde{\gamma}_{k,n} - \boldsymbol{\gamma}_k)\end{pmatrix} + (1 - h/n)O_p(n^{-1/2})$ which determines that the kernel estimator of the long-run variance $\tilde{\omega}_n^2(m_n)$ can be written as

$$\tilde{\omega}_n^2(m_n) = (1, -(\tilde{\gamma}_{k,n} - \boldsymbol{\gamma}_k)')\sum_{h=-(n-1)}^{n-1} w(h/m_n)\mathbf{K}_n^+(h)\begin{pmatrix}1\\ -(\tilde{\gamma}_{k,n} - \boldsymbol{\gamma}_k)\end{pmatrix} + O_p(m_n n^{-1/2}) \quad (2.156)$$

With

$$\mathbf{K}_n^+(h) = (1/n)\sum_{t=h+1}^{n}\begin{pmatrix}z_t\\ \boldsymbol{\varepsilon}_{k,t}\end{pmatrix}(z_{t-h}, \boldsymbol{\varepsilon}_{k,t-h}') \quad (2.157)$$

So that it can be verified that

$$\boldsymbol{\Omega}_n^+(m_n) = \sum_{h=-(n-1)}^{n-1} w(h/m_n)\mathbf{K}_n^+(h) \to^p \begin{pmatrix}\omega_{u.k}^2 & \mathbf{0}_k'\\ \mathbf{0}_k & \boldsymbol{\Omega}_{kk}\end{pmatrix} \quad (2.158)$$

Finally, taking into account that $\tilde{\gamma}_{k,n} - \boldsymbol{\gamma}_k \Rightarrow \boldsymbol{\Theta}_{\gamma,k}^0 = \omega_{u.k}\boldsymbol{\Omega}_{kk}^{-1/2}\boldsymbol{\theta}_{\gamma,k}^0$ under the cointegration assumption, then it is immediate to obtain the result in (2.92) as

$$\tilde{\omega}_n^2(m_n) \Rightarrow (1, -\Theta_{\gamma,k}'^{0}) \begin{pmatrix} \omega_{u.k}^2 & \mathbf{0}_k' \\ \mathbf{0}_k & \Omega_{kk} \end{pmatrix} \begin{pmatrix} 1 \\ -\Theta_{\gamma,k}^{0} \end{pmatrix} = \omega_{u.k}^2 (1 + \theta_{\gamma,k}'^{0} \theta_{\gamma,k}^{0}) \quad (2.159)$$

On the other hand, given that under the assumption of no cointegration the sequence of first differences of the IM-OLS residuals is now given by

$$\begin{aligned} \tilde{z}_{t,p}(k) &= z_t - n\varepsilon_{k,t}'[n^{-1}(\tilde{\gamma}_{k,n} - \gamma_k)] \\ &\quad - \sqrt{n}(\tau_{p,tn}'[n^{-1/2}\Gamma_{p,n}^{-1}(\tilde{\alpha}_{p,n} - \alpha_p)] + n^{-1/2}\eta_{k,t}'(\tilde{\beta}_{k,n} - \beta_k) \end{aligned} \quad (2.160)$$

With $z_t = O_p(\sqrt{n})$, then we have that $\tilde{z}_{t,p}(k) = O_p(n)$ and hence

$$(1/n)\tilde{z}_{t,p}(k) = -\varepsilon_{k,t}'[n^{-1}(\tilde{\gamma}_{k,n} - \gamma_k)] + O_p(n^{-1/2}) \quad (2.161)$$

With this we have that

$$\begin{aligned} (1/n^2)\tilde{\kappa}_n(h) &= (1/n)\sum_{t=h+1}^{n}(n^{-1}\tilde{z}_{t,p}(k))(n^{-1}\tilde{z}_{t-h,p}(k)) \\ &= [n^{-1}(\tilde{\gamma}_{k,n} - \gamma_k)']\mathbf{G}_{kn}(h)[n^{-1}(\tilde{\gamma}_{k,n} - \gamma_k)] + (1 - h/n)O_p(n^{-1/2}) \end{aligned} \quad (2.162)$$

With

$$\mathbf{G}_{kn}(h) = (1/n)\sum_{t=h+1}^{n} \varepsilon_{k,t}\varepsilon_{k,t-h} \quad (2.163)$$

Thus, the scaled kernel estimator of the long-run variance is now given by

$$\begin{aligned} (1/n^2)\tilde{\omega}_n^2(m_n) &= [n^{-1}(\tilde{\gamma}_{k,n} - \gamma_k)'] \sum_{h=-(n-1)}^{n-1} w(h/m_n)\mathbf{G}_{kn}(h)[n^{-1}(\tilde{\gamma}_{k,n} - \gamma_k)] + O_p(m_n n^{-1/2}) \\ &= [n^{-1}(\tilde{\gamma}_{k,n} - \gamma_k)']\tilde{\Omega}_{kk,n}(m_n)[n^{-1}(\tilde{\gamma}_{k,n} - \gamma_k)] + O_p(m_n n^{-1/2}) \\ &\Rightarrow \Theta_{\gamma,k}'^{1}\Omega_{kk}\Theta_{\gamma,k}^{1} \end{aligned} \quad (2.164)$$

by making use of the convergence result $n^{-1}(\tilde{\gamma}_{k,n} - \gamma_k) \Rightarrow \Theta_{\gamma,k}^{1} = \omega_u \Omega_{k,k}^{-1/2}\theta_{\gamma,k}^{1}$, and the consistency result $\tilde{\Omega}_{kk,n}(m_n) = \sum_{h=-(n-1)}^{n-1} w(h/m_n)\mathbf{G}_{kn}(h) \to^p \Omega_{kk}$, which gives

$$(1/n^2)\tilde{\omega}_n^2(m_n) \Rightarrow \Theta_{\gamma,k}'^{1}\Omega_{kk}\Theta_{\gamma,k}^{1} = \omega_u^2\theta_{\gamma,k}'^{1}\theta_{\gamma,k}^{1}$$

■

A.5 Proof of Proposition 3.3. For the proof of parts (a) and (c), then partial summing from (2.27) gives

$$\hat{S}_{t,p} = \beta_k'\hat{\mathbf{S}}_{kt,p} + U_{t,p}, t = 1, ..., n \quad (2.165)$$

So that

$$\begin{pmatrix}\tilde{\beta}_{k,n}\\ \tilde{\gamma}_{k,n}\end{pmatrix}=\begin{pmatrix}\beta_k\\ \mathbf{0}_k\end{pmatrix}+n^{(1-v)}(\mathbf{W}_n')^{-1}\left((1/n)\sum_{t=1}^{n}\begin{pmatrix}n^{-3/2}\hat{\mathbf{S}}_{kt,p}\\ n^{-1/2}\hat{\mathbf{T}}_{kt,p}\end{pmatrix}(n^{-3/2}\hat{\mathbf{S}}_{kt,p}',n^{-1/2}\hat{\mathbf{T}}_{kt,p}')\right)^{-1} \times(1/n)\sum_{t=1}^{n}\begin{pmatrix}n^{-3/2}\hat{\mathbf{S}}_{kt,p}\\ n^{-1/2}\hat{\mathbf{T}}_{kt,p}\end{pmatrix}n^{-(1-v)}U_{t,p} \quad (2.166)$$

And thus

$$n^{-(1-v)}\mathbf{W}_n'\begin{pmatrix}\tilde{\beta}_{k,n}-\beta_k\\ \tilde{\gamma}_{k,n}\end{pmatrix}=\begin{pmatrix}n^{1/2+v}(\tilde{\beta}_{k,n}-\beta_k)\\ n^{-1/2+v}\tilde{\gamma}_{k,n}\end{pmatrix} =\left((1/n)\sum_{t=1}^{n}\begin{pmatrix}n^{-3/2}\hat{\mathbf{S}}_{kt,p}\\ n^{-1/2}\hat{\mathbf{T}}_{kt,p}\end{pmatrix}(n^{-3/2}\hat{\mathbf{S}}_{kt,p}',n^{-1/2}\hat{\mathbf{T}}_{kt,p}')\right)^{-1}\left\{(1/n)\sum_{t=1}^{n}\begin{pmatrix}n^{-3/2}\hat{\mathbf{S}}_{kt,p}\\ n^{-1/2}\hat{\mathbf{T}}_{kt,p}\end{pmatrix}n^{-(1-v)}U_{t,p}\right\} \quad (2.167)$$

Making use of the convergence results in (2.23), (2.24) and (2.107), then under the cointegration assumption, that is when $v = 1/2$, we have that

$$\begin{pmatrix}n(\tilde{\beta}_{k,n}-\beta_k)\\ \tilde{\gamma}_{k,n}\end{pmatrix}\Rightarrow \left(\int_0^1\mathbf{g}_p(r)\mathbf{g}_p(r)'dr\right)^{-1}\int_0^1\mathbf{g}_p(r)V_{u,p}(r)dr = \left(\int_0^1\mathbf{g}_p(r)\mathbf{g}_p(r)'dr\right)^{-1}\left(\int_0^1\mathbf{g}_p(r)V_{u.k,p}(r)dr+\int_0^1\mathbf{g}_p(r)\mathbf{V}_{k,p}(r)'dr\,\gamma_k\right) \quad (2.168)$$

Where the last two terms are based on the decomposition in (2.26). For the last term above, as in VW (equation (25), we can write

$$\int_0^1\mathbf{g}_p(r)\mathbf{V}_{k,p}(r)'dr\,\gamma_k=\int_0^1\mathbf{g}_p(r)\left\{\mathbf{g}_p(r)'\begin{pmatrix}\mathbf{0}_{k,k}\\ \mathbf{I}_{k,k}\end{pmatrix}\right\}dr\,\gamma_k=\int_0^1\mathbf{g}_p(r)\mathbf{g}_p(r)'dr\begin{pmatrix}\mathbf{0}_k\\ \gamma_k\end{pmatrix} \quad (2.169)$$

So that

$$\begin{pmatrix}n(\tilde{\beta}_{k,n}-\beta_k)\\ \tilde{\gamma}_{k,n}-\gamma_k\end{pmatrix}\Rightarrow\mathbf{\Pi}^{-1}\left(\int_0^1\bar{\mathbf{g}}_p(r)\bar{\mathbf{g}}_p(r)'dr\right)^{-1}\int_0^1\bar{\mathbf{g}}_p(r)V_{u.k,p}(r)dr \quad (2.170)$$

Or, equivalently,

$$\begin{pmatrix}n(\tilde{\beta}_{k,n}-\beta_k)\\ \tilde{\gamma}_{k,n}-\gamma_k\end{pmatrix}\Rightarrow\mathbf{\Pi}^{-1}\left(\int_0^1\bar{\mathbf{g}}_p(r)\bar{\mathbf{g}}_p(r)'dr\right)^{-1}\int_0^1[\mathbf{G}_p(1)-\mathbf{G}_p(r)]dV_{u.k,p}(r) \quad (2.171)$$

Where the last equality comes from defining $\mathbf{G}_p(r)=\int_0^r\mathbf{g}_p(s)ds=\mathbf{\Pi}\int_0^r\bar{\mathbf{g}}_p(s)ds$, with $\mathbf{\Pi}=diag(\mathbf{\Omega}_{k,k}^{1/2},\mathbf{\Omega}_{k,k}^{1/2})$, and $\mathbf{g}_p(r)=\mathbf{\Pi}\cdot\bar{\mathbf{g}}_p(r)$. Also, by defining $V_{u.k,p}(r)=\omega_{u.k}^2\cdot W_{u.k,p}(r)$, with $W_{u.k,p}(r)=BM(b_p(r))$, then we have

$$\begin{pmatrix}n(\tilde{\beta}_{k,n}-\beta_k)\\ \tilde{\gamma}_{k,n}-\gamma_k\end{pmatrix}\Rightarrow\omega_{u.k}\mathbf{\Pi}^{-1}\left(\int_0^1\bar{\mathbf{g}}_p(r)\bar{\mathbf{g}}_p(r)'dr\right)^{-1}\int_0^1[\mathbf{G}_p(1)-\mathbf{G}_p(r)]dW_{u.k,p}(r) \quad (2.172)$$

As in equation (24) in VW, conditional on $\mathbf{B}_k(r)$, the above limiting distribution (2.172) is $N(\mathbf{0}_{2k}, \mathbf{\Theta}_{2k})$, with $\mathbf{\Theta}_{2k}$ a well defined conditional asymptotic stochastic covariance matrix. Under no cointegration, that is, with $v = -1/2$ and nonstationarity of the error sequence u_t, then we have

$$\begin{pmatrix} \tilde{\boldsymbol{\beta}}_{k,n} - \boldsymbol{\beta}_k \\ n^{-1}\tilde{\boldsymbol{\gamma}}_{k,n} \end{pmatrix} \Rightarrow \left(\int_0^1 \mathbf{g}_p(r)\mathbf{g}_p(r)'dr\right)^{-1} \int_0^1 \mathbf{g}_p(r) J_{u,p}(r)dr \tag{2.173}$$

Where $J_{u,p}(r) = \int_0^r B_{u,p}(s)ds$. As can be seen from (2.172) and (2.173), the convergence rates for the IM-OLS estimator of $\boldsymbol{\beta}_k$ are the same as when using OLS or any of the asymptotically equivalent and efficient estimation methods.

For the proof of parts (b) and (d), then given the sequence of IM-OLS residuals in (2.112), the IM cointegrating regression equation in (2.105) and (2.167), we can write $\tilde{\zeta}_{t,p}(k)$ as

$$\tilde{\zeta}_{t,p}(k) = \zeta_{t,p} - n^{1-v}(n^{-3/2}\hat{\mathbf{S}}'_{kt,p}, n^{-1/2}\hat{\mathbf{T}}'_{kt,p})\begin{pmatrix} n^{1/2+v}(\tilde{\boldsymbol{\beta}}_{k,n} - \boldsymbol{\beta}_k) \\ n^{-1/2+v}(\tilde{\boldsymbol{\gamma}}_{k,n} - \boldsymbol{\gamma}_k) \end{pmatrix} \quad t = 1,\ldots,n \tag{2.174}$$

Under the cointegration assumption, making use of (3.22), (3.23) and the weak convergence of the IM-OLS estimators of $\boldsymbol{\beta}_k$ and $\boldsymbol{\gamma}_k$, the result (d) then follows by the continuous mapping theorem.A.6 Proof of Proposition 4.1. Part (a) follows directly from the results (a), (b) in Proposition 3.1 and the continuous mapping theorem for the functionals considered. For the proof of part (b), we use the result (c) in Proposition 3.1 with $v = -1/2$, which gives

$$n^{-3/2}\tilde{\zeta}_{t,p}(k) = n^{-3/2}\zeta_{t,p} - (n^{-3/2}\hat{\mathbf{S}}'_{kt,p}, n^{-1/2}\hat{\mathbf{T}}'_{kt,p})\begin{pmatrix} \tilde{\boldsymbol{\beta}}_{k,n} - \boldsymbol{\beta}_k \\ n^{-1}\tilde{\boldsymbol{\gamma}}_{k,n} \end{pmatrix} \tag{2.175}$$

Where

$$n^{-3/2}\zeta_{t,p} = n^{-3/2}U_{t,p} - \boldsymbol{\gamma}'_k n^{-1}(n^{-1/2}\hat{\mathbf{T}}_{kt,p}) = n^{-3/2}U_{t,p} + O_p(n^{-1}) \tag{2.176}$$

So that, using (A.6) above and the continuous mapping theorem we have that

$$n^{-3/2}\tilde{\zeta}_{t,p}(k) \Rightarrow J_{k,p}(r) = J_{u,p}(r) - \mathbf{g}(r)'\left(\int_0^1 \mathbf{g}(s)\mathbf{g}(s)'ds\right)^{-1}\int_0^1 \mathbf{g}(s)J_{u,p}(s)ds \tag{2.177}$$

With $J_{u,p}(r) = \int_0^r B_{u,p}(s)ds$ as in (2.9), which gives the desired results in the case of IM-OLS estimation of the cointegrating regression with a general polynomial trend function of order p and no deterministic component in the observations of the integrated regressors, whereas the results for the case of no deterministic component in the cointegrating regression then follow trivially by making use of the IM-OLS residuals $\tilde{\zeta}_t(k)$ and $\tilde{g}(r) = (\tilde{g}_k(r)', W_k(r)')'$ in Proposition 3.1. The extension to the case of the IM-OLS residuals from estimating the cointegrating regression with OLS detrended observations of the model variables also follows trivially employing the results in Proposition 3.3.

2.8 Appendix B. Critical values and other numerical results

B.1 Quantiles of the limit distribution of the variance ratio test statistic for testing the null of no cointegration and empirical power performance

Table 2.1 Quantiles of the null distribution of the OLS-based variance ratio test statistic by Breitung (2002), $\hat{VR}_n(p,k)$, for testing the null of no cointegration against the alternative of cointegration (left-sided test)

	No deterministic component				
Quantiles, $c_\alpha(k)$	$k = 1$	2	3	4	5
0.01	0.0054	0.0035	0.0025	0.0019	0.0016
0.025	0.0073	0.0047	0.0033	0.0025	0.0020
0.05	0.0098	0.0061	0.0043	0.0032	0.0025
0.1	0.0143	0.0088	0.0060	0.0044	0.0033
0.25	0.0280	0.0163	0.0108	0.0077	0.0057
0.50	0.0670	0.0374	0.0233	0.0160	0.0115
0.75	0.1406	0.0844	0.0533	0.0355	0.0250
0.90	0.2168	0.1486	0.0998	0.0680	0.0483
0.95	0.2558	0.1907	0.1357	0.0953	0.0686
0.975	0.2838	0.2248	0.1672	0.1228	0.0904
0.99	0.3086	0.2633	0.2077	0.1549	0.1198
	Constant term ($p = 0$)				
Quantiles, $c_\alpha(k)$	$k = 1$	2	3	4	5
0.01	0.0033	0.0023	0.0018	0.0014	0.0011
0.025	0.0043	0.0029	0.0022	0.0017	0.0014
0.05	0.0057	0.0037	0.0027	0.0021	0.0017
0.1	0.0076	0.0049	0.0035	0.0026	0.0021
0.25	0.0127	0.0078	0.0054	0.0040	0.0030
0.50	0.0231	0.0138	0.0090	0.0064	0.0048
0.75	0.0448	0.0246	0.0157	0.0108	0.0078
0.90	0.0675	0.0427	0.0262	0.0175	0.0124
0.95	0.0782	0.0559	0.0362	0.0237	0.0165
0.975	0.0848	0.0665	0.0465	0.0311	0.0213
0.99	0.0900	0.0779	0.0583	0.0413	0.0288
	Constant term and linear trend ($p = 1$)				
Quantiles, $c_\alpha(k)$	$k = 1$	2	3	4	5
0.01	0.0017	0.0013	0.0011	0.0009	0.0008
0.025	0.0021	0.0017	0.0013	0.0011	0.0009
0.05	0.0026	0.0020	0.0016	0.0013	0.0011
0.1	0.0033	0.0025	0.0020	0.0016	0.0013
0.25	0.0049	0.0037	0.0029	0.0023	0.0019
0.50	0.0075	0.0056	0.0044	0.0035	0.0028
0.75	0.0112	0.0086	0.0067	0.0053	0.0042
0.90	0.0154	0.0122	0.0096	0.0076	0.0060
0.95	0.0176	0.0146	0.0118	0.0094	0.0075
0.975	0.0192	0.0164	0.0138	0.0111	0.0090
0.99	0.0207	0.0185	0.0158	0.0133	0.0110

Table 2.2 Finite sample power at 5% nominal level of the OLS-based variance ratio test statistic by Breitung (2002), $V\hat{R}_n(p,k)$, for testing the null of no cointegration against the alternative of cointegration (left-sided test). Results based on a local-to-unity framework of analysis, with $u_t = \alpha_n u_{t-1} + \upsilon_t$, $\alpha_n = 1 - c/n$

		No deterministic component					**Constant term** ($p = 0$)					**Constant term and linear trend** ($p = 1$)				
Sample size, n		$k = 1$	2	3	4	5	$k = 1$	2	3	4	5	$k = 1$	2	3	4	5
$n = 100$	$c = 1$	0.0478	0.0528	0.0530	0.0500	0.0510	0.0630	0.0594	0.0614	0.0600	0.0660	0.0506	0.0506	0.0522	0.0510	0.0534
	5	0.1018	0.0824	0.0742	0.0716	0.0736	0.1310	0.1122	0.0910	0.1006	0.0964	0.0850	0.0786	0.0860	0.0806	0.0832
	10	0.1808	0.1544	0.1284	0.1186	0.1330	0.2454	0.2060	0.1894	0.1756	0.1674	0.1546	0.1488	0.1454	0.1272	0.1290
	20	0.3436	0.2972	0.2806	0.2718	0.2406	0.5148	0.4442	0.4410	0.3898	0.3822	0.3810	0.3438	0.3362	0.3162	0.3078
	30	0.4468	0.4480	0.4118	0.3708	0.3680	0.6700	0.6342	0.6370	0.6032	0.5662	0.6354	0.6030	0.5688	0.5648	0.5316
	40	0.5884	0.5592	0.5328	0.4890	0.4626	0.7972	0.7658	0.7354	0.7400	0.7232	0.7816	0.7488	0.7276	0.7192	0.6994
	50	0.6880	0.6404	0.6196	0.6092	0.5886	0.8654	0.8606	0.8574	0.8610	0.8420	0.8996	0.8784	0.8642	0.8560	0.8298
$n = 250$	$c = 1$	0.0532	0.0486	0.0522	0.0502	0.0474	0.0682	0.0606	0.0608	0.0582	0.0580	0.0506	0.0498	0.0522	0.0532	0.0518
	5	0.0962	0.0824	0.0774	0.0792	0.0802	0.1432	0.1242	0.1120	0.1064	0.1036	0.0844	0.0824	0.0842	0.0840	0.0740
	10	0.1674	0.1380	0.1294	0.1252	0.1256	0.2460	0.1938	0.1898	0.1818	0.1722	0.1726	0.1446	0.1446	0.1342	0.1402
	20	0.3338	0.2764	0.2626	0.2340	0.2226	0.4538	0.4316	0.3926	0.3436	0.3466	0.3700	0.3176	0.3058	0.2910	0.3098
	30	0.4478	0.4048	0.3586	0.3578	0.3418	0.6236	0.5866	0.5650	0.5352	0.5336	0.5704	0.5286	0.5212	0.4964	0.4872
	40	0.5356	0.4972	0.4798	0.4598	0.4336	0.7514	0.7216	0.7154	0.7006	0.6900	0.7274	0.7076	0.6866	0.6732	0.6492
	50	0.6316	0.5650	0.5528	0.5536	0.5258	0.8198	0.8142	0.8002	0.7878	0.7744	0.8294	0.7892	0.7872	0.7716	0.7670
$n = 500$	$c = 1$	0.0506	0.0478	0.0520	0.0512	0.0446	0.0650	0.0682	0.0626	0.0668	0.0618	0.0506	0.0562	0.0512	0.0514	0.0538
	5	0.0864	0.0792	0.0754	0.0782	0.0724	0.1470	0.1206	0.1238	0.1138	0.1010	0.0830	0.0782	0.0784	0.0748	0.0740
	10	0.1774	0.1338	0.1244	0.1228	0.1236	0.2368	0.1928	0.1844	0.1728	0.1768	0.1380	0.1396	0.1300	0.1366	0.1294
	20	0.3234	0.2840	0.2634	0.2482	0.2378	0.4498	0.4298	0.4024	0.3830	0.3628	0.3860	0.3536	0.3062	0.2866	0.2802
	30	0.4452	0.4012	0.3668	0.3426	0.3262	0.6290	0.5878	0.5738	0.5488	0.5286	0.5714	0.5288	0.4872	0.4766	0.4696
	40	0.5310	0.4864	0.4422	0.4428	0.4224	0.7462	0.7222	0.7010	0.6858	0.6658	0.7062	0.6544	0.6394	0.6396	0.6148
	50	0.5996	0.5688	0.5306	0.5076	0.4986	0.8066	0.7870	0.7874	0.7794	0.7656	0.8064	0.8090	0.7938	0.7808	0.7478
$n = 750$	$c = 1$	0.0528	0.0546	0.0520	0.0510	0.0498	0.0654	0.0632	0.0650	0.0634	0.0646	0.0498	0.0538	0.0502	0.0540	0.0530
	5	0.0950	0.0852	0.0756	0.0824	0.0802	0.1356	0.1204	0.1202	0.1186	0.1138	0.0878	0.0816	0.0840	0.0850	0.0834
	10	0.1732	0.1416	0.1408	0.1212	0.1196	0.2474	0.2394	0.2012	0.1748	0.1556	0.1614	0.1354	0.1238	0.1152	0.1138
	20	0.3128	0.2562	0.2568	0.2484	0.2372	0.4522	0.4238	0.3828	0.3714	0.3522	0.3546	0.3326	0.3064	0.2962	0.2908

30	0.4304	0.3918	0.3858	0.3480	0.3222	0.6144	0.6066	0.5722	0.5292	0.5206	0.5554	0.5068	0.4900	0.4726	0.4620
40	0.5464	0.4784	0.4626	0.4352	0.4212	0.7368	0.6966	0.6630	0.6660	0.6542	0.7054	0.6632	0.6508	0.6268	0.6086
50	0.5910	0.5536	0.5328	0.5178	0.4930	0.8116	0.7912	0.7704	0.7676	0.7606	0.8290	0.7908	0.7780	0.7530	0.7538

B.2 Quantiles of the IM-OLS based fluctuation test statistics and finite-sample power results

Table 2.3 Quantiles of the null distribution of the fluctuation test statistics with scaling factor given by the residual variance of the first differences of the IM-OLS residuals. Case of deterministically trendless integrated regressors and results from the IM-OLS regression

		No deterministic component					**Constant term** ($p = 0$)					**Constant term and linear trend** ($p = 1$)				
Statistic	Quantiles, $c_\alpha(k)$	$k = 1$	2	3	4	5	$k = 1$	2	3	4	5	$k = 1$	2	3	4	5
$\tilde{F}_{1,n}(p,k)$	0.01	0.0163	0.0114	0.0089	0.0074	0.0064	0.0129	0.0098	0.0080	0.0068	0.0059	0.0111	0.0089	0.0073	0.0064	0.0056
	0.025	0.0191	0.0130	0.0100	0.0082	0.0070	0.0150	0.0110	0.0089	0.0074	0.0065	0.0126	0.0099	0.0081	0.0070	0.0061
	0.05	0.0222	0.0145	0.0110	0.0089	0.0076	0.0170	0.0123	0.0097	0.0081	0.0070	0.0140	0.0108	0.0089	0.0076	0.0066
	0.1	0.0265	0.0167	0.0124	0.0100	0.0084	0.0196	0.0139	0.0109	0.0090	0.0077	0.0160	0.0121	0.0099	0.0083	0.0072
	0.25	0.0366	0.0216	0.0154	0.0120	0.0099	0.0258	0.0173	0.0132	0.0107	0.0090	0.0202	0.0149	0.0118	0.0098	0.0084
	0.50	0.0550	0.0292	0.0198	0.0150	0.0121	0.0356	0.0227	0.0166	0.0132	0.0109	0.0269	0.0190	0.0146	0.0119	0.0100
	0.75	0.0892	0.0413	0.0263	0.0191	0.0150	0.0511	0.0304	0.0214	0.0164	0.0133	0.0369	0.0246	0.0185	0.0146	0.0121
	0.90	0.1476	0.0581	0.0343	0.0242	0.0183	0.0732	0.0404	0.0272	0.0203	0.0162	0.0498	0.0318	0.0228	0.0179	0.0145
	0.95	0.2060	0.0745	0.0414	0.0281	0.0208	0.0908	0.0482	0.0318	0.0232	0.0181	0.0600	0.0375	0.0265	0.0203	0.0162
	0.975	0.2867	0.0929	0.0488	0.0325	0.0235	0.1115	0.0569	0.0363	0.0262	0.0201	0.0713	0.0431	0.0300	0.0228	0.0181
	0.99	0.4265	0.1261	0.0605	0.0380	0.0273	0.1431	0.0706	0.0428	0.0299	0.0229	0.0869	0.0525	0.0351	0.0262	0.0207
$\tilde{F}_{2,n}(p,k)$	0.01	0.3585	0.3098	0.2764	0.2553	0.2396	0.3263	0.2878	0.2618	0.2446	0.2312	0.3039	0.2761	0.2568	0.2389	0.2247
	0.025	0.3871	0.3289	0.2936	0.2697	0.2521	0.3492	0.3066	0.2776	0.2581	0.2428	0.3234	0.2925	0.2696	0.2504	0.2360
	0.05	0.4150	0.3474	0.3085	0.2827	0.2634	0.3697	0.3239	0.2925	0.2710	0.2531	0.3426	0.3066	0.2819	0.2622	0.2462
	0.1	0.4494	0.3727	0.3285	0.2994	0.2771	0.3987	0.3447	0.3105	0.2863	0.2674	0.3644	0.3252	0.2973	0.2766	0.2590
	0.25	0.5218	0.4198	0.3651	0.3302	0.3044	0.4509	0.3846	0.3428	0.3147	0.2925	0.4084	0.3602	0.3279	0.3024	0.2828
	0.50	0.6219	0.4858	0.4156	0.3713	0.3393	0.5228	0.4372	0.3860	0.3514	0.3243	0.4667	0.4069	0.3662	0.3372	0.3131
	0.75	0.7633	0.5707	0.4776	0.4210	0.3809	0.6134	0.5018	0.4386	0.3942	0.3626	0.5398	0.4619	0.4122	0.3766	0.3487
	0.90	0.9364	0.6665	0.5461	0.4745	0.4262	0.7162	0.5741	0.4952	0.4415	0.4030	0.6206	0.5242	0.4633	0.4203	0.3863

	0.95	1.0764	0.7426	0.5962	0.5130	0.4592	0.7892	0.6238	0.5341	0.4727	0.4308	0.6717	0.5657	0.4988	0.4490	0.4122
	0.975	1.2319	0.8172	0.6441	0.5476	0.4894	0.8575	0.6710	0.5725	0.5040	0.4553	0.7257	0.6089	0.5309	0.4764	0.4383
	0.99	1.4285	0.9302	0.7049	0.5955	0.5278	0.9471	0.7390	0.6198	0.5446	0.4888	0.7904	0.6599	0.5718	0.5105	0.4712
$\tilde{F}_{3,n}(p,k)$	0.01	0.3859	0.3299	0.2974	0.2724	0.2540	0.3515	0.3070	0.2797	0.2600	0.2457	0.3260	0.2976	0.2743	0.2535	0.2383
	0.025	0.4181	0.3558	0.3159	0.2894	0.2698	0.3792	0.3295	0.2996	0.2759	0.2592	0.3515	0.3151	0.2894	0.2682	0.2523
	0.05	0.4503	0.3784	0.3343	0.3047	0.2836	0.4053	0.3498	0.3167	0.2915	0.2732	0.3717	0.3319	0.3047	0.2836	0.2653
	0.1	0.4934	0.4092	0.3598	0.3260	0.3020	0.4384	0.3767	0.3395	0.3115	0.2898	0.4001	0.3557	0.3249	0.3013	0.2821
	0.25	0.5797	0.4704	0.4089	0.3673	0.3373	0.5065	0.4296	0.3822	0.3494	0.3239	0.4588	0.4020	0.3646	0.3368	0.3137
	0.50	0.7012	0.5596	0.4782	0.4265	0.3886	0.6065	0.5042	0.4439	0.4018	0.3716	0.5408	0.4681	0.4212	0.3860	0.3584
	0.75	0.8603	0.6748	0.5701	0.5045	0.4564	0.7358	0.6013	0.5241	0.4730	0.4345	0.6479	0.5553	0.4939	0.4508	0.4166
	0.90	1.0578	0.8159	0.6796	0.5952	0.5338	0.8844	0.7158	0.6214	0.5554	0.5097	0.7771	0.6588	0.5820	0.5281	0.4875
	0.95	1.2031	0.9142	0.7603	0.6588	0.5899	0.9872	0.7957	0.6866	0.6125	0.5619	0.8681	0.7342	0.6479	0.5829	0.5377
	0.975	1.3674	1.0089	0.8348	0.7254	0.6416	1.0776	0.8809	0.7535	0.6752	0.6118	0.9559	0.8060	0.7024	0.6360	0.5838
	0.99	1.5734	1.1525	0.9379	0.8092	0.7179	1.2116	0.9821	0.8375	0.7473	0.6752	1.0620	0.8862	0.7761	0.7079	0.6432

Table 2.4 Quantiles of the null distribution of the fluctuation test statistics with scaling factor given by the residual variance estimator based on the first difference of IM-OLS residuals. Case of deterministically trending integrated regressors and results from the IM-OLS regression based on OLS detrended observations

Statistic	Quantiles $c_\alpha(k)$	**Constant term** ($p = 0$) $k = 1$	2	3	4	5	**Constant term and linear trend** ($p = 1$) $k = 1$	2	3	4	5
$\tilde{F}_{1,n}(p,k)$	0.01	0.0146	0.0107	0.0083	0.0071	0.0062	0.0118	0.0092	0.0076	0.0066	0.0058
	0.025	0.0167	0.0119	0.0093	0.0078	0.0067	0.0135	0.0103	0.0085	0.0072	0.0063
	0.05	0.0188	0.0132	0.0103	0.0085	0.0072	0.0154	0.0115	0.0093	0.0078	0.0069
	0.1	0.0221	0.0152	0.0115	0.0094	0.0080	0.0176	0.0129	0.0103	0.0087	0.0075
	0.25	0.0296	0.0192	0.0141	0.0113	0.0094	0.0226	0.0161	0.0125	0.0103	0.0087
	0.50	0.0424	0.0254	0.0180	0.0140	0.0115	0.0306	0.0207	0.0156	0.0126	0.0104
	0.75	0.0629	0.0344	0.0233	0.0176	0.0141	0.0427	0.0274	0.0199	0.0156	0.0128
	0.90	0.0940	0.0464	0.0301	0.0219	0.0171	0.0592	0.0360	0.0251	0.0191	0.0154
	0.95	0.1229	0.0562	0.0359	0.0250	0.0193	0.0734	0.0428	0.0290	0.0218	0.0173
	0.975	0.1579	0.0684	0.0405	0.0286	0.0217	0.0878	0.0506	0.0332	0.0245	0.0193
	0.99	0.2061	0.0848	0.0494	0.0334	0.0247	0.1104	0.0616	0.0390	0.0280	0.0219
$\tilde{F}_{2,n}(p,k)$	0.01	0.3415	0.2993	0.2727	0.2498	0.2352	0.3172	0.2831	0.2605	0.2432	0.2297
	0.025	0.3635	0.3177	0.2868	0.2646	0.2472	0.3355	0.3000	0.2743	0.2563	0.2415
	0.05	0.3871	0.3351	0.3007	0.2759	0.2579	0.3556	0.3145	0.2869	0.2668	0.2513
	0.1	0.4156	0.3565	0.3183	0.2920	0.2715	0.3779	0.3335	0.3032	0.2809	0.2644
	0.25	0.4735	0.3988	0.3517	0.3211	0.2970	0.4266	0.3705	0.3342	0.3085	0.2879
	0.50	0.5546	0.4557	0.3968	0.3590	0.3300	0.4884	0.4201	0.3754	0.3437	0.3188
	0.75	0.6599	0.5243	0.4511	0.4045	0.3700	0.5670	0.4805	0.4240	0.3848	0.3551
	0.90	0.7807	0.6047	0.5120	0.4511	0.4104	0.6541	0.5457	0.4760	0.4283	0.3935
	0.95	0.8637	0.6603	0.5490	0.4831	0.4384	0.7160	0.5908	0.5104	0.4562	0.4214
	0.975	0.9487	0.7108	0.5876	0.5140	0.4660	0.7739	0.6301	0.5499	0.4869	0.4451
	0.99	1.0635	0.7835	0.6423	0.5621	0.5018	0.8391	0.6902	0.5895	0.5245	0.4791
$\tilde{F}_{3,n}(p,k)$	0.01	0.3415	0.2993	0.2727	0.2498	0.2352	0.3172	0.2831	0.2605	0.2432	0.2297
	0.025	0.3635	0.3177	0.2868	0.2646	0.2472	0.3355	0.3000	0.2743	0.2563	0.2415
	0.05	0.3871	0.3351	0.3007	0.2759	0.2579	0.3556	0.3145	0.2869	0.2668	0.2513
	0.1	0.4156	0.3565	0.3183	0.2920	0.2715	0.3779	0.3335	0.3032	0.2809	0.2644
	0.25	0.4735	0.3988	0.3517	0.3211	0.2970	0.4266	0.3705	0.3342	0.3085	0.2879
	0.50	0.5546	0.4557	0.3968	0.3590	0.3300	0.4884	0.4201	0.3754	0.3437	0.3188
	0.75	0.6599	0.5243	0.4511	0.4045	0.3700	0.5670	0.4805	0.4240	0.3848	0.3551
	0.90	0.7807	0.6047	0.5120	0.4511	0.4104	0.6541	0.5457	0.4760	0.4283	0.3935
	0.95	0.8637	0.6603	0.5490	0.4831	0.4384	0.7160	0.5908	0.5104	0.4562	0.4214
	0.975	0.9487	0.7108	0.5876	0.5140	0.4660	0.7739	0.6301	0.5499	0.4869	0.4451
	0.99	1.0635	0.7835	0.6423	0.5621	0.5018	0.8391	0.6902	0.5895	0.5245	0.4791

Table 2.5 Finite-sample power of the test statistic $\tilde{F}_{1,n}(p,k)$ deterministically trendless integrated regressors, at the 5% nominal level, with a nonparametric kernel estimator of the CLRV based on first differences of IM-OLS residuals based on the Bartlett kernel and bandwidth $m_n = [d(n/100)^{1/4}]$, $d = 12$

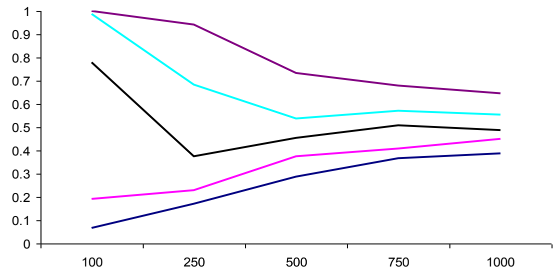
		Sample size, n 100	250	500	750	1000
No deterministics	$k = 1$	0.0678	0.1702	0.2888	0.3650	0.3894
	2	0.1932	0.2300	0.3738	0.4076	0.4486
	3	0.7804	0.3754	0.4528	0.5064	0.4876
	4	0.9884	0.6846	0.5386	0.5694	0.5534
	5	0.9998	0.9434	0.7314	0.6808	0.6460
Case $p = 0$	$k = 1$	0.2120	0.3874	0.5234	0.6108	0.6456
	2	0.5548	0.4322	0.5164	0.5942	0.6096
	3	0.9578	0.6714	0.58100	0.6340	0.6498
	4	0.9990	0.9090	0.7362	0.6848	0.7118

		5	1.0000	0.9898	0.8588	0.7906	0.7970
Case p = 1		k = 1	0.3006	0.4066	0.6026	0.7138	0.7464
		2	0.8380	0.4848	0.5740	0.6534	0.6656
		3	0.9906	0.7410	0.6716	0.6788	0.7030
		4	0.9998	0.9544	0.7850	0.7310	0.7580
		5	1.0000	0.9960	0.9066	0.8590	0.8240

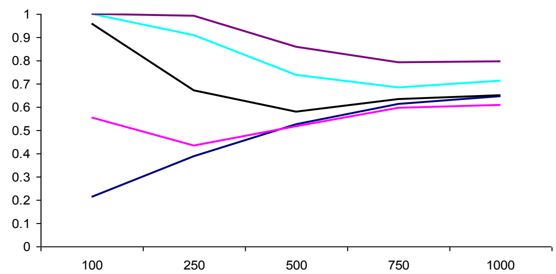
Table 2.6 Finite-sample power of the test statistic $\tilde{F}_{1,n}(p,k)$ with deterministically trending integrated regressors, at the 5% nominal level, with nonparametric kernel estimator of the CLRV based on the first difference of the IM-OLS residuals, Bartlett kernel and bandwidth m_n = $[d(n/100)^{1/4}]$

		d	Sample size, n 100	250	500	750	1000
Case p =0	k = 1	1	0.5985	0.6600	0.6705	0.6805	0.6605
		4	0.4495	0.5620	0.6455	0.6250	0.6270
		12	0.2546	0.3634	0.5122	0.5662	0.5798
	2	1	0.5620	0.6570	0.6190	0.6390	0.6385
		4	0.4460	0.5465	0.6235	0.6095	0.6065
		12	0.5896	0.4344	0.4926	0.5504	0.5708
	3	1	0.6000	0.6465	0.6605	0.6425	0.6705
		4	0.5600	0.5240	0.6320	0.6655	0.6325
		12	1.0000	0.6522	0.5504	0.6116	0.5946
	4	1	0.6480	0.6810	0.7065	0.6935	0.6990
		4	0.7390	0.6240	0.6705	0.6895	0.6930
		12	1.0000	0.9440	0.7106	0.6956	0.6656
	5	1	0.6825	0.6935	0.7355	0.7395	0.7330
		4	0.9320	0.7145	0.6985	0.7300	0.7300
		12	1.0000	1.0000	0.8590	0.7834	0.7492
Case p =1	k = 1	1	0.7155	0.7565	0.7905	0.7800	0.7985
		4	0.4765	0.6770	0.7490	0.7845	0.7795
		12	0.3470	0.4666	0.6180	0.6784	0.6960
	2	1	0.6365	0.7140	0.7355	0.7280	0.7390
		4	0.5360	0.6370	0.6905	0.7455	0.7325
		12	0.9286	0.5276	0.5892	0.6312	0.6614
	3	1	0.6210	0.7130	0.7445	0.7545	0.7410
		4	0.6565	0.6560	0.7130	0.7520	0.7430
		12	1.0000	0.8208	0.6582	0.6728	0.6712
	4	1	0.6540	0.7425	0.7625	0.7670	0.7405
		4	0.8555	0.7435	0.7675	0.7645	0.7590
		12	1.0000	0.9894	0.8002	0.7326	0.7404
	5	1	0.6975	0.7720	0.7690	0.7725	0.7710
		4	0.9805	0.8285	0.7820	0.7810	0.7810
		12	1.0000	1.0000	0.9300	0.8368	0.8148

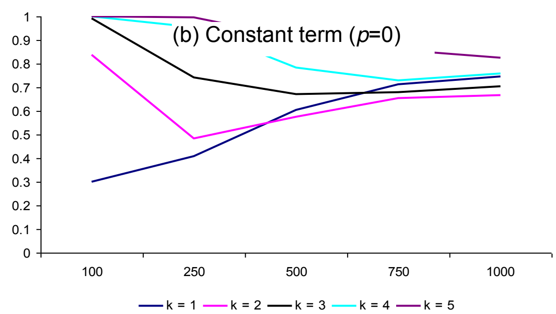
Figure 2.1 Finite-sample power of $\tilde{F}_{1,n}(p,k)$ with deterministically trendless integrated regressors



(a) No deterministics

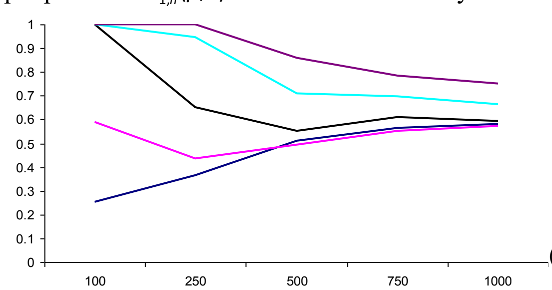


(b) Constant term (p=0)

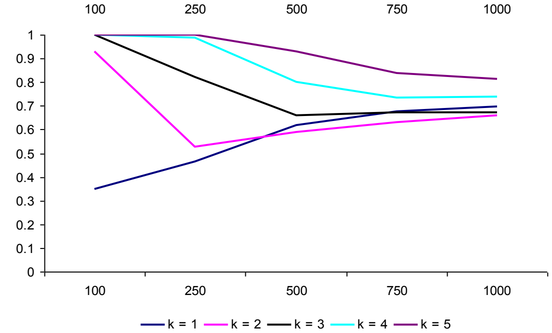


(c) Constant term and linear trend (p=1)

Figure 2.2 Finite-sample power of $\tilde{F}_{1,n}(p,k)$ with deterministically trending integrated regressors



(a) Constant term (p=0)



(b) Constant term and linear trend (p=1)

Table 2.7 Finite-sample power of the test statistic $\hat{F}_{1,n}(p,k)$ with deterministically trendless integrated regressors, at the 5% nominal level, with nonparametric kernel estimator of the CLRV based on OLS residuals, Bartlett kernel and bandwidth $m_n = [d(n/100)^{1/4}]$, $d = 12$

		Sample size, n				
		100	250	500	750	1000
No deterministics	$k = 1$	0.1810	0.3354	0.5324	0.6570	0.7348
	2	0.1396	0.3270	0.6354	0.7332	0.8266
	3	0.1120	0.3148	0.6338	0.7762	0.8548
	4	0.0840	0.2712	0.6092	0.7674	0.8674
	5	0.0770	0.2166	0.5558	0.7490	0.8620
Case $p = 0$	$k = 1$	0.2396	0.5388	0.8252	0.9088	0.9444
	2	0.1640	0.4612	0.7908	0.9006	0.9370
	3	0.1472	0.4144	0.7294	0.8808	0.9364
	4	0.1306	0.3466	0.6686	0.8456	0.9304
	5	0.1528	0.2910	0.6324	0.8226	0.9014
Case $p = 1$	$k = 1$	0.2024	0.5216	0.8184	0.9192	0.9574
	2	0.1478	0.4354	0.7602	0.8828	0.9366
	3	0.1250	0.3836	0.7218	0.8648	0.9322
	4	0.1546	0.3046	0.6606	0.8438	0.9250
	5	0.1916	0.2526	0.6158	0.8086	0.8976

Table 2.8 Finite-sample power of the test statistic $\hat{F}_{1,n}(p,k)$ with deterministically trending integrated regressors, at the 5% nominal level, with nonparametric kernel estimator of the CLRV based on OLS residuals, Bartlett kernel and bandwidth $m_n = [d(n/100)^{1/4}]$, $d = 12$

		Sample size, n				
		100	250	500	750	1000
Case $p = 0$	$k = 1$	0.2704	0.5786	0.7970	0.8882	0.9410
	2	0.2230	0.4998	0.7952	0.8926	0.9426
	3	0.1596	0.3970	0.7596	0.8866	0.9298
	4	0.1290	0.3560	0.7042	0.8696	0.9168
	5	0.1674	0.2936	0.6440	0.8288	0.9118
Case $p = 1$	$k = 1$	0.2640	0.5726	0.8272	0.9118	0.9550
	2	0.1750	0.4746	0.7616	0.8878	0.9362
	3	0.1928	0.4138	0.7290	0.8688	0.9240
	4	0.2152	0.3504	0.6736	0.8458	0.9188
	5	0.2290	0.3262	0.6118	0.8200	0.9076