Article **ECORFAN Journal-Mexico**
**OPTIMIZATION** December 2013 Vol.4 No.11 887-908

# Asymptotic and finite-sample properties of a new simple estimator of cointegrating regressions under near cointegration

AFONSO- Julio*†

*University of La Laguna, Department of Institutional Economics, Economic Statistics and Econometrics, Faculty of Economics and Business Administration.*

Received October 06, 2012; Accepted May 31, 2013

Asymptotically efficient estimation of a static cointegrating regression represents a critical requirement for later development of valid inferential procedures. Existing methods, such as fully-modified ordinary least-squares (FM-OLS), canonical cointegrating regression (CCR), or dynamic OLS (DOLS), that are asymptotically equivalent, require the choice of several tuning parameters to perform parametric or nonparametric correction of the two sources of bias that contaminate the limiting distribution of the OLS estimates and residuals. The so-called Integrated Modified OLS (IM-OLS) estimation method, recently proposed by Vogelsang and Wagner (2011), avoids these inconveniencies through a simple transformation (integration) of the system variables in the cointegrating regression equation, so that it represents a very appealing alternative estimation procedure that produces asymptotically almost efficient estimates of the model parameter. In this paper we study the performance of this estimator, both asymptotically and in finite samples, in the case of near cointegration when mechanism generating the error term of the cointegrating regression equation represents a certain generalization of the I(0) assumption in the standard case. Particularly, we consider three different specifications for the error term that generate a stationary sequence with finite variance in large samples, but are nonstationary for small sample sizes, and a fourth specification known as a stochastically trendless process that represents an intermediate situation between ordinary stationarity and nonstationarity and that determines what has been termed as stochastic cointegration. With this, we characterize the limiting distribution of the IM-OLS estimator, determining the main differences with respect the reference case of stationary cointegration, and evaluate its performance in finite samples as measured by bias and root mean squared error through a small simulation experiment.

**Cointegrating regression, asymptotically efficient estimation, integrated trending regressor, near cointegration, stochastic cointegration**

**Citation:** Afonso J. Asymptotic and finite-sample properties of a new simple estimator of cointegrating regressions under near cointegration. ECORFAN Journal -Mexico 2013, 4-11:857-878

*Correspondence to Author (email: jafonsor@ull.es)
† Researcher contributing first author.

December 2013 Vol.4 No.11 887-908

## Introduction

Since the seminal work of Engle and Granger (1987), theoretical and empirical analysis of cointegrating regressions have become a commonly used tool for analyzing integrated variables. The structure of the integrated variables, and in particular that of the regressors, plays an important role in determining the distributional properties of the estimators in these regression equations. It is also relevant to consider the role of the stochastic properties of the error term in the cointegrating regression model, particularly when we consider that can follows a highly persistent but stationary process. In any of these situations, the usefulness and optimality properties of some existing estimation methods could be questioned. Another characteristic of the regressors, many times not considered, is when they contain some deterministic component and it is not explicitly taken into account in specifying the cointegrating regression model and in determining the limiting distribution of these estimators, as has been indicated by Hansen (1992a).

Given that the use of the basic OLS estimator presents serious problems in many of the most important practical situations, particularly under endogeneity of the regressors and serially correlated error terms, there has been proposed a number of alternative estimation procedures whose main disadvantage is the need to make some choices on tuning parameters that are fundamental to their implementation. Recently, Vogelsang and Wagner (2011) have proposed a very simple alternative procedure, the integrated-modified OLS (IM-OLS) estimator, that seems to work as well as the other procedures when consider a standard framework of analysis.

In this paper we are interested in exploring the performance of this new estimator under a no standard framework when the error term of the cointegrating regression model is perturbed in several ways.

In this paper we derive the limiting distribution of the OLS and IM-OLS estimators under this no standard situations, and also perform a simulation experiment to evaluate their behavior in small samples, with particular attention to the small sample bias induced by the parameters characterizing the behavior of the error term.

## The model, assumptions and preliminary results

We assume that the observed time series $Y_t$ and $\mathbf{X}_{k,t}$, with $\mathbf{X}_{k,t}$ a $k$-dimensional vector with $k \geq 1$, are generate according to the following unobserved components model

$$\begin{pmatrix} Y_t \\ \mathbf{X}_{k,t} \end{pmatrix} = \begin{pmatrix} d_{0,t} \\ \mathbf{d}_{k,t} \end{pmatrix} + \begin{pmatrix} \eta_{0,t} \\ \boldsymbol{\eta}_{k,t} \end{pmatrix} \qquad (1)$$

Where $(d_{0,t}, \mathbf{d}_{k,t}')'$, with $\mathbf{d}_{k,t} = (d_{1,t}, \ldots, d_{k,t})'$, is the deterministic component of each series, and $(\eta_{0,t}, \boldsymbol{\eta}_{k,t}')'$ is the zero mean stochastic trend component. It is assumed that $(\eta_{0,t}, \boldsymbol{\eta}_{k,t}')'$ is generated by the potentially cointegrated triangular system

$$\eta_{0,t} = \boldsymbol{\beta}_k' \boldsymbol{\eta}_{k,t} + u_t \qquad (2)$$

$$\Delta \boldsymbol{\eta}_{k,t} = \boldsymbol{\varepsilon}_{k,t} \qquad (3)$$

By combining (1) and (2) we get the following relation

ISSN-Print: 2007-1582- ISSN-On line: 2007-3682
ECORFAN® All rights reserved.

$$Y_t = (d_{0,t} - \beta_k' \mathbf{d}_{k,t}) + \beta_k' \mathbf{X}_{k,t} + u_t \qquad (4)$$

With $\mathbf{c}_k = (1, -\beta_k')'$ the unknown cointegrating vector. Next, in order to complete the specification of the cointegrating regression equation (4) we introduce a very general assumption on the structure of the nonstochastic time trends $(d_{0,t}, \mathbf{d}_{k,t}')'$.

Assumption 2.1. *Deterministic trend components*

We assume that $d_{i,t} = \alpha_{i,p_i}' \tau_{p_i,t}$, with $\alpha_{i,p_i}$ a $(p_i + 1) \times 1$ vector of trend coefficients, with $\tau_{p_i,t} = (1, t, \ldots, t^{p_i})'$, $p_i \geq 0$, for each $i = 0, 1, \ldots, k$. By defining $p = \max(p_0, p_1, \ldots, p_k)$, then we can write $d_{i,t} = \alpha_{i,p}' \tau_{p,t}$, with $\alpha_{i,p} = (\alpha_{i,p_i}' : \mathbf{0}_{p-p_i}')'$, and $\tau_{p,t} = (\tau_{p_i,t}' : \tau_{p-p_i,t}')'$, so that $\mathbf{d}_{k,t} = \mathbf{A}_{k,p} \tau_{p,t}$, where $\mathbf{A}_{k,p} = (\alpha_{1,p}, \ldots, \alpha_{k,p})'$.

Under this assumption 2.1, we get the following standard specification of the cointegrating regression model

$$Y_t = \alpha_p' \tau_{p,t} + \beta_k' \mathbf{X}_{k,t} + u_t \qquad (5)$$

Where $\alpha_p = \alpha_{0,p} - \mathbf{A}_{k,p}' \beta_k$. With this choice for the order of the polynomial trend function, we ensure that the OLS estimator of $\beta_k$ and the OLS residuals are free of the trend parameters $\mathbf{A}_{k,p}$. Taking into account that the vector of trending regressors in (5), $\mathbf{m}_t = (\tau_{p,t}', \mathbf{X}_{k,t}')$, can be decomposed as

$$\mathbf{m}_t = \begin{pmatrix} \Gamma_{p,n}^{-1} \tau_{p,tn} \\ \mathbf{A}_{k,p} \Gamma_{p,n}^{-1} \tau_{p,tn} + \eta_{k,t} \end{pmatrix} = \begin{pmatrix} \Gamma_{p,n}^{-1} & \mathbf{0}_{p+1,k} \\ \mathbf{A}_{k,p} \Gamma_{p,n}^{-1} & \sqrt{n}\mathbf{I}_{k,k} \end{pmatrix} \begin{pmatrix} \tau_{p,tn} \\ \eta_{k,tn} \end{pmatrix} = \mathbf{W}_n \mathbf{m}_{t,n} \qquad (6)$$

Where $\eta_{k,tn} = n^{-1/2} \eta_{k,t}$, with $\mathbf{W}_n$ a $(p+1+k)\times(p+1+k)$ nonstochastic and non-singular weighting matrix, where $\tau_{p,[nr]n} = \Gamma_{p,n} \tau_{p,[nr]} \rightarrow \tau_p(r) = (1, r, \ldots, r^p)'$, and $\Gamma_{p,n} = diag(1, n^{-1}, \ldots, n^{-p})$, then $\mathbf{m}_{t,n} = (\tau_{p,tn}', \eta_{k,tn}')'$ is stochastically bounded for $t = [nr]$ as $n \rightarrow \infty$, such as $\mathbf{m}_{[nr],n} \Rightarrow \mathbf{m}(r) = (\tau_p'(r), \mathbf{B}_k'(r))'$, with $\mathbf{m}(r)$ a full-ranked process in the sense that $\int_0^1 \mathbf{m}(r)\mathbf{m}'(r)dr > 0$ a.s. Thus, given the OLS estimator of the parameter vectors in (2.5), $(\hat{\alpha}_{p,n}', \hat{\beta}_{k,n}')'$, the scaled and normalized OLS estimation error, $\hat{\Theta}_n = (\hat{\Theta}_{p,n}', \hat{\Theta}_{k,n}')'$, can be represented as

$$\hat{\Theta}_n = n^v \mathbf{W}_n \begin{pmatrix} \hat{\alpha}_{p,n} - \alpha_p \\ \hat{\beta}_{k,n} - \beta_k \end{pmatrix} = \begin{pmatrix} n^v \Gamma_{p,n}^{-1}[(\hat{\alpha}_{p,n} - \alpha_p) + \mathbf{A}_{k,p}'(\hat{\beta}_{k,n} - \beta_k)] \\ n^{1/2+v}(\hat{\beta}_{k,n} - \beta_k) \end{pmatrix}$$
$$= \left( (1/n) \sum_{t=1}^{n} \mathbf{m}_{t,n} \mathbf{m}_{t,n}' \right)^{-1} n^{-(1-v)} \sum_{t=1}^{n} \mathbf{m}_{t,n} u_t \qquad (7)$$

Where the exponent $v$ will take different values depending on the stochastic properties of the cointegrating error term, $u_t$, as will be stated later. Besides the assumptions concerning the deterministic trend components of the observed time series, in order to complete the usual specification of the cointegrating regression and to obtain the limiting results characterizing the OLS estimators and residuals in the standard cases analyzed in the literature, we introduce the following assumption concerning the behavior of the error components $u_t$ and $\varepsilon_{k,t}$ in (2) and (3). In this case, we assume that the cointegrating error sequence $u_t$ is driven by a particular function of an underlying error sequence $\upsilon_t$ that we describe as follows.


ECORFAN® All rights reserved.

Assumption 2.2. *Error components*. It is assumed that $\zeta_t = (\upsilon_t, \varepsilon'_{k,t})'$ is a zero mean covariance stationary process that satisfy sufficient regularity conditions to verify the following multivariate invariance principle such that

$$\mathbf{B}_n(r) = \begin{pmatrix} B_{n,\upsilon}(r) \\ \mathbf{B}_{n,k}(r) \end{pmatrix} = n^{-1/2}\sum_{t=1}^{[nr]} \zeta_t \Rightarrow \mathbf{B}_{k+1}(r) = (B_\upsilon(r), \mathbf{B}_k(r)')' \qquad (8)$$

Where $\mathbf{B}_{k+1}(r) = \mathrm{BM}_{k+1}(\mathbf{\Omega})$ is a $k$+1-dimensional Brownian process with covariance matrix $\mathbf{\Omega}$ such that, $\mathbf{B}_{k+1}(r) = \mathbf{\Omega}^{1/2}\mathbf{W}_{k+1}(r)$, and $\mathbf{W}_{k+1}(r) = (W_{\upsilon.k}(r), \mathbf{W}_k(r)')'$, with $W_{\upsilon.k}(r)$ and $\mathbf{W}_k(r)$ two standard independent Wiener processes, and $\mathbf{\Omega}$ a positive definite covariance matrix.[25] The covariance matrix $\mathbf{\Omega}$ is given by the long-run covariance matrix of the sequence $\zeta_t$,

$$\mathbf{\Omega} = \begin{pmatrix} \omega_\upsilon^2 & \boldsymbol{\omega}'_{k\upsilon} \\ \boldsymbol{\omega}_{k\upsilon} & \mathbf{\Omega}_{kk} \end{pmatrix} = \lim_{n\to\infty} n^{-1}\sum_{t=1}^{n}\sum_{s=1}^{n} E[\zeta_t \zeta'_s] = \Delta + \Lambda'$$

Where $\Delta$ is the one-sided long-run covariance matrix defined as

$$\Delta = \Sigma + \Lambda = \lim_{n\to\infty} n^{-1}\sum_{t=1}^{n}\sum_{s=1}^{t} E[\zeta_s \zeta'_t] = \begin{pmatrix} \delta_{\upsilon\upsilon} & \Delta_{\upsilon k} \\ \Delta_{k\upsilon} & \Delta_{kk} \end{pmatrix}$$

With

$$\Sigma = E[\zeta_t \zeta'_t] = \begin{pmatrix} \sigma_\upsilon^2 & \sigma_{\upsilon k} \\ \sigma_{k\upsilon} & \Sigma_{kk} \end{pmatrix}$$

The short-run covariance matrix, and

$$\Lambda = \lim_{n\to\infty} n^{-1}\sum_{t=2}^{n}\sum_{s=1}^{t-1} E[\zeta_s \zeta'_t] = \begin{pmatrix} \lambda_{\upsilon\upsilon} & \Lambda_{\upsilon k} \\ \Lambda_{k\upsilon} & \Lambda_{kk} \end{pmatrix}$$

Making use of the upper triangular Cholesky decomposition of $\mathbf{\Omega}$ we have that $B_\upsilon(r) = B_{\upsilon.k}(r) + \boldsymbol{\omega}'_{k\upsilon}\mathbf{\Omega}_{kk}^{-1}\mathbf{B}_k(r)$, with $B_{\upsilon.k}(r) = \omega_{\upsilon.k}W_{\upsilon.k}(r)$, and $\omega^2_{\upsilon.k} = \omega^2_\upsilon - \boldsymbol{\omega}'_{k\upsilon}\mathbf{\Omega}_{kk}^{-1}\boldsymbol{\omega}_{k\upsilon}$ the conditional long-run variance of $B_{\upsilon.k}(r)$, $\omega^2_{\upsilon.k} = E[B_{\upsilon.k}(r)^2] = E[B_{\upsilon.k}(r)B_\upsilon(r)]$, where $B_{\upsilon.k}(r)$ and $\mathbf{B}_k(r)$ are independent, that is, $E[\mathbf{B}_k(r)B_{\upsilon.k}(r)] = \mathbf{0}_k$.

The assumption that $\mathbf{\Omega}$ is positive definite is a standard, but important, regularity condition which implies that $\eta_{k,t}$ (and hence $\mathbf{X}_{k,t}$) is a non-cointegrated integrated process (no subcointegration) and rules out multicointegration under a stable long-run relation between $Y_t$ and $\mathbf{X}_{k,t}$. For the initial values $\upsilon_0$ and $\eta_{k,0}$, we assume the sufficiently general conditions $\upsilon_0 = O_p(1)$, and $\eta_{k,0} = o_p(n^{1/2})$, which include the particular case of constant finite values.

Among all the elements described above, the off-diagonal $k\times 1$ vector $\Delta_{k\upsilon}$ in the one-sided long-run covariance matrix is of particular relevance in determining de limiting behavior of the OLS estimator in (7) under standard stationary cointegration, that is, when the long-run equilibrium error is stable. In this case, when $u_t = \upsilon_t$ or, more generally, when $u_t$ is any stationary transformation of $\upsilon_t$, such as $u_t = \phi u_{t-1} + \upsilon_t$ with $|\phi| < 1$ and fixed, it is well known that the key component determining the limiting distribution of the OLS estimator of the cointegrating vector $\beta_k$ is given, from (7) with $v$ = 1/2, by

$$n^{-1/2}\sum_{t=1}^{n} (n^{-1/2}\eta_{k,t})u_t \Rightarrow \mathbf{G}_{ku} + \Delta_{ku}, \qquad (9)$$

[25] This assumption is imposed, rather than derive from more primitive assumption, since it is a standard result that holds under general conditions, such as a linear process driven by an iid or martingale difference sequence as in Phillips and Solo (1992).

ISSN-Print: 2007-1582- ISSN-On line: 2007-3682
ECORFAN® All rights reserved.

With $B_u(r) = (1-\phi)^{-1}B_\upsilon(r)$,

$$\mathbf{G}_{ku} = \int_0^1 \mathbf{B}_k(s)dB_u(s) = (1-\phi)^{-1}\int_0^1 \mathbf{B}_k(s)d(B_{\upsilon.k}(r) + \boldsymbol{\omega}'_{k\upsilon}\boldsymbol{\Omega}_{kk}^{-1}\mathbf{B}_k(r))$$

And

$$\begin{aligned}\Delta_{ku,n} &= (1/n)\sum_{t=1}^{n} E[\eta_{k,t}u_t] = E\left[(\eta_{k,0}/\sqrt{n})(1/\sqrt{n})\sum_{t=1}^{n} u_t\right] + (1/n)\sum_{t=1}^{n}\sum_{j=1}^{t} E[\varepsilon_{k,t}u_t] \\ &= \sum_{j=0}^{n-1}\left[(1/n)\sum_{t=j+1}^{n} E[\varepsilon_{k,t-j}u_t]\right] + E[o_p(1)\cdot B_{n,u}(1)] \xrightarrow{p} \Delta_{ku} = \sigma_{ku} + \Lambda_{ku}\end{aligned}$$

Where $\Delta_{ku} = (1-\phi)^{-1}(\Delta_{k\upsilon} + \sum_{j=1}^{\infty}\phi^j E[\varepsilon_{k,t}\upsilon_{t-j}])$, and $\Delta_{k\upsilon} = \sigma_{k\upsilon} + \Lambda_{k\upsilon}$. In this case, the OLS estimator is consistent at the rate *n* (superconsistent), but under endogeneity of the regressors the vector $\Delta_{ku}$ introduces an asymptotic bias and the limiting distribution is not a zero mean Gaussian mixture.[26] For the trend parameters $\boldsymbol{\alpha}_p$ appearing in the cointegrating regression model (5), this framework does not allow their consistent estimation in the presence of deterministically trending integrated regressors (see, e.g., Hansen (1992a). As it follows from (7), and under standard cointegration, the composite trend parameters $\boldsymbol{\alpha}_p + \mathbf{A}'_{k,p}\boldsymbol{\beta}_k$ can be estimated consistently at the usual rate $n^{1/2}$, but the limiting distribution of the OLS estimator $\hat{\boldsymbol{\alpha}}_{p,n} + \mathbf{A}'_{k,p}\hat{\boldsymbol{\beta}}_{k,n}$ also depends on the nuisance parameters measuring the degree of endogeneity of the regressors.

Despite this last result, the OLS residuals are exactly invariant to the trend parameters, and allows for consistent estimation of the equilibrium error sequence under standard stationary cointegration.[27]

However, the limiting distribution of some commonly used residual-based statistics and functionals is plagued of these nuisance parameters, invalidating the inferential procedures based on standard normal asymptotic theory. On the other hand, under non-stationarity of the long-run relationship among $Y_t$ and $\mathbf{X}_{k,t}$ (no cointegration), the limiting results are quite different. Particularly, when the equilibrium error sequence $u_t = \eta_{0,t} - \boldsymbol{\beta}'_k\boldsymbol{\eta}_{k,t}$ contains a unit root, that is $u_t = u_{t-1} + \upsilon_t$, with $n^{-1/2}u_{[nr]} \Rightarrow B_\upsilon(r)$, then we get the following limiting result $n^{-3/2}\sum_{t=1}^{n}(n^{-1/2}\boldsymbol{\eta}_{k,t})u_t \Rightarrow \int_0^1 \mathbf{B}_k(s)B_u(s)ds$ when taking $v = -1/2$ in (7), determining the inconsistent estimation of the cointegrating vector $\boldsymbol{\beta}_k$, while that the OLS estimator of $\boldsymbol{\alpha}_p + \mathbf{A}'_{k,p}\boldsymbol{\beta}_k$ diverge at the rate $n^{1/2}$.

Once established these theoretical results, there remains to consider the fundamental question of consistently discriminate in practice between these two situations making use of some of the existing testing procedures for the null of no cointegration against cointegration (see, e.g., Phillips and Ouliaris (1990) and Stock (1999) for a review).

---

26 Given that the first term in (2.9) can be decomposed as $\int_0^1 \mathbf{B}_k(s)dB_u(s) = (1-\phi)^{-1}\int_0^1 \mathbf{B}_k(s)dB_{\upsilon.k}(s)$ $+(1-\phi)^{-1}\int_0^1 \mathbf{B}_k(s)d\mathbf{B}_k(s)'\boldsymbol{\Omega}_{kk}^{-1}\boldsymbol{\omega}_{k\upsilon}$, then under strict exogeneity of the regressors, $\boldsymbol{\omega}_{k\upsilon} = \mathbf{0}_k$, this stochastic integral behaves as a Gaussian mixture random process, where the remaining nuisance parameters can be removed by simple scaling.

27 From equation (2.1) and Assumption 2.1, we have that the observation *t* for the set of *k* deterministically trending integrated regressors can be decomposed as $\mathbf{X}_{k,t} = \mathbf{A}_{k,p}\boldsymbol{\Gamma}_{p,n}^{-1}\boldsymbol{\tau}_{p,tn} + \boldsymbol{\eta}_{k,t}$, which gives that the sequence of OLS residuals from (2.5) can be written as $\hat{u}_{t,p}(k) = u_t - n^{-v}\boldsymbol{\tau}'_{p,tn}(n^v\boldsymbol{\Gamma}_{p,n}^{-1}[(\hat{\boldsymbol{\alpha}}_{p,n} - \boldsymbol{\alpha}_p) + \mathbf{A}'_{k,p}(\hat{\boldsymbol{\beta}}_{k,n} - \boldsymbol{\beta}_k)]) - n^{-(1/2+v)}\boldsymbol{\eta}'_{k,t}[n^{1/2+v}(\hat{\boldsymbol{\beta}}_{k,n} - \boldsymbol{\beta}_k)]$ Making use of (2.7) or, alternatively given that (2.5) may be rewritten as $\hat{Y}_{t,p} = \boldsymbol{\beta}'_k\hat{\mathbf{X}}_{kt,p} + u_{t,p}$, with $\hat{Y}_{t,p}$, $\hat{\mathbf{X}}_{kt,p} = \boldsymbol{\eta}_{kt,p}$ and $u_{t,p}$ the OLS detrended error terms $u_t$, then we have that $\hat{u}_{t,p}(k) = u_{t,p} - n^{-(1/2+v)}\boldsymbol{\eta}'_{kt,p}[n^{1/2+v}(\hat{\boldsymbol{\beta}}_{k,n} - \boldsymbol{\beta}_k)]$.


ECORFAN® All rights reserved.

Alternatively we could test the opposite hypotheses, with cointegration as the null, by making use of the procedures proposed, among others, by Shin (1994), Choi and Ahn (1995), McCabe, Leybourne and Shin (MLS) (1997), Xiao (1999), Xiao and Phillips (2002) or Wu and Xiao (2008).

This is not the topic analyzed in this paper, but it must be stated that all these last testing procedures are based on asymptotically efficient estimates of the model parameters in the sense that this estimators asymptotically eliminate both the endogeneous bias and the non-centrality parameter appearing in (9). These estimation methods are based on several modifications to OLS and include the fully modified OLS (FM-OLS) approach of Phillips and Hansen (1990) and Kitamura and Phillips (1997), and the canonical cointegrating regression (CCR) method of Park (1992), which are based on two different nonparametric corrections. Also, it must be mentioned the dynamic OLS (DOLS) approach of Phillips and Loretan (1991), Saikkonen (1991) and Stock and Watson (1993) which is based on a parametric correction consisting on augmenting the specification of the cointegrating regression (5) with leads and lags of the first difference of the regressors.[28] A major drawback of any of these procedures is the choice of several tuning parameters, such as a kernel function and a bandwidth for long run variance estimation for FM-OLS or CCR estimation, and the number of leads and lags for the DOLS procedure.

All the above mentioned testing procedures for the null hypothesis of stationarity make use of the residuals obtained from one of these alternatives.[29]

Even though these estimators are considered asymptotically equivalent, there may be sensible differences in their use in finite samples.

Kurozumi and Hayakawa (2009) study the asymptotic behaviour of the asymptotically efficient estimators cited above under a *m* local-to-unity framework for describing moderately serially correlated equilibrium errors in a standard cointegrating regression equation, which is similar to the formulation in (2.12) with $\rho = \rho_m = 1 - c/m$, where $m \to \infty$, and $m/n \to 0$ as $n \to \infty$. This formulation imply that $\rho = \rho_m$ approaches 1 at a slower rate that does the *n* local-to-unity system, and it seems to be a more convenient tool of analysis when we relate the properties of the estimators for the cointegrating regression model with the local power properties of cointegration tests. We reserve the consideration of this case for further investigation.

After this discussion, the following assumption presents four alternative characterizations of the cointegrating, or equilibrium, error sequence representing different slight departures from the stationarity assumption underlying the standard stationary cointegration result.

---

28 Pesaran and Shin (1997) examines a further modification of the two-sided underlying distributed lag model in the DOLS approach, incorporating a number of lags of the dependent variable and eliminating the terms based on leads of the first differences of the regressors. That is, they propose to use a traditional autoregressive distributed lag (ARDL) model for the analysis of long-run relations and find several interesting results for the estimators of the long-run coefficients in terms of its consistency and asymptotic distribution.

29 Particularly, the Shin's (1994) and MLS (1997) test statistics are based on DOLS residuals, while that the testing procedure proposed by Choi and Ahn (1995) makes use of the feasible CCR residuals. The test statistics proposed by Xiao (1999), Xiao and Phillips (2002) and Wu and Xiao (2008) employ the FM-OLS residuals.

ISSN-Print: 2007-1582- ISSN-On line: 2007-3682
ECORFAN® All rights reserved.

Assumption 2.3. *Cointegrating error sequence*

*We assume that the error sequence in (2.5),* $u_t$*, is given by any of the following alternative characterizations:*

**(a)** *A moving average (MA) unit root under n local-to-unity asymptotics*

$$\Delta u_t = (1-\theta L)\upsilon_t,\ \theta = 1 - n^{-1}\lambda,\ \lambda \in [0,\bar{\lambda}] \qquad (10)$$

**(b)** *A local-to-finite variance process*

$$u_t = \upsilon_t + \frac{\lambda}{a n^{1/\alpha - 1/2}} b_t \upsilon_{\alpha,t} \qquad (11)$$

*With* $b_t \sim iid B(\pi)$ *a Bernoulli random sequence, mutually independent of* $\upsilon_t$ *and* $\upsilon_{\alpha,t}$*, where* $\upsilon_{\alpha,t}$ *is an iid sequence of symmetrically distributed infinite variance random variables, with distribution belonging to the normal domain of attraction of a stable law with characteristic exponent* $\alpha \in (0,2)$*, denoted as* $\upsilon_{\alpha,t} \in ND(\alpha)$.

**(c)** *An autoregressive (AR) unit root under n local-to-unity asymptotics with a highly persistent initial observation*

$$(1-\rho L)u_t = \upsilon_t,\ u_0 = \sum_{s=0}^{\infty} \rho^s \upsilon_{-s},\ \rho = \rho_n = 1 - c/n,\ c > 0 \qquad (12)$$

**(d)** *A stochastically integrated process*

$$u_t = \upsilon_t + \mathbf{v}'_{q,t}\mathbf{h}_{q,t} \qquad (13)$$

*With* $\mathbf{h}_{q,t} = \mathbf{h}_{q,t-1} + \xi_{q,t}$ *a q-dimensional integrated process, and* $\zeta_t = (\upsilon_t, \mathbf{v}'_{q,t}, \xi'_{q,t})'$ *a* 2*q*+1-*dimensional mean zero stationary sequence.*

The process considered in part (a) was first proposed by Jansson and Haldrup (2002) as a way to introduce a notion of near cointegration, and further exploited by Jansson (2005a, b) to derive point optimal tests of the null hypothesis of cointegration, when $\lambda = 0$, based on efficient tests for a unit MA root.

The mixture process in part (b) was proposed by Cappuccio and Lubian (2007) to assess the performance of some commonly used nonparametric univariate test statistics for testing the null hypothesis of stationarity of an observed process, so that in this paper we extended their results to determine the effects of an infinite variance error in a cointegration framework. Making use of the distributional results obtained by Paulauskas and Rachev (1998), Caner (1998) propose how to test for no cointegration under infinite variance errors.

These two first cases represent departures from the standard cointegration situation, preserving the same rates of consistency for the estimates as in the referenced case but determining some relevant changes in the asymptotic null distributions of the estimators. Case (c) is a slight modification of the well known local-to-unity approach to stationarity, where a stationary sequence is modelled as a first-order AR process with a root that approaches one with the sample size but that strictly less than one in finite samples.

For a finite sample size, the behavior is governed by the parameter *c*, determining the degree of persistence of the innovations to the process (Phillips, 1987).

ISSN-Print: 2007-1582- ISSN-On line: 2007-3682
ECORFAN® All rights reserved.

Elliott (1999) and Müller (2005) propose to extend the high persistence behavior of the strictly mean reverting error process in finite samples to the initial observation as well and to investigate its effects on the size and power properties of some tests for a unit root and for stationarity. Here this characterization is used to represent no cointegration when $c = 0$, or asymptotic no cointegration for a fixed $c > 0$ and $n \to \infty$, while a fixed value of $c > 0$ indicates stationary cointegration for a finite sample size. Finally, case (d) represents a generalized version of the heteroskedastic cointegrating regression model of Hansen (1992b) as has been proposed by McCabe et.al. (2006).[30] These authors consider the case where the unobserved stochastic trend components of the observed model variables in (1) can be decomposed as follows

$$\eta_t = \begin{pmatrix} \eta_{0,t} \\ \eta_{k,t} \end{pmatrix} = \Pi_m \mathbf{w}_{m,t} + \varepsilon_t + \mathbf{V}_t \mathbf{h}_{q,t} = \begin{pmatrix} \pi'_{0,m} \\ \Pi_{k,m} \end{pmatrix} \mathbf{w}_{m,t} + \begin{pmatrix} \varepsilon_{0,t} \\ \varepsilon_{k,t} \end{pmatrix} + \begin{pmatrix} \mathbf{v}'_{0q,t} \\ \mathbf{V}_{kq,t} \end{pmatrix} \mathbf{h}_{q,t}$$

Where $\mathbf{w}_{m,t} = \mathbf{w}_{m,t-1} + \upsilon_{m,t}$ is a $m\times 1$ vector integrated process, with initial value $\mathbf{w}_{m,0}, \mathbf{h}_{q,0} = O_p(n^{1/2-\delta})$ for any $0 < \delta \le 1/2$, $\Pi_m$ is a $(k+1)\times m$ real matrix with rank $k$, and $\upsilon_{m,t}$ $(m\times 1)$, $\varepsilon_t$ $(k+1)\times 1$, and $\mathbf{V}_t$ $(k+1)\times q$ are mean zero stationary processes which may be correlated. Given the linear combination of such a vector, $\mathbf{c}'_k \eta_t$, with $\mathbf{c}_k = (1, -\beta'_k)'$ as in equation (2), then the error term $u_t$ can be decomposed as follows

$$\begin{aligned} u_t &= \mathbf{c}'_k \eta_t = (\pi'_{0,m} - \beta'_k \Pi_{k,m}) \mathbf{w}_{m,t} + \varepsilon_{0,t} - \beta'_k \varepsilon_{k,t} + (\mathbf{v}'_{0q,t} - \beta'_k \mathbf{V}_{kq,t}) \mathbf{h}_{q,t} \\ &= \mathbf{c}'_k \Pi_m \mathbf{w}_{m,t} + \mathbf{c}'_k \varepsilon_t + \mathbf{c}'_k \mathbf{V}_t \mathbf{h}_{q,t} = \pi'_m \mathbf{w}_{m,t} + \upsilon_t + \mathbf{v}'_{q,t} \mathbf{h}_{q,t} \end{aligned} \quad (14)$$

With $\pi_m = \Pi'_m \mathbf{c}_k$, $\upsilon_t = \mathbf{c}'_k \varepsilon_t$, and $\mathbf{v}_{q,t} = \mathbf{V}'_t \mathbf{c}_k$. In this setup, the condition $\pi_m = \mathbf{0}_m$ is interpreted as stochastic cointegration, with $\beta_k$ the stochastically cointegrating vector. If in addition we set $E[\mathbf{v}'_{q,t}\mathbf{v}_{q,t}] = 0$, then we get what can be called as stationary cointegration, with $\mathbf{v}_{q,t} = \mathbf{0}_q$ corresponding to the case of standard stationary cointegration.[31] Otherwise, if $E[\mathbf{v}'_{q,t}\mathbf{v}_{q,t}] > 0$, then the equilibrium error term is said to be heteroskedastically integrated and the variables in (2.1) are said to be stochastically cointegrated. The definition of stochastic cointegration nests standard cointegration and heteroskedastic cointegration. Hansen (1992b) calls the last additive term in (2), $\mathbf{v}'_{q,t}\mathbf{h}_{q,t}$, a bi-integrated process, while that McCabe et.al. (2003) establish the long-run memoryless property of this type of processes through stating that the optimal *s* step ahead forecasts, in the sense of minimizing the mean square error, converge to the unconditional mean as the forecast horizon *s* increases. This means that the behavior of the process up to time *t* has negligible effect on its behavior into the infinite future. The presence of the stochastic trend component $\mathbf{h}_{q,t}$ induces long memory in the product process, but the effect of shocks on the level of the process is transitory rather than permanent, justifying the so-called stochastically trendless property of this type of processes. It is this property that gives meaning to the concept of common heteroskedastic persistence.

[30] See also Harris et.al. (2002), and McCabe et.al. (2003) for the treatment of some particular cases of this general model of stochastic cointegration.

[31] If this additional condition is extended to $\mathbf{V}_t = \mathbf{0}_{k+1,q}$, then the variables are all integrated and cointegrated in the Engle-Granger (EG) sense.

ISSN-Print: 2007-1582- ISSN-On line: 2007-3682
ECORFAN® All rights reserved.

Once stated this underlying structure of the unobserved trend components in $\eta_t$, there is an additional technical reason supporting the concept of stochastic cointegration.

This argument makes use of the concept of summability, originally introduced by Gonzalo and Pitarakis (2006). As can be seen from part(d) in Proposition 2.1, under stochastic cointegration, the partial sum process of the sequence of equilibrium errors is dominated by this last component that is summable of order 1/2, while that the stochastically integrated trend components $\eta_{0,t}$ and $\eta_{k,t}$ are summable of order 1. This formulation implies the generalization of the traditional concept of stationary cointegration allowing for equilibrium errors that are not purely stationary but display a lower degree of persistence that the underlying common stochastic trend as measured by a lower order of summability.

Finally, for a further justification of the theoretical and empirical relevance of this specification, we may refer to the work of Park (2002), Chung and Park (2007), and Kim and Lee (2011), where it is introduced the concept of nonlinear and nonstationary heteroskedasticity (NNH) describing a conditionally heteroskedastic process given by a nonlinear function of an integrated processes. This formulation represents a convenient generalization of the nonstationary regression by Hansen (1995) allowing for nonstationary regressors, and as an alternative to the class of highly persistent dynamic conditional heteroskedastic processes. Following Park's (2002) approach, the last term in (13) can be interpreted as the simplest particular version of the heterogeneity generating functions (HGF) that are asymptotically homogeneous (the identity function in our case).

The following lemma states the basis to obtain the main results of this paper concerning the limiting behavior of the OLS estimator in (7) and of the alternative estimator that will be presented and examined in the next section.

Lemma 2.1. *Given the error term of the static linear cointegrating regression equation,* $u_t$*, in (2.5), then:*

**(a)** *When generated according to* $\Delta u_t = (1-\theta L)\upsilon_t$, *with* $\theta = 1 - n^{-1}\lambda$, $\lambda \in [0,\bar{\lambda}]$, *as in Assumption 2.3(a) and under Assumption 2.2, then we have*

$$n^{-1/2}U_{[nr]} = n^{-1/2}\sum_{t=1}^{[nr]} u_t \Rightarrow U_\lambda(r) = B_\upsilon(r) + \lambda\int_0^r B_\upsilon(s)ds \qquad (15)$$

*with* $dU_\lambda(r) = dB_\upsilon(r) + \lambda B_\upsilon(r)$.

**(b)** *When generated according to the local-to-finite variance process in 2.3(b), then*

$$\left(a_n^{-1}\sum_{t=1}^{[nr]}\upsilon_{\alpha,t},\ a_n^{-2}\sum_{t=1}^{[nr]}\upsilon_{\alpha,t}^2\right) \Rightarrow \left(V_{1,\alpha}(r), V_{2,\alpha}(r)\right)$$

*with norming sequence* $a_n = an^{1/\alpha}$, *and where* $V_{1,\alpha}(r)$ *is the Lévy* $\alpha$*-stable process on the space* $D[0,1]$, *with* $V_{2,\alpha}(r)$ *its quadratic variation process,* $V_{2,\alpha}(r) = V_{1,\alpha}^2(r) - 2\int_0^r V_{1,\alpha}^-(s)dV_{1,\alpha}(s)$, *with* $V_{1,\alpha}^-(r)$ *the left limit of the process* $V_{1,\alpha}(r)$ *in r. Then, we have*

$$n^{-1/2}U_{[nr]} \Rightarrow U_{\alpha,\lambda}(r) = B_\upsilon(r) + \lambda V_{1,\alpha}(r) \qquad (16)$$

*And*

$$n^{-1/2}\sum_{t=1}^{n}\eta_{k,tn}u_t \Rightarrow \mathbf{G}_{k\upsilon} + \Delta_{k\upsilon} + \lambda\left\{V_{1,\alpha}(1)\mathbf{B}_k(1) - \int_0^1 \mathbf{B}_k(s)dV_{1,\alpha}(s)\right\} \qquad (17)$$

*For any* $0 < \pi \leq 1$, *with* $\mathbf{G}_{k\upsilon}$ *and* $\Delta_{k\upsilon}$ *as in (9).*


ECORFAN® All rights reserved.

**(c)** *When generated according to* $(1-\rho L)u_t = \upsilon_t$, *with* $\rho = \rho_n = 1 - c/n$, *c* ≥ 0, *as in Assumption 2.3(c) and under Assumption 2.2, then we have that*

$$n^{-1/2}(u_{[nr]} - u_0) \Rightarrow \omega_\upsilon(e^{cr} - 1)\xi + J_{\upsilon,c}(r) \qquad (18)$$

*Where* $\xi \sim N[0,(2c)^{-1}]$, *and* $J_{\upsilon,c}(r) = \int_0^r e^{(r-s)c} dB_\upsilon(s)$ $= B_\upsilon(r) + c\int_0^r e^{(r-s)c} B_\upsilon(s)ds$ *is an Ornstein-Uhlenbeck process, which is independent of* ξ*. Further, as c* > 0 *tends to zero, this is continuous in c and converges to* $J_{\upsilon,0}(r) = B_\upsilon(r)$.

**(d)** *When generated according to* $u_t = \upsilon_t + \mathbf{v}'_{q,t}\mathbf{h}_{q,t}$, *with* $\mathbf{h}_{q,t} = \mathbf{h}_{q,t-1} + \xi_{q,t}$ *a q-dimensional integrated process, and* $\zeta_t = (\upsilon_t, \mathbf{v}'_{q,t}, \xi'_{q,t})'$ *a* 2*q*+1*-dimensional mean zero stationary sequence satisfying the functional central limit theorem as in (8).*

*Then*

$$n^{-(1-v)}U_{[nr]} = n^{-(1/2-v)}\left\{ n^{-1/2}\sum_{t=1}^{[nr]} \upsilon_t \right\} + n^v \left\{ n^{-1}\sum_{t=1}^{[nr]} \mathbf{v}'_{q,t}\mathbf{h}_{q,t} \right\}$$

*Where for the last term we have that*

$$n^{-1}\sum_{t=1}^{[nr]} \mathbf{v}'_{q,t}\mathbf{h}_{q,t} = \frac{\mathbf{h}'_{q,0}}{\sqrt{n}} n^{-1/2}\sum_{t=1}^{[nr]} \mathbf{v}_{q,t} + n^{-1}\sum_{t=1}^{[nr]}\sum_{j=1}^{t} \xi'_{q,j}\mathbf{v}_{q,t} \Rightarrow \int_0^r \mathbf{B}_q(s)' d\mathbf{V}_q(s) + r\Delta_{q,0} \qquad (19)$$

*With* $\mathbf{B}_q(r)$ *and* $\mathbf{V}_q(r)$ *two q-dimensional Brownian processes given by the weak limits of* $n^{-1/2}\sum_{t=1}^{[nr]} \xi_{q,t}$ *and* $n^{-1/2}\sum_{t=1}^{[nr]} \mathbf{v}_{q,t}$, *respectively, and* $\Delta_{q,0} = \sum_{j=0}^{\infty} E[\xi'_{q,t-j}\mathbf{v}_{q,t}]$ $= \sum_{j=0}^{\infty} Tr(E[\mathbf{v}_{q,t}\xi'_{q,t-j}])$. *Thus,* $n^{-1/2}U_{[nr]} = O_p(n^{1/2})$ *and* $n^{-1}U_{[nr]} = n^{-1}\sum_{t=1}^{[nr]} \mathbf{v}'_{q,t}\mathbf{h}_{q,t} + O_p(n^{-1/2})$ *under stochastic cointegration.*

Proof. For the result in part (a), see Appendix A. For the results in part (b), see Lemmas 2.1 and C.1 in Cappuccio and Lubian (2007) for (16), and Appendix B for (17). These results make clear that the weighted sum of the two component processes in (2.11) allows to obtain these composite results. If, instead, we consider $u_t = \upsilon_t + \lambda b_t \upsilon_{\alpha,t}$, then the infinite variance process will dominate the behavior of the scaled partial sum process as can be seen from the following decomposition

$$n^{-1/2}U_{[nr]} = B_{n,\upsilon}(r) + \lambda a n^{1/\alpha - 1/2}(a n^{1/\alpha})^{-1}\sum_{t=1}^{[nr]} b_t \upsilon_{\alpha,t} = O_p(n^{1/\alpha - 1/2})$$

With no finite limiting results available in this case. For the result (18) in part (c), see Lemma 2 in Elliott (1999). With *c* > 0, the weak limit of the covariance-stationary series $u_t$ is $n^{-1/2}u_{[nr]} \Rightarrow M_{u,c}(r) = \omega_\upsilon \xi e^{cr} + J_{\upsilon,c}(r)$, which is a stationary continuous time process.

Finally, the result in part(d) follows from standard application of the convergence to stochastic integrals of a stochastically trendless process.

Remark 2.1. Given that $B_\upsilon(r)$ can be decomposed as $B_\upsilon(r) = B_{\upsilon.k}(r) + \gamma'_k\mathbf{B}_k(r)$, with $\gamma_k = \mathbf{\Omega}_{kk}^{-1}\boldsymbol{\omega}_{k\upsilon}$, then the limiting process $U_\lambda(r)$ in (2.15) can be decomposed as $U_\lambda(r) = B_{\upsilon.k,\lambda}(r) + \gamma'_k\mathbf{B}_{k,\lambda}(r)$, with $B_{\upsilon.k,\lambda}(r) = B_{\upsilon.k}(r) + \lambda\int_0^r B_{\upsilon.k}(s)ds$ and $\mathbf{B}_{k,\lambda}(r) = \mathbf{B}_k(r) + \lambda\int_0^r \mathbf{B}_k(s)ds$.

ISSN-Print: 2007-1582- ISSN-On line: 2007-3682
ECORFAN® All rights reserved.

Similarly, the limiting processes $Z_{\alpha,\lambda}(r)$ and $J_{\upsilon,c}(r)$ in (16) and (17) can also be written as $Z_{\alpha,\lambda}(r) = B_{\upsilon.k}(r) + \gamma_k' \mathbf{B}_k(r) + \lambda V_{1,\alpha}(r)$, and $J_{\upsilon,c}(r) = J_{\upsilon.k,c}(r) + \gamma_k' \mathbf{J}_{k,c}(r)$, with $J_{\upsilon.k,c}(r)$ an Ornstein-Uhlenbeck process defined on $B_{\upsilon,c}(r)$, that is $J_{\upsilon.k,c}(r) = B_{\upsilon.k}(r) + c\int_0^r e^{(r-s)c} B_{\upsilon.k}(s)ds$, and similarly for $\mathbf{J}_{k,c}(r)$ based on the *k*-dimensional Brownian process $\mathbf{B}_k(r)$.

The first two cases considered determine a modification of the standard formulation of stationary cointegration, but are susceptible to produce consistent estimation results.

The next result establish the consistency rate and weak limit distribution of the OLS estimator in (7) in the cases (10)-(12).

Proposition 2.1(a) *Under Assumption 2.2 and the generating mechanism given in (10) and (11) for the cointegrating error term, we have that the limiting distribution of the OLS estimator of the cointegrating regression equation in (5) is given by*

$$\begin{pmatrix} n^{1/2}\Gamma_{p,n}^{-1}[(\hat{\alpha}_{p,n} - \alpha_p) + \mathbf{A}'_{k,p}(\hat{\beta}_{k,n} - \beta_k)] \\ n(\hat{\beta}_{k,n} - \beta_k) \end{pmatrix} \Rightarrow \left(\int_0^1 \mathbf{m}(s)\mathbf{m}(s)'ds\right)^{-1} \left\{ \begin{pmatrix} \int_0^1 \tau_p(s)dB_\upsilon(s) \\ \mathbf{G}_{k\upsilon} \end{pmatrix} + \begin{pmatrix} \mathbf{0}_{p+1} \\ \Delta_{k\upsilon} \end{pmatrix} + \lambda \begin{pmatrix} \int_0^1 \tau_p(s)dT(s) \\ \mathbf{H}_k(1) \end{pmatrix} \right\} \quad (20)$$

*Where* $\mathbf{m}(r) = (\tau_p'(r), \mathbf{B}_k'(r))'$. $T(r)$ *and* $\mathbf{H}_k(1)$ *are given by* $T(r) = T_\upsilon(r) = \int_0^r B_\upsilon(s)ds$, *and* $\mathbf{H}_k(1) = \int_0^1 \mathbf{B}_k(s)B_\upsilon(s)ds$ *when* $u_t$ *is generated as in (10), while* $T(s) = V_{1,\alpha}(r)$ *and*

$$\mathbf{H}_k(1) = V_{1,\alpha}(1)\mathbf{B}_k(1) - \int_0^1 \mathbf{B}_k(s)dV_{1,\alpha}(s)$$

*When* $u_t$ *is generated as in (11).* **(b)** *Under Assumption 2.2, and the generating mechanism given in (12) for the cointegrating error term, then the limiting distribution for the OLS estimator of the cointegrating regression equation (5) is given by*

$$\begin{pmatrix} n^{-1/2}\Gamma_{p,n}^{-1}[(\hat{\alpha}_{p,n} - \alpha_p) + \mathbf{A}'_{k,p}(\hat{\beta}_{k,n} - \beta_k)] \\ \hat{\beta}_{k,n} - \beta_k \end{pmatrix} \Rightarrow \left(\int_0^1 \mathbf{m}(s)\mathbf{m}(s)'ds\right)^{-1} \int_0^1 \mathbf{m}(s)M_{u,c}(s)ds \quad (21)$$

*Where*

$$\int_0^1 \mathbf{m}(s)M_{u,c}(s)ds = \omega_\upsilon \xi \int_0^1 e^{cs}\mathbf{m}(s)ds + \int_0^1 \mathbf{m}(s)J_{\upsilon,c}(s)ds \quad (22)$$

Proof. The results follows directly from parts (a)-(c) of Lemma 2.1, and the continuous mapping theorem.

From (20), it is evident that the direct impact of the cases (a) and (b) in Assumption 2.3 on the limiting distribution of the OLS estimator is through the value of the parameter $\lambda$, indicating the degree of persistence of the error sequence $u_t$ in case (a), and the relative importance of the infinite variance component in case (b). The final effect will be different in each case due to the very different behavior and properties of the terms $T(s)$ and $\mathbf{H}_k$ integrating the last component in (2.20).

The question of assessing the impact of these choices on the FM-OLS, CCR and DOLS estimators is not considered here, and it is left as an extension of the above results in future research. On the other hand, the results from (2.21)-(2.22) indicate that the impact of a highly persistent initial observation introduce an additional perturbation into de asymptotic behavior of the OLS estimator, which is inconsistent for the cointegrating vector $\beta_k$.

ISSN-Print: 2007-1582- ISSN-On line: 2007-3682
ECORFAN® All rights reserved.

Afonso J. Asymptotic and finite-sample properties of a new simple estimator of cointegrating regressions under near cointegration. ECORFAN Journal-Mexico 2013, 4-11: 887-908

Without the consideration of this additional source of persistence, the case of stationary but highly persistent error terms in finite samples determinate limiting distributional results that are equivalent to what are obtained under no cointegration.

Remark 2.2. As has been established in Harris et.al. (2002) (part (ii) of Theorem 1), the result in (19) is only of application for the OLS estimator in (7) under stationary cointegration ( $E[\mathbf{v}'_{q,t}\mathbf{v}_{q,t}] = 0$ and $\mathbf{V}_t \neq \mathbf{0}_{k+1,q}$) and only if $\boldsymbol{\sigma}_{kq} = E[\text{vec}(\mathbf{V}_{kq,t})\upsilon_t] \neq \mathbf{0}_{kq}$. In this case we get $\sqrt{n}(\hat{\beta}_{k,n} - \beta_k) = O_p(1)$, and $\boldsymbol{\Gamma}^{-1}_{p,n}[(\hat{\alpha}_{p,n} - \alpha_p) + \mathbf{A}'_{k,p}(\hat{\beta}_{k,n} - \beta_k)] = O_p(1)$, so that $\hat{\alpha}_{p,n} - \alpha_p = O_p(n^{-1/2})$ in the case of stochastically integrated regressors ( $\mathbf{V}_{kq,t} \neq \mathbf{0}_{k,q}$ ) containing a deterministic trend component ( $\mathbf{A}_{k,p} \neq \mathbf{0}_{k,p+1}$). Thus, the relevant results for the limiting distribution of the OLS estimators in (7) are given by $n^{-1}\sum_{t=1}^{[nr]} \boldsymbol{\tau}_{p,tn}u_t = O_p(n^{-1/2})$, and [32]

$$(1/n)\sum_{t=1}^{n} \boldsymbol{\eta}_{k,tn}u_t = \left\{(1/n)\sum_{t=1}^{n}(n^{-1/2}\mathbf{h}_{q,t} \otimes \mathbf{I}_{k,k})\right\}\boldsymbol{\sigma}_{kq} + O_p(n^{-1/2}) \Rightarrow \left\{\int_0^1 (\mathbf{B}_q(s) \otimes \mathbf{I}_{k,k})ds\right\}\boldsymbol{\sigma}_{kq}$$

Under heteroskedastic cointegration with stochastically integrated regressors, that is when $E[\mathbf{v}'_{q,t}\mathbf{v}_{q,t}] > 0$, then it can be proved that $n^{-3/2}\sum_{t=1}^{n} \boldsymbol{\tau}_{p,tn}u_t = O_p(n^{-1/2})$, and

$$n^{-3/2}\sum_{t=1}^{n} \boldsymbol{\eta}_{k,tn}u_t = (1/n)\sum_{t=1}^{n}(n^{-1/2}\mathbf{h}'_{q,t} \otimes \mathbf{I}_{k,k})E[\text{vec}(\mathbf{V}_{kq,t})\mathbf{v}'_{q,t}](n^{-1/2}\mathbf{h}_{q,t}) + O_p(n^{-1/2})$$
$$\Rightarrow \int_0^1 (\mathbf{B}_q(s)' \otimes \mathbf{I}_{k,k})E[\text{vec}(\mathbf{V}_{kq,t})\mathbf{v}'_{q,t}]\mathbf{B}_q(s)ds$$

Which determine that $\hat{\alpha}_{p,n} - \alpha_p = O_p(\sqrt{n})$, and $\hat{\beta}_{k,n} - \beta_k = O_p(1)$. In order to obtain consistent estimation results in this case, Harris et.al. (2002) propose to utilize an instrumental variable (IV) technique by defining $\mathbf{m}_{t-s} = (\boldsymbol{\tau}'_{p,t-s}, \mathbf{X}'_{k,t-s})'$, $s \geq 0$, and using $\mathbf{m}_{t-m}$ for $s > 0$ as an instrument with

$$\begin{pmatrix}\hat{\alpha}_{p,n}(s)\\ \hat{\beta}_{k,n}(s)\end{pmatrix} = \left(\sum_{t=s+1}^{n} \mathbf{m}_{t-s}\mathbf{m}'_{t}\right)^{-1} \sum_{t=s+1}^{n} \mathbf{m}_{t-s}Y_t$$

The so-called AIV($s$) (asymptotic IV) estimator. With this estimator we have that the parameter $\boldsymbol{\sigma}_{kq}$ is replaced by $\boldsymbol{\sigma}_{kq,s} = E[\text{vec}(\mathbf{V}_{kq,t-s})\upsilon_t]$, where $\boldsymbol{\sigma}_{kq,s} \to \mathbf{0}_{kq}$ if we let $s \to \infty$. As a consequence, this estimator should be consistent by letting $s = s(n) \to \infty$, and $s/n \to 0$ as $n \to \infty$. These authors require that $s = O(n^{1/2})$. However, the limiting distribution of this estimator is contaminated by the presence of the parameters $\boldsymbol{\Lambda}_{q,i} = \sum_{j=i}^{\infty} E[\mathbf{v}_{q,t}\boldsymbol{\xi}'_{q,t-j}]$, for $i = 0$, 1, due to the endogeneity of the stochastically integrated regressors, so to obtain a useful result in practical applications it must be imposed the extra exogeneity condition $E[\mathbf{v}_{q,t}\boldsymbol{\xi}'_{q,t-j}] = E[\mathbf{V}'_t\mathbf{c}_k\boldsymbol{\xi}'_{q,t-j}] = \mathbf{0}_{q,q}$ for all $j = 0, \pm 1, \pm 2, \ldots$ These authors argue that any other existing standard procedure for asymptotically efficient estimation of the model parameters in this setup will work as usual. Particularly, given that the feasible FM-OLS and CCR estimators require the use of a consistent estimator of the long-run covariance matrix $\boldsymbol{\Omega}$ based on the sequence $\boldsymbol{\zeta}_t = (u_t, \boldsymbol{\zeta}'_{k,t})'$, with

[32] The details of the derivation of these results in our more general setup, not included in this paper, can be requested from the author.

ISSN-Print: 2007-1582- ISSN-On line: 2007-3682
ECORFAN® All rights reserved.

$\zeta_{k,t} = \Delta\eta_{k,t} = \Pi_{k,m}\upsilon_{m,t} + \Delta\varepsilon_{k,t} + (\mathbf{V}_{kq,t} - \mathbf{V}_{kq,t-1})\mathbf{h}_{q,t-1} + \mathbf{V}_{kq,t}\xi_{q,t}$, it may be expected seriously biased estimates given that, in general, $E[\zeta_t] \neq \mathbf{0}_{k+1}$, with

$$E[u_t] = E[\mathbf{v}'_{q,t}\mathbf{h}_{q,t}] = E[\mathbf{v}'_{q,t}\mathbf{h}_{q,0}] + \sum_{j=1}^{t} E[\mathbf{v}'_{q,t}\xi_{q,j}]$$

$$E[\zeta_{k,t}] = E[(\mathbf{V}_{kq,t} - \mathbf{V}_{kq,t-1})\mathbf{h}_{q,0}] + E[\mathbf{V}_{kq,t}\xi_{q,t-1}] - E[\mathbf{V}_{kq,t}\xi_{q,t}]$$

, where $E[u_t] = O(t)$, and $E[u_t] = 0$ only under the above exogeneity condition and also $E[\mathbf{v}'_{q,t}\mathbf{h}_{q,0}] = \mathbf{c}'_k E[\mathbf{V}_t\mathbf{h}_{q,0}] = 0$, that trivially holds if $\mathbf{h}_{q,0} = \mathbf{0}_q$. Thus, only a kernel-type estimator defined as the sample analog of $\tilde{\Omega}_n = (1/n)\sum_{t=1}^{n}\sum_{s=1}^{n}\tilde{\zeta}_t\tilde{\zeta}'_s$, with $\tilde{\zeta}_t = \zeta_t - E[\zeta_t]$, can produce the desired results. Next section is devoted to the analysis of an alternative estimation method to those considered here, which has been recently proposed by Vogelsang and Wagner (2011), that allows for a unified treatment of all the different data generating processes treated in this section and represents a very interesting and easy to use estimation procedure for cointegrating regression models.

**An alternative asymptotically almost efficient estimation method**

The new estimator of a cointegrating regression model proposed by Vogelsang and Wagner (2011) is based on a simple transformation of the model variables and allows to obtain an asymptotically unbiased estimator of the cointegrating vector $\beta_k$ in (5), with a zero mean Gaussian mixture limiting distribution under standard stationary cointegration. The first step requires to rewrite the cointegrating regression model in (5) as

$$S_t = \alpha'_p\mathbf{S}_{p,t} + \beta'_k\mathbf{S}_{k,t} + U_t \qquad (22)$$

Where $S_t = \sum_{j=1}^{t} Y_j$, $\mathbf{S}_{p,t} = \sum_{j=1}^{t}\tau_{p,j}$, $\mathbf{S}_{k,t} = \sum_{j=1}^{t}\mathbf{X}_{k,j}$, and $U_t = \sum_{j=1}^{t} u_j$ are obtain by applying partial summation on both sides of (5). This formulation can be called the integrated-cointegrating regression model, where the vector of transformed trending regressors in (22), $\mathbf{g}_t = (\mathbf{S}'_{p,t}, \mathbf{S}'_{k,t})'$, can be factorized as:

$$\mathbf{g}_t = \begin{pmatrix} n\Gamma^{-1}_{p,n} & \mathbf{0}_{p+1,k} \\ n\mathbf{A}_{k,p}\Gamma^{-1}_{p,n} & n\sqrt{n}\mathbf{I}_{k,k} \end{pmatrix}\begin{pmatrix} (1/n)\mathbf{S}_{p,tn} \\ \mathbf{H}_{k,tn} \end{pmatrix} = \mathbf{W}^0_n\mathbf{g}_{t,n} \qquad (23)$$

Where $\mathbf{S}_{p,t} = \Gamma^{-1}_{p,n}\sum_{j=1}^{t}\tau_{p,jn} = \Gamma^{-1}_{p,n}\mathbf{S}_{p,tn}$, $\mathbf{S}_{k,t} = \mathbf{A}_{k,p}\mathbf{S}_{p,t} + \mathbf{H}_{k,t}$, with $\mathbf{H}_{k,tn} = (1/n\sqrt{n})\mathbf{H}_{k,t}$, and $\mathbf{H}_{k,t} = \sum_{j=1}^{t}\eta_{k,t}$, as it comes from Assumption 2.1. The OLS estimators of $\alpha_p$ and $\beta_k$ from (22) are exactly invariant to the trend parameters in $\mathbf{X}_{k,t}$, and partial summing before estimating the model performs the same role that the nonparametric correction used by FM-OLS to remove $\Delta_{ku}$ in (9), but still leaves the problem caused by the endogeneity of the regressors. The solution pointed by these authors only requires that $\mathbf{X}_{k,t}$ be added as a regressor to the partial sum regression (22), that is

$$S_t = \alpha'_p\mathbf{S}_{p,t} + \beta'_k\mathbf{S}_{k,t} + \gamma'_k\mathbf{X}_{k,t} + e_t \qquad (24)$$

With $e_t = U_t - \gamma'_k\mathbf{X}_{k,t}$. Thus, (24) can be called the integrated modified (IM) cointegrating regression equation. When the integrated regressors do not contain any deterministic components (that is, $\mathbf{d}_{k,t} = \mathbf{0}_k$ in (1), with $\mathbf{A}_{k,p} = \mathbf{0}_{k,p+1}$ under Assumption 2.1), which is the case considered in Vogelsang and Wagner (2011), then the augmented vector of regressors in (24), $\mathbf{g}_t = (\mathbf{S}'_{p,t}, \mathbf{S}'_{k,t}, \mathbf{X}'_{k,t})'$, can be factorized as


ECORFAN® All rights reserved.

$$\mathbf{g}_t=\begin{pmatrix}\mathbf{S}_{p,t}\\ \mathbf{S}_{k,t}\\ \mathbf{X}_{k,t}\end{pmatrix}=\begin{pmatrix} n\boldsymbol{\Gamma}^{-1}_{p,n} & \mathbf{0}_{p+1,k} & \mathbf{0}_{p+1,k}\\ \mathbf{0}_{k,p+1} & n\sqrt{n}\mathbf{I}_{k,k} & \mathbf{0}_{k,k}\\ \mathbf{0}_{k,p+1} & \mathbf{0}_{k,k} & \sqrt{n}\mathbf{I}_{k,k}\end{pmatrix}\begin{pmatrix}(1/n)\mathbf{S}_{p,tn}\\ \mathbf{H}_{k,tn}\\ \boldsymbol{\eta}_{k,tn}\end{pmatrix}=\mathbf{W}^1_n\mathbf{g}_{t,n} \qquad (25)$$

Where $\mathbf{g}_{t,n}$ is stochastically bounded, with:

$$\mathbf{g}_{[nr],n}\Rightarrow \mathbf{g}(r)=\begin{pmatrix}\mathbf{g}_p(r)\\ \mathbf{g}_k(r)\\ \mathbf{B}_k(r)\end{pmatrix}=\begin{pmatrix}\int_0^r \boldsymbol{\tau}_p(s)ds\\ \int_0^r \mathbf{B}_k(s)ds\\ \mathbf{B}_k(r)\end{pmatrix} \qquad (26)$$

Where, as with (6), it is verified that $\int_0^1 \mathbf{g}(r)\mathbf{g}'(r)dr>0$. In the case of deterministically trending integrated regressors, that is with $\mathbf{A}_{k,p}\neq \mathbf{0}_{k,p+1}$, then the vector of regressors in (24), $\mathbf{g}_t=(\mathbf{S}'_{p,t},\mathbf{S}'_{k,t},\mathbf{X}'_{k,t})'$, is decomposed as

$$\mathbf{g}_t=\mathbf{W}^1_n\mathbf{g}_{t,n}+\begin{pmatrix}\mathbf{0}_{p+1}\\ \mathbf{A}_{k,p}[(1/\sqrt{n})\boldsymbol{\Gamma}^{-1}_{p,n}](1/n)\mathbf{S}_{p,tn}\\ \mathbf{A}_{k,p}[(1/\sqrt{n})\boldsymbol{\Gamma}^{-1}_{p,n}]\boldsymbol{\tau}_{p,tn}\end{pmatrix}$$

Where $(1/\sqrt{n})\boldsymbol{\Gamma}^{-1}_{p,n}$ is $O(n^{-1/2})$ in the case of stochastic regressors containing at most a constant term, that is $p = 0$, and $O(n^{1/2})$ for any $p \geq 1$. Thus, at the expense to develop an appropriate treatment in the general case, we proceed under the assumption that $\mathbf{A}_{k,p}=\mathbf{0}_{k,p+1}$ or, when $\mathbf{A}_{k,p}\neq\mathbf{0}_{k,p+1}$ that $\mathbf{g}_t=\mathbf{W}^1_n\mathbf{g}_{t,n}+O(n^{-1/2})$ for $p = 0$. This formulation allows to write the scaled and normalized bias vector from OLS estimation of (24), which is called the integrated modified OLS estimator (IM-OLS), as

$$\tilde{\boldsymbol{\theta}}_n=\begin{pmatrix}\tilde{\boldsymbol{\theta}}_{p,n}\\ \tilde{\boldsymbol{\theta}}_{\beta kn}\\ \tilde{\boldsymbol{\theta}}_{\gamma k,n}\end{pmatrix}=n^{-(1-v)}\mathbf{W}^{1\prime}_n\begin{pmatrix}\tilde{\boldsymbol{\alpha}}_{p,n}-\boldsymbol{\alpha}_p\\ \tilde{\boldsymbol{\beta}}_{k,n}-\boldsymbol{\beta}_k\\ \tilde{\boldsymbol{\gamma}}_{k,n}-\boldsymbol{\gamma}_k\end{pmatrix}=\begin{pmatrix}n^v\boldsymbol{\Gamma}^{-1}_{p,n}(\tilde{\boldsymbol{\alpha}}_{p,n}-\boldsymbol{\alpha}_p)\\ n^{1/2+v}(\tilde{\boldsymbol{\beta}}_{k,n}-\boldsymbol{\beta}_k)\\ n^{-1/2+v}(\tilde{\boldsymbol{\gamma}}_{k,n}-\boldsymbol{\gamma}_k)\end{pmatrix}$$

$$=\left((1/n)\sum_{t=1}^n \mathbf{g}_{t,n}\mathbf{g}'_{t,n}\right)^{-1}(1/n)\sum_{t=1}^n \mathbf{g}_{t,n}n^{-(1-v)}e_t \qquad (27)$$

Taking into account that the error term in the augmented integrated representation of the cointegrating regression equation (24) is given by:

$e_t = U_t-\boldsymbol{\gamma}'_k\boldsymbol{\eta}_{k,t}-\boldsymbol{\gamma}'_k\mathbf{A}_{k,p}\boldsymbol{\tau}_{p,t}=z_t-\boldsymbol{\gamma}'_k\mathbf{A}_{k,p}\boldsymbol{\tau}_{p,t}$

Then $n^{-(1-v)}e_t=n^{-(1-v)}z_t-n^{-(1-v)}\boldsymbol{\gamma}'_k\mathbf{A}_{k,p}\boldsymbol{\Gamma}^{-1}_{p,n}\boldsymbol{\tau}_{p,tn}$, with $n^{-(1-v)}z_t=n^{-(1-v)}U_t-n^{-(1/2-v)}\boldsymbol{\gamma}'_k\boldsymbol{\eta}_{k,tn}$, where under the cointegration assumption (with $v$ = 1/2) we get $n^{-1/2}z_{[nr]}\Rightarrow B_u(r)-\boldsymbol{\gamma}'_k\mathbf{B}_k(r)=B_{u.k}(r)$

Whenever $\boldsymbol{\gamma}_k=\boldsymbol{\gamma}_{ku}=\boldsymbol{\Omega}^{-1}_{kk}\boldsymbol{\omega}_{ku}$, where the second equality comes from the decomposition $B_u(r)=B_{u.k}(r)+\boldsymbol{\omega}'_{ku}\boldsymbol{\Omega}^{-1}_{kk}\mathbf{B}_k(r)$, with $B_{u.k}(r)=(1-\phi)^{-1}B_{\upsilon.k}(r)$, $\boldsymbol{\omega}_{ku}=(1-\phi)^{-1}\boldsymbol{\omega}_{k\upsilon}$, and $E[\mathbf{B}_k(r)B_{u.k}(r)]=\mathbf{0}_k$. This is also the weak limit of $n^{-1/2}e_{[nr]}$ whenever $\mathbf{A}_{k,p}=\mathbf{0}_{k,p+1}$ or when $p$ = 0, where $\boldsymbol{\Gamma}^{-1}_{0,n}=\boldsymbol{\tau}_{0,tn}=1$, while that when $\mathbf{A}_{k,p}\neq\mathbf{0}_{k,p+1}$ and $p \geq 1$ we have that $n^{-1/2}\boldsymbol{\gamma}'_k\mathbf{A}_{k,p}\boldsymbol{\Gamma}^{-1}_{p,n}\boldsymbol{\tau}_{p,tn}=O(n^{-1/2+p})$, and this term will dominate the behavior of $n^{-1/2}e_t$. On the other hand, under no cointegration (with $v$ = –1/2), we have $n^{-3/2}z_t=n^{-3/2}U_t+O_p(n^{-1})$, and this term will dominate the limiting behavior of $n^{-3/2}e_t$ unless $p \geq 2$ when $\mathbf{A}_{k,p}\neq\mathbf{0}_{k,p+1}$.

Under standard stationary cointegration, where $u_t=\phi u_{t-1}+\upsilon_t$, with $0\leq\phi<1$, $\upsilon_t$ as in Assumption 2.2 and $v$ = 1/2 in equation (27), the consistency rates of the estimators of the trend parameters $\boldsymbol{\alpha}_p$ and the cointegrating vector $\boldsymbol{\beta}_k$ are the usual ones for the OLS estimator in (7). More importantly, what is especially remarkable is that the asymptotic distribution of the IM-OLS estimator in (27) is zero mean mixed Gaussian, but with a different conditional asymptotic variance compared to that of the FM-OLS estimator.


ECORFAN® All rights reserved.

From Theorem 2 in Vogelsang and Wagner (2011), the limiting distribution under cointegration of the scaled and centered IM-OLS estimator of $(\alpha_p', \beta_k', \gamma_k')'$ is given by

$$\begin{pmatrix} n^{1/2}\Gamma_{p,n}^{-1}(\tilde{\alpha}_{p,n}-\alpha_p) \\ n(\tilde{\beta}_{k,n}-\beta_k) \\ \tilde{\gamma}_{k,n}-\gamma_k \end{pmatrix} \Rightarrow \tilde{\Theta} = \left(\int_0^1 \mathbf{g}(r)\mathbf{g}(r)'dr\right)^{-1}\int_0^1 \mathbf{g}(r)B_{u.k}(r)dr \qquad (28)$$

Where the limiting random vector $\tilde{\Theta}$ can also be written as

$$\tilde{\Theta} = \left(\int_0^1 \mathbf{g}(r)\mathbf{g}(r)'dr\right)^{-1}\int_0^1 [\mathbf{G}(1)-\mathbf{G}(r)]dB_{u.k}(r) \qquad (29)$$

With $\mathbf{G}(r)=\int_0^r \mathbf{g}(s)ds$ in (29). The correction for endogeneity based on the inclusion of the original regressors in the integrated-cointegrating regression works because it is of same stochastic order that $U_t$ under cointegration and all the correlation is soaked up into the vector parameter $\gamma_{ku}=\Omega_{kk}^{-1}\omega_{ku}$. On the other hand, under standard no cointegration when the cointegrating error term is a fixed unit root process, that is when $u_t = u_{t-1}+\upsilon_t$ with $\phi = 1$ and $v$ takes the value $v = -1/2$, then we get

$$\begin{pmatrix} n^{-1/2}\Gamma_{p,n}^{-1}(\tilde{\alpha}_{p,n}-\alpha_p) \\ \tilde{\beta}_{k,n}-\beta_k \\ n^{-1}(\tilde{\gamma}_{k,n}-\gamma_k) \end{pmatrix} \Rightarrow \tilde{\Theta} = \left(\int_0^1 \mathbf{g}(r)\mathbf{g}(r)'dr\right)^{-1}\int_0^1 \mathbf{g}(r)T_\upsilon(r)dr \qquad (30)$$

With $T_\upsilon(r)=\int_0^r B_\upsilon(s)ds$, that can be decomposed as $T_\upsilon(r)=\int_0^r B_{\upsilon.k}(s)ds+\int_0^r \mathbf{B}_k'(s)ds\cdot\gamma_{k\upsilon}$ $= T_{\upsilon.k}(r)+\mathbf{g}_k'(r)\gamma_{k\upsilon}$, with $\gamma_{k\upsilon}=\Omega_{kk}^{-1}\omega_{k\upsilon}$, so that the limiting random vector $\tilde{\Theta}_d$ can also be written as

$$\tilde{\Theta}_d = \begin{pmatrix} \mathbf{0}_{p+1} \\ \gamma_{k\upsilon} \\ \mathbf{0}_k \end{pmatrix} + \left(\int_0^1 \mathbf{g}(r)\mathbf{g}(r)'dr\right)^{-1}\int_0^1 \mathbf{g}(r)T_{\upsilon.k}(r)dr$$

With $\int_0^1 \mathbf{g}(r)T_{\upsilon.k}(r)dr=\int_0^1[\mathbf{G}(1)-\mathbf{G}(r)]B_{\upsilon.k}(r)dr$

This result indicates that, besides the change in the rates of convergence of the estimates and in the Gaussian process driving the mixed Gaussian distribution, there is an additional asymptotic bias term affecting the IM-OLS estimator of the cointegrating vector $\beta_k$ in the case of endogenous regressors ($\omega_{k\upsilon} \neq \mathbf{0}_k$).

Next result establish the limiting distribution and properties of the IM-OLS estimator in equation (27) under the Assumption 2.3 concerning the behavior of the cointegrating error sequence $u_t$.

Proposition 3.1. *Under Assumptions 2.2 and 2.3 for the cointegrating error term, then for the IM-OLS estimator of* $(\alpha_p', \beta_k', \gamma_k')'$ *computed from (24) we have that:*

**(a)** *For v = 1/2, and* $u_t$ *given in Assumption 2.3(a)-(b), then*

$$\tilde{\Theta}_n = \begin{pmatrix} n^{1/2}\Gamma_{p,n}^{-1}(\tilde{\alpha}_{p,n}-\alpha_p) \\ n(\tilde{\beta}_{k,n}-\beta_k) \\ \tilde{\gamma}_{k,n}-\gamma_k \end{pmatrix} \Rightarrow \tilde{\Theta} + \lambda\left(\int_0^1 \mathbf{g}(r)\mathbf{g}(r)'dr\right)^{-1}\int_0^1 \mathbf{g}(r)T_\upsilon(r)dr \qquad (31)$$

*With* $\tilde{\Theta}$ *as in (28)-(29), where* $T_\upsilon(r)=\int_0^r B_\upsilon(s)ds$ *in the case of the Assumption 2.3(a), and* $T_\upsilon(r)=V_{1,\alpha}(r)$ *in case of the Assumption 2.3(b). Also, in the cases of the Assumption 2.3(c)-(d) we have that* $\tilde{\Theta}_n = O_p(n)$, *and* $\tilde{\Theta}_n = O_p(\sqrt{n})$, *respectively.*

**(b)** *For v = –1/2, and* $u_t$ *generated as in Assumption 2.3(c), then*

$$\tilde{\Theta}_n = \begin{pmatrix} n^{-1/2}\Gamma_{p,n}^{-1}(\tilde{\alpha}_{p,n}-\alpha_p) \\ \tilde{\beta}_{k,n}-\beta_k \\ n^{-1}(\tilde{\gamma}_{k,n}-\gamma_k) \end{pmatrix} \Rightarrow \tilde{\Theta}_c = \left(\int_0^1 \mathbf{g}(r)\mathbf{g}(r)'dr\right)^{-1}\int_0^1 \mathbf{g}(r)T_{u,c}(r)dr \qquad (32)$$

*Where* $T_{u,c}(r)=\int_0^r M_{u,c}(s)ds=\omega_\upsilon\xi\int_0^r e^{cs}ds+\int_0^r J_{u,c}(s)ds$, *with* $\int_0^r e^{cs}ds = -(1/c)(1-e^{rc})$.

ISSN-Print: 2007-1582- ISSN-On line: 2007-3682
ECORFAN® All rights reserved.

**(c)** *For v = 0, with* $u_t$ *generated as in Assumption 2.3(d), and standard integrated regressors with* $\mathbf{V}_{kq,t} = \mathbf{0}_{k,q}$ *, then*

$$\tilde{\Theta}_n = \begin{pmatrix} \Gamma_{p,n}^{-1}(\tilde{\alpha}_{p,n} - \alpha_p) \\ n^{1/2}(\tilde{\beta}_{k,n} - \beta_k) \\ n^{-1/2}(\tilde{\gamma}_{k,n} - \gamma_k) \end{pmatrix} \Rightarrow \left(\int_0^1 \mathbf{g}(r)\mathbf{g}(r)'dr\right)^{-1} \int_0^1 \mathbf{g}(r)T_q(r)dr \tag{33}$$

*Where the limiting random process* $T_q(r)$ *given by* $T_q(r) = \int_0^r \mathbf{B}_q(s)'d\mathbf{V}_q(s) + r\Delta_{q,0}$.

Proof. These results simply follows from Lemma 2.1, the continuous mapping theorem, with $n^{-1/2}e_{[nr]} = n^{-1/2}U_{[nr]} - \gamma_k'\eta_{k,[nr]n} \Rightarrow B_{\upsilon.k}(r) + \lambda T_{\upsilon}(r)$, in the cases of the Assumption 2.3(a)-(b), and the same development as in the proof of Theorem 2 in Vogelsang and Wagner (2011).

Remark 3.1. From part (a) of Proposition 3.1, equation (31), in the case of the local-to-unity MA root in Assumption 2.3(a), we get $T_{\upsilon}(r) = T_{\upsilon.k}(r) + \gamma_{k\upsilon}'\mathbf{g}_k(r)$, where $T_{\upsilon.k}(r) = \int_0^r B_{\upsilon.k}(s)ds$, $\gamma_{k\upsilon} = \mathbf{\Omega}_{kk}^{-1}\mathbf{\omega}_{k\upsilon}$, and $\mathbf{g}_k(r)$ is given in equation (26). Then, it is immediate to rewrite equation (31) as

$$\tilde{\Theta}_n \Rightarrow \tilde{\Theta} + \lambda\begin{pmatrix} \mathbf{0}_{p+1} \\ \gamma_{k\upsilon} \\ \mathbf{0}_k \end{pmatrix} + \lambda\left(\int_0^1 \mathbf{g}(r)\mathbf{g}(r)'dr\right)^{-1}\int_0^1 \mathbf{g}(r)T_{\upsilon.k}(r)dr$$

Where the second term above determines an asymptotic bias component in the limiting distribution, while that the last multiplicative term can also be written as $\int_0^1 \mathbf{g}(r)T_{\upsilon.k}(r)dr = \int_0^1[\mathbf{G}(1) - \mathbf{G}(r)]B_{\upsilon.k}(r)dr$, as in equation (39). As can be seen from equations (38) and (40), for any λ > 0, this limiting distribution is a mixture of the corresponding ones under standard cointegration and no cointegration given above.

Also, denoting by $n^{-(1-v)}\tilde{e}_{t,p}(k) = n^{-(1-v)}e_t - \mathbf{g}_{t,n}'\tilde{\Theta}_n$ the sequence of scaled OLS residuals (IM-OLS residuals) from estimating the IM cointegrating regression in (34), then we get the following limiting distribution

$$\begin{aligned} n^{-1/2}\tilde{e}_{[nr],p}(k) &\Rightarrow B_{\upsilon.k}(r) + \lambda T_{\upsilon}(r) - \mathbf{g}'(r)\left\{\tilde{\Theta} + \lambda\begin{pmatrix} \mathbf{0}_{p+1} \\ \gamma_{k\upsilon} \\ \mathbf{0}_k \end{pmatrix} + \lambda\left(\int_0^1 \mathbf{g}(s)\mathbf{g}'(s)ds\right)^{-1}\int_0^1 \mathbf{g}(s)T_{\upsilon.k}(s)ds\right\} \\ &= B_{\upsilon.k}(r) - \mathbf{g}'(r)\tilde{\Theta} + \lambda[T_{\upsilon.k}(r) - \mathbf{g}'(r)\tilde{\Theta}_T] \end{aligned} \tag{34}$$

With $\tilde{\Theta}_T = (\int_0^1 \mathbf{g}(s)\mathbf{g}'(s)ds)^{-1}\int_0^1 \mathbf{g}(s)T_{\upsilon.k}(s)ds$, so that it is free of the effect of the additive limiting bias component while that the two additive components in the last line of (35) have the same structure and are not mutually independent. Additionally, from part (b) of the Proposition 3.1, we have that the last term in equation (32) can be decomposed as

$$\int_0^1 \mathbf{g}(r)T_{\upsilon,c}(r)dr = -\omega_{\upsilon}(1/c)\xi\int_0^1 \mathbf{g}(r)(1 - e^{rc})dr + \int_0^1[\mathbf{G}(1) - \mathbf{G}(r)]J_{\upsilon,c}(r)dr$$

So that the IM-OLS estimators has the usual divergence rates as under standard no cointegration, but with limiting distribution given by

$$\tilde{\Theta}_c \Rightarrow -\omega_{\upsilon}\frac{\xi}{c}\left(\int_0^1 \mathbf{g}(r)\mathbf{g}(r)'dr\right)^{-1}\int_0^1 \mathbf{g}(r)(1 - e^{rc})dr + \left(\int_0^1 \mathbf{g}(r)\mathbf{g}(r)'dr\right)^{-1}\int_0^1[\mathbf{G}(1) - \mathbf{G}(r)]J_{\upsilon,c}(r)dr$$

Where the first term can be interpreted as a stochastic bias-type component, while that the second one resembles the limiting distribution under standard no cointegration, with $B_{\upsilon}(r)$ replaced by $J_{\upsilon,c}(r)$.

Remark 3.2. The condition imposed on the integrated regressors in the framework of stochastic cointegration in part(c) is to simplify the calculations needed to obtain the limiting distribution and to preserve a similar structure that in the standard case. Thus, given that we can write

$$n^{-3/2}\mathbf{H}_{k,t} = \mathbf{\Pi}_{k,m}\left\{\frac{1}{n^{3/2}}\sum_{j=1}^{t}\mathbf{w}_{m,j}\right\} + \frac{1}{n}\left\{\frac{1}{\sqrt{n}}\sum_{j=1}^{t}\mathbf{\varepsilon}_{k,j}\right\} + \frac{1}{\sqrt{n}}\left\{\frac{1}{n}\sum_{j=1}^{t}(\mathbf{h}_{q,t}' \otimes \mathbf{I}_{k,k})\text{vec}(\mathbf{V}_{kq,t})\right\}$$


ECORFAN® All rights reserved.

And

$$n^{-1/2}\eta_{k,t} = \Pi_{k,m} n^{-1/2}\mathbf{w}_{m,t} + n^{-1/2}\boldsymbol{\varepsilon}_{k,t} + \mathbf{V}_{kq,t} n^{-1/2}\mathbf{h}_{q,t}$$

, then with $\mathbf{V}_{kq,t} = \mathbf{0}_{k,q}$ we have that

$$(1/n)\sum_{t=1}^{n}(n^{-3/2}\mathbf{H}_{k,t})(n^{-1/2}\eta'_{k,t}) = \Pi_{k,m}(1/n)\sum_{t=1}^{n}(n^{-3/2}\mathbf{W}_{m,t})(n^{-1/2}\mathbf{w}'_{m,t})\Pi'_{k,m} + o_p(1)$$
$$\Rightarrow \Pi_{k,m}\int_0^1\left(\int_0^r \mathbf{B}_m(s)ds\right)\mathbf{B}_m(r)'dr\,\Pi'_{k,m}$$

with $\mathbf{B}_k(s) = \Pi_{k,m}\mathbf{B}_m(s)$ in $\mathbf{g}(r)$.

These results makes clear that each of the alternatives considered will produce a different effect on the corresponding limiting distribution and, consequently, on the stochastic properties and behavior not only of the IM-OLS estimators but also on any other statistic based on it. However, from these limiting results it is not easy to deduce the impact on the precision of these estimates. Thus, in order to complete these findings we also present the results of a small simulation experiment designed to evaluate the finite sample estimation error of this estimator through the computation of the bias and RMSE for each of this alternatives describing the stochastic properties of the error term in a cointegrating regression equation.

**Finite sample results**

To evaluate this finite sample properties in parts (a), (b) in Proposition 3.1, we use the same model as in Vogelsang and Wagner (2011)for $k$ = 2, with $\eta_{k,t} = \eta_{k,t-1} + \boldsymbol{\varepsilon}_{k,t}$, where $\boldsymbol{\varepsilon}_{k,t} = \mathbf{C}_k(L)\boldsymbol{e}_{k,t}$, $\mathbf{C}_k(L) = \mathbf{I}_{k,k} + \mathbf{C}_{k1}L$, and $\mathbf{C}_{k1} = diag(c_{11}, c_{22})$, with $c_{11} = c_{22} = 0.5$, while that for the error term $u_t$ we use

$$u_t = \rho u_{t-1} + \upsilon_t + \gamma'_k \boldsymbol{e}_{k,t},$$

With $\gamma_k = (\gamma_1, \gamma_2)'$ controlling the degree of endogeneity of the regressors, and the iid sequence $(\upsilon_t, \boldsymbol{e}'_{k,t})'$ that follows a multivariate standard normal. Particularly, we set $\beta_k = (1,1)'$, and $\gamma = \gamma_1 = \gamma_2 = 0, 0.3$.

The results for this case are shown in Table 1 of Appendix C. On the other hand, to evaluate the performance of the IM-OLS estimator under heteroskedastic cointegration we use the same model as in Harris et.al. (2002), with

$$\begin{pmatrix} y_t \\ X_t \end{pmatrix} = \begin{pmatrix} \eta_{0,t} \\ \eta_{1,t} \end{pmatrix} = \begin{pmatrix} \pi_{01} & 0 \\ \pi_{11} & 0 \end{pmatrix}\mathbf{w}_{2,t} + \begin{pmatrix} \varepsilon_{0,t} \\ \varepsilon_{1,t} \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ v_{11,t} & 0 \end{pmatrix}\mathbf{h}_{2,t}$$

Where $\varepsilon_{i,t} = \phi\varepsilon_{i,t-1} + e_{i,t}, i = 0,1$, $v_{11,t} = \phi v_{11,t-1} + \sqrt{\lambda}e_{2,t}$, $\Delta h_{1,t} = e_{3,t}$, and $\Delta w_{1,t} = e_{4,t}$, with $u_t = -\beta v_{11,t}h_{1,t} + \varepsilon_{0,t} - \beta\varepsilon_{1,t}$, for the cointegrating error term, where $\beta = \pi_{01}/\pi_{11}$ under stochastic cointegration. Also, for the noise components we assume that $\boldsymbol{e}_t = (e_{0,t}, \ldots, e_{4,t})' \sim N_5(\mathbf{0}_5, \mathbf{R}_{5,5})$, where $\mathbf{R}_{5,5} = (\rho_{i,j})_{i,j=0,1,\ldots,4}$, with $\rho_{i,j} = E[e_{i,t}e_{j,t}]$. We set the values $\rho_{0,3} = \rho_{1,3} = 0.5$, $\rho_{0,1} = 0.25$, $\rho_{0,2} = \rho_{1,2} = \rho_{i,4} = 0$, $i$ = 0, 1, 2, 3, and $\rho_{2,3} = 0, 0.5$ where this last correlation coefficient measures the degree of endogeneity of regressors. For the AIV estimator we set $k_{i,n} = [c_i n^{1/2}], i = 1,2,3$, with $c_1 = 0.75$, $c_2 = 1.00$, $c_3 = 1.25$ for the lag order. In both cases we evaluate the performance of the integrated-OLS (I-OLS) and the IM-OLS estimators computed from (32) and (33), respectively.

From Table 1, we can see that the IM-OLS estimator always outperforms the standard OLS results in terms of finite sample bias, but with a higher RMSE, for increasing values of $\lambda$ in case 2.3(a). Very similar results are obtained in the case of the infinite-variance mixture process in 2.3(b), even under exogeneity of the regressor. In the last case of highly persistent but stationary equilibrium errors in finite samples,


ECORFAN® All rights reserved.

Table 1.C, both estimators are biased with a slightly lower bias for the IM-OLS estimator. When $u_0 = O_p(1)$, and particularly $u_0 = 0$, the results are absolutely comparable to these in terms of the finite sample bias, with a slight, but systematic, reduction of the RMSE due to the lower degree of persistence.

From Table 2, in the case of the finite sample performance of the AIV and IM-OLS estimators, the IM-OLS estimator performs as well as the AIV estimator in almost all the situations, except under endogeneity of the regressor and high correlation in $v_{11,t}$, where the AIV estimator, specially designed to taking into account for this effect, slightly outperforms the new estimator considered here.

**References**

Afonso Rodríguez, Julio A. (2013). Integrated Modified OLS estimation of cointegrating regressions with trending regressors and residual-based tests for the null of cointegration. Mimeography.

Caner, M. (1998). Tests for cointegration with infinite variance errors. Journal of Econometrics, 86(1), 155-175.

Cappuccio, N., D. Lubian (2007). Asymptotic null distributions of stationarity and nonstationarity tests under local-to-finite variance errors. Annals of the Institute of Statistical Mathematics, 59(3), 403-423.

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

Chung, H., J.Y. Park (2007). Nonstationary nonlinear heteroskedasticity in regression. Journal of Econometrics, 137(1), 230-259.

Elliott, G. (1999). Efficient tests for a unit root when the initial observation is drawn from its unconditional distribution. International Economic Review, 40(3), 767-783.

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

Gonzalo, J., J.Y. Pitarakis (2006). Threshold effects in cointegrating relationships. Oxford Bulletin of Economics and Statistics, 68, Supplement, 813-833.

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

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

Hansen, B.E. (1995). Regression with nonstationary volatility. Econometrica, 63(5), 1113-1132.

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

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

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


ECORFAN® All rights reserved.

Jansson, M., N. Haldrup (2002). Regression theory for nearly cointegrated time series. Econometric Theory, 18(6), 1309-1335.

Kim, C.S., S. Lee (2011). Spurious regressions driven by excessive volatility. Economics Letters, 113(3), 292-297.

Kitamura, Y., P.C.B. Phillips (1997). Fully modified IV, GIVE and GMM estimation with possibly nonstationary regressors and instruments. Journal of Econometrics, 80(1), 85-123.

Kurozumi, E., K. Hayakawa (2009). Asymptotic properties of the efficient estimators for cointegrating regression models with serially dependent errors. Journal of Econometrics, 149(2), 118-135.

McCabe, B., S. Leybourne, D. Harris (2003). Testing for stochastic cointegration and evidence for present value models. EconWPA Working Paper 0311009.

McCabe, B., S. Leybourne, D. Harris (2006). A residual-based test for stochastic cointegration. Econometric Theory, 22(3), 429-456.

McCabe, B.P.M., S.J. Leybourne, Y. Shin (1997). A parametric approach to testing the null of cointegration. Journal of Time Series Analysis, 18(4), 395-413.

Müller, U.K. (2005). Size and power of tests of stationarity in highly autocorrelated time series. Journal of Econometrics, 128(2), 195-213.

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

Park, J.Y. (2002). Nonstationary nonlinear heteroskedasticity. Journal of Econometrics, 110(2), 383-415.

Paulauskas, V., S.T. Rachev (1998). Cointegrated processes with infinite variance innovations. The Annals of Applied Probability, 8(3), 775-792.

Pesaran, M.H., Y. Shin (1997). An autoregressive distributed lag modelling approach to cointegration analysis. Cambridge Working Paper in Economics, no.9514. Faculty of Economics, University of Cambridge.

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

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

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

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

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

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

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

ISSN-Print: 2007-1582- ISSN-On line: 2007-3682
ECORFAN® All rights reserved.

Stock, J.H. (1999). A class of tests for integration and cointegration. In R.F. Engle, H. White (eds.): Cointegration, Causality and Forecasting. Oxford University Press, Oxford.

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

Vogelsang, T.J., M. Wagner (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, 307-320.

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

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

## Appendixes

**A. Proof of Lemma 2.1(a).** Using the representation $\Delta u_t = c_n(L)\upsilon_t$, then we can write $u_t = u_0 + c_n(L)\sum_{j=1}^{t}\upsilon_j$. Making use of the Beveridge-Nelson (BN) decomposition of the first-order lag polynomial $c_n(L)$ with $\theta = 1 - n^{-1}\lambda$, we have that $c_n(L) = 1 - \theta L = 1 - \theta - \theta(L-1) = n^{-1}\lambda - \theta(L-1)$, which gives $u_t = \lambda n^{-1/2}(n^{-1/2}\sum_{j=1}^{t}\upsilon_j) + \theta\upsilon_t + u_0 - \theta\upsilon_0$. Then, the scaled partial sum of $u_t$, $n^{-1/2}U_{[nr]} = n^{-1/2}\sum_{t=1}^{[nr]}u_t$, weakly converges to $U_\lambda(r)$ by direct application of Assumption 2.2.

**B. Proof of Lemma 2.1(b).** Making use of the

| | | λ | OLS | I-OLS | IM-OLS | OLS | I-OLS | IM-OLS |
|---|---|---|---|---|---|---|---|---|
| | | | Panel A. Bias | | | Panel B. RMSE | | |
| ρ = 0.0 | γ = 0.0 | 0.10 | 0.00003 | −0.00097 | −0.00088 | 0.0300 | 0.0416 | 0.0437 |
| | | 0.20 | −0.00060 | −0.00081 | −0.00057 | 0.0382 | 0.0554 | 0.0484 |
| | | 0.30 | 0.00028 | −0.00046 | −0.00020 | 0.0720 | 0.0604 | 0.0559 |
| | | 0.40 | 0.00123 | 0.00052 | −0.00010 | 0.0483 | 0.0750 | 0.0738 |
| | | 0.50 | 0.00078 | 0.00208 | 0.00199 | 0.0468 | 0.0742 | 0.0786 |
| | | 1.00 | 0.00061 | 0.00070 | −0.00048 | 0.1029 | 0.2368 | 0.2028 |
| | γ = 0.3 | 0.10 | 0.01164 | 0.00151 | −0.00027 | 0.0345 | 0.0485 | 0.0414 |
| | | 0.20 | 0.01181 | 0.00066 | 0.00019 | 0.0415 | 0.0605 | 0.0597 |
| | | 0.30 | 0.01185 | 0.00167 | 0.00108 | 0.0485 | 0.0781 | 0.0649 |
| | | 0.40 | 0.01279 | 0.00276 | 0.00133 | 0.0486 | 0.0787 | 0.0721 |
| | | 0.50 | 0.01209 | 0.00195 | 0.00114 | 0.0510 | 0.0748 | 0.0691 |
| | | 1.00 | 0.01163 | 0.00059 | 0.00031 | 0.1074 | 0.1918 | 0.1016 |
| ρ = 0.3 | γ = 0.0 | 0.10 | 0.00066 | 0.00111 | 0.00062 | 0.0371 | 0.0565 | 0.0547 |
| | | 0.20 | 0.00015 | 0.00015 | 0.00073 | 0.0444 | 0.1152 | 0.0800 |
| | | 0.30 | 0.00040 | 0.00010 | 0.00041 | 0.0422 | 0.0667 | 0.0654 |
| | | 0.40 | −0.00026 | −0.00029 | 0.00048 | 0.0543 | 0.0824 | 0.0769 |
| | | 0.50 | 0.00143 | −0.00048 | −0.00152 | 0.1325 | 0.1133 | 0.0898 |
| | | 1.00 | −0.00328 | −0.00008 | −0.00129 | 0.1567 | 0.1739 | 0.1569 |
| | γ = 0.3 | 0.10 | 0.02008 | 0.00221 | 0.00127 | 0.0493 | 0.0662 | 0.0553 |
| | | 0.20 | 0.02124 | 0.00290 | 0.00211 | 0.1047 | 0.0676 | 0.0708 |
| | | 0.30 | 0.01969 | 0.00227 | 0.00109 | 0.0623 | 0.0766 | 0.0700 |
| | | 0.40 | 0.02130 | 0.00345 | −0.00028 | 0.0809 | 0.1068 | 0.1127 |
| | | 0.50 | 0.01973 | 0.00035 | −0.00030 | 0.0669 | 0.1338 | 0.1147 |
| | | 1.00 | 0.02038 | 0.00225 | 0.00048 | 0.0979 | 0.1761 | 0.1699 |

decomposition of $u_t$ as in (11) we trivially have that

$$n^{-1/2}\sum_{t=1}^{n}(n^{-1/2}\eta_{k,t})u_t = n^{-1/2}\sum_{t=1}^{n}(n^{-1/2}\eta_{k,t})\upsilon_t + \lambda(an^{1/\alpha})^{-1}\sum_{t=1}^{n}(n^{-1/2}\eta_{k,t})b_t\upsilon_{\alpha,t}$$

Where for the first term we have the same result as in (9) using $u_t = \upsilon_t$, while that for the second term we have that it can be written as

$$(an^{1/\alpha})^{-1}\sum_{t=1}^{n}(n^{-1/2}\eta_{k,t})b_t\upsilon_{\alpha,t} = \frac{\eta_{k,0}}{\sqrt{n}}(an^{1/\alpha})^{-1}\sum_{t=1}^{n}b_t\upsilon_{\alpha,t} + (an^{1/\alpha})^{-1}\sum_{t=1}^{n}\left\{n^{-1/2}\sum_{j=1}^{t}\varepsilon_{k,j}\right\}b_t\upsilon_{\alpha,t}$$
$$= \left((an^{1/\alpha})^{-1}\sum_{t=1}^{n}b_t\upsilon_{\alpha,t}\right)\left(n^{-1/2}\sum_{j=1}^{n}\varepsilon_{k,j}\right)$$
$$-\ (an^{1/\alpha})^{-1}\sum_{t=1}^{n}b_t\upsilon_{\alpha,t}\left\{n^{-1/2}\sum_{j=1}^{t-1}\varepsilon_{k,j}\right\} + o_p(1)$$

As in Lemma 1 in Paulauskas and Rachev (1998). Then, the desired result follows by the joint convergence of each of these functionals to their corresponding weak limits.

ISSN-Print: 2007-1582- ISSN-On line: 2007-3682
ECORFAN® All rights reserved.

Article
**ECORFAN Journal-Mexico**
**OPTIMIZATION**
December 2013 Vol.4 No.11 887-908

## C. Simulation results

| | | | OLS | I-OLS | IM-OLS | OLS | I-OLS | IM-OLS |
|---|---|---|---|---|---|---|---|---|
| | | λ | Panel A. Bias | | | Panel B. RMSE | | |
| ρ = 0.0 | γ = 0.0 | 1 | −0.00011 | −0.00003 | 0.00013 | 0.0267 | 0.0398 | 0.0378 |
| | | 2 | 0.00051 | 0.00039 | 0.00015 | 0.0267 | 0.0414 | 0.0395 |
| | | 3 | −0.00091 | −0.00143 | −0.00065 | 0.0286 | 0.0461 | 0.0429 |
| | | 4 | 0.00034 | 0.00041 | 0.00115 | 0.0301 | 0.0498 | 0.0471 |
| | | 5 | 0.00011 | 0.00040 | 0.00018 | 0.0324 | 0.0537 | 0.0507 |
| | | 10 | −0.00010 | −0.00079 | −0.00013 | 0.0455 | 0.0861 | 0.0763 |
| | γ = 0.3 | 1 | 0.01477 | 0.00378 | 0.00274 | 0.0342 | 0.0435 | 0.0360 |
| | | 2 | 0.01778 | 0.00754 | 0.00618 | 0.0361 | 0.0472 | 0.0391 |
| | | 3 | 0.02092 | 0.01064 | 0.00925 | 0.0388 | 0.0518 | 0.0438 |
| | | 4 | 0.02340 | 0.01364 | 0.01236 | 0.0418 | 0.0555 | 0.0471 |
| | | 5 | 0.02652 | 0.01721 | 0.01454 | 0.0448 | 0.0607 | 0.0516 |
| | | 10 | 0.04198 | 0.03247 | 0.03146 | 0.0641 | 0.0952 | 0.0829 |
| ρ = 0.3 | γ = 0.0 | 1 | −0.00085 | −0.00021 | −0.00073 | 0.0364 | 0.0545 | 0.0531 |
| | | 2 | 0.00019 | −0.00015 | −0.00011 | 0.0374 | 0.0565 | 0.0550 |
| | | 3 | 0.00015 | 0.00076 | 0.00046 | 0.0400 | 0.0627 | 0.0597 |
| | | 4 | 0.00043 | −0.00011 | −0.00070 | 0.0417 | 0.0711 | 0.0668 |
| | | 5 | 0.00165 | 0.00206 | 0.00213 | 0.0454 | 0.0791 | 0.0711 |
| | | 10 | 0.00073 | 0.00136 | 0.00112 | 0.0661 | 0.1267 | 0.1128 |
| | γ = 0.3 | 1 | 0.02299 | 0.00570 | 0.00458 | 0.0490 | 0.0643 | 0.0527 |
| | | 2 | 0.02818 | 0.01123 | 0.01035 | 0.0523 | 0.0676 | 0.0561 |
| | | 3 | 0.03264 | 0.01611 | 0.01445 | 0.0571 | 0.0720 | 0.0620 |
| | | 4 | 0.03581 | 0.01957 | 0.01907 | 0.0591 | 0.0773 | 0.0690 |
| | | 5 | 0.04081 | 0.02569 | 0.02416 | 0.0647 | 0.0872 | 0.0754 |
| | | 10 | 0.06063 | 0.04727 | 0.04458 | 0.0914 | 0.1369 | 0.1187 |

**Table 1**

**Table 1.b**

| | | | OLS | I-OLS | IM-OLS | OLS | I-OLS | IM-OLS |
|---|---|---|---|---|---|---|---|---|
| | | $c$ | Panel A. Bias | | | Panel B. RMSE | | |
| ρ = 0.0 | γ = 0.0 | 1 | 0.00516 | 0.00569 | −0.00076 | 0.3380 | 0.6761 | 0.5863 |
| | | 2 | 0.00309 | 0.00344 | 0.00215 | 0.3024 | 0.6180 | 0.5506 |
| | | 3 | −0.00206 | −0.00303 | 0.00066 | 0.2728 | 0.5461 | 0.4807 |
| | | 4 | 0.00114 | 0.00237 | 0.00222 | 0.2475 | 0.4877 | 0.4356 |
| | | 5 | −0.00300 | −0.01230 | −0.00912 | 0.2248 | 0.4439 | 0.3969 |
| | | 10 | −0.00243 | 0.00284 | 0.00035 | 0.1600 | 0.3015 | 0.2751 |
| | γ = 0.3 | 1 | 0.27185 | 0.26264 | 0.26708 | 0.4384 | 0.7736 | 0.6746 |
| | | 2 | 0.23169 | 0.21143 | 0.21699 | 0.3936 | 0.6688 | 0.5959 |
| | | 3 | 0.21151 | 0.18603 | 0.19207 | 0.3628 | 0.6140 | 0.5324 |
| | | 4 | 0.18963 | 0.15766 | 0.16007 | 0.3292 | 0.5587 | 0.4803 |
| | | 5 | 0.17236 | 0.13321 | 0.13827 | 0.3090 | 0.5140 | 0.4504 |
| | | 10 | 0.12102 | 0.07778 | 0.08145 | 0.2277 | 0.3582 | 0.3063 |
| ρ = 0.3 | γ = 0.0 | 1 | 0.00359 | 0.00737 | 0.00973 | 0.4689 | 0.9324 | 0.8213 |
| | | 2 | 0.00263 | 0.01379 | 0.00568 | 0.4163 | 0.8412 | 0.7467 |
| | | 3 | 0.00322 | 0.00499 | 0.00176 | 0.3663 | 0.7306 | 0.6457 |
| | | 4 | 0.00552 | 0.02035 | 0.01937 | 0.3386 | 0.6801 | 0.6087 |
| | | 5 | 0.01248 | 0.01921 | 0.01723 | 0.3126 | 0.6258 | 0.5611 |
| | | 10 | 0.00474 | 0.00732 | 0.00815 | 0.2264 | 0.4268 | 0.3956 |
| | γ = 0.3 | 1 | 0.36510 | 0.35872 | 0.36815 | 0.5989 | 0.9968 | 0.8914 |
| | | 2 | 0.34023 | 0.32872 | 0.33163 | 0.5510 | 0.9195 | 0.8238 |
| | | 3 | 0.28632 | 0.25705 | 0.26476 | 0.5003 | 0.8318 | 0.7471 |
| | | 4 | 0.26499 | 0.21921 | 0.23548 | 0.4589 | 0.7440 | 0.6668 |
| | | 5 | 0.24230 | 0.18946 | 0.20133 | 0.4310 | 0.7236 | 0.6279 |
| | | 10 | 0.17118 | 0.10890 | 0.12016 | 0.3174 | 0.4974 | 0.4328 |

**Table 1.c**

ISSN-Print: 2007-1582- ISSN-On line: 2007-3682
ECORFAN® All rights reserved.

| | | | OLS | I-OLS | IM-OLS | OLS | I-OLS | IM-OLS |
|---|---|---|---|---|---|---|---|---|
| | | $\lambda$ | Panel A. Bias | | | Panel B. RMSE | | |
| $\rho = 0.0$ | $\gamma = 0.0$ | 0.10 | 0.00098 | 0.00094 | 0.00058 | 0.0747 | 0.0980 | 0.0905 |
| | | 0.20 | −0.00531 | −0.00733 | −0.00787 | 0.3362 | 0.4272 | 0.4150 |
| | | 0.30 | 0.01171 | 0.03165 | 0.03437 | 0.9818 | 1.4617 | 1.5600 |
| | | 0.40 | −0.00772 | −0.00705 | −0.00319 | 0.5296 | 0.7593 | 0.7488 |
| | | 0.50 | 0.00753 | 0.00571 | −0.00213 | 0.3902 | 0.7871 | 0.7316 |
| | | 1.00 | −0.04686 | 0.09637 | 0.03942 | 2.0277 | 11.0871 | 6.0260 |
| | $\gamma = 0.3$ | 0.10 | 0.00918 | −0.00323 | −0.00672 | 0.1062 | 0.1526 | 0.2891 |
| | | 0.20 | 0.00430 | −0.00018 | −0.00694 | 0.7481 | 0.9946 | 1.0094 |
| | | 0.30 | 0.02252 | 0.03624 | −0.01807 | 0.5379 | 2.4175 | 1.2893 |
| | | 0.40 | 0.03819 | 0.04805 | 0.04222 | 1.2524 | 2.8126 | 2.5005 |
| | | 0.50 | 0.03405 | 0.00279 | −0.00002 | 2.6439 | 0.7495 | 1.1833 |
| | | 1.00 | −0.01565 | 0.01585 | 0.02489 | 2.2374 | 1.6647 | 2.2443 |
| $\rho = 0.3$ | $\gamma = 0.0$ | 0.10 | −0.01296 | −0.01109 | −0.00875 | 0.6542 | 0.5336 | 0.5114 |
| | | 0.20 | 0.00255 | −0.00467 | −0.00111 | 0.1494 | 0.5570 | 0.3301 |
| | | 0.30 | 0.00157 | 0.00449 | 0.00829 | 0.2629 | 0.4152 | 0.4414 |
| | | 0.40 | 0.00762 | 0.00953 | 0.01343 | 0.5200 | 0.5780 | 0.6077 |
| | | 0.50 | 0.07859 | 0.08516 | 0.10830 | 6.1917 | 5.5190 | 8.0634 |
| | | 1.00 | 0.06707 | 0.19270 | 0.23362 | 4.2177 | 14.1266 | 16.7206 |
| | $\gamma = 0.3$ | 0.10 | 0.02081 | 0.01041 | 0.00263 | 0.2011 | 0.6788 | 0.4197 |
| | | 0.20 | 0.01469 | −0.00373 | −0.00577 | 0.2230 | 0.3045 | 0.3668 |
| | | 0.30 | 0.01852 | 0.00066 | 0.00184 | 0.3395 | 0.4656 | 0.4784 |
| | | 0.40 | 0.00966 | −0.01392 | −0.00435 | 0.5061 | 0.7540 | 0.3778 |
| | | 0.50 | 0.02908 | 0.02655 | 0.02374 | 0.5599 | 0.9811 | 1.0204 |
| | | 1.00 | 0.02143 | 0.00733 | −0.00350 | 0.8300 | 0.9785 | 1.0673 |

**Table 1.d**

| | | $n$: 100 | 200 | 400 | | $n$: 100 | 200 | 400 |
|---|---|---|---|---|---|---|---|---|
| | (a) Stationary standard cointegration, $\lambda = 0.00$ | | | | (b) Heteroskedastic cointegration, $\lambda = 0.05$ | | | |
| $\rho_{14} = 0.00$ | $AIV(k_{1,n})$ | 0.0129 | 0.0094 | −0.0001 | $\rho_{13} = 0.00$ | −0.0088 | 0.1514 | 0.0100 |
| | $AIV(k_{2,n})$ | 0.1374 | 0.0058 | 0.0002 | $\phi = 0.00$ | 0.0173 | −0.0939 | −0.0014 |
| | $AIV(k_{3,n})$ | −0.1928 | 0.0029 | 0.0001 | | 0.4682 | 0.1312 | −0.0063 |
| | I-OLS | −0.0029 | −0.0008 | −0.0004 | | −0.0115 | −0.0063 | −0.0032 |
| | IM-OLS | −0.0031 | −0.0001 | −0.0003 | | −0.0122 | −0.0068 | −0.0033 |
| $\rho_{14} = 0.50$ | $AIV(k_{1,n})$ | −0.0386 | −0.0206 | −0.0093 | $\rho_{13} = 0.50$ | 0.0782 | 0.4751 | 0.0098 |
| $\phi = 0.00$ | $AIV(k_{2,n})$ | 0.0399 | −0.0238 | −0.0081 | $\phi = 0.00$ | −0.0883 | 0.0918 | 0.0129 |
| | $AIV(k_{3,n})$ | −0.0187 | −0.0728 | −0.0164 | | −0.3106 | 0.1616 | −0.0081 |
| | I-OLS | −0.0035 | −0.0002 | 0.0002 | | −0.0113 | −0.0068 | −0.0019 |
| | IM-OLS | −0.0049 | −0.0006 | 0.0000 | | −0.0112 | −0.0072 | −0.0022 |
| $\rho_{14} = 0.50$ | $AIV(k_{1,n})$ | 0.1441 | −0.0320 | −0.0205 | $\rho_{13} = 0.50$ | −0.0961 | −0.0127 | 0.0042 |
| $\phi = 0.60$ | $AIV(k_{2,n})$ | 0.2339 | −0.0486 | −0.0268 | $\phi = 0.60$ | −0.4289 | −0.9996 | −0.0223 |
| | $AIV(k_{3,n})$ | −0.2081 | −0.0847 | −0.0308 | | −0.7932 | 0.0362 | −0.0235 |
| | I-OLS | −0.0076 | −0.0014 | −0.0007 | | −0.0487 | −0.0278 | −0.0135 |
| | IM-OLS | −0.0107 | −0.0020 | −0.0009 | | −0.0547 | −0.0309 | −0.0150 |

| | | $n$: 100 | 200 | 400 | | $n$: 100 | 200 | 400 |
|---|---|---|---|---|---|---|---|---|
| | (c) Heteroskedastic cointegration, $\lambda = 0.50$ | | | | (d) Heteroskedastic cointegration, $\lambda = 1.00$ | | | |
| $\rho_{13} = 0.00$ | $AIV(k_{1,n})$ | −0.2294 | −0.4070 | 0.1940 | $\rho_{13} = 0.00$ | −0.2319 | −0.3969 | 0.0202 |
| $\phi = 0.00$ | $AIV(k_{2,n})$ | −0.3278 | −0.0405 | −0.0764 | $\phi = 0.00$ | −0.6073 | −0.3017 | −0.1654 |
| | $AIV(k_{3,n})$ | −0.1334 | −1.0122 | −0.7849 | | 0.6853 | −0.0765 | −0.3884 |
| | I-OLS | −0.0714 | −0.0422 | −0.0205 | | −0.1294 | −0.0783 | −0.0454 |
| | IM-OLS | −0.0715 | −0.0426 | −0.0210 | | −0.1269 | −0.0781 | −0.0458 |
| $\rho_{13} = 0.50$ | $AIV(k_{1,n})$ | −0.4842 | −0.1329 | −0.3910 | $\rho_{13} = 0.50$ | −0.2094 | 0.3707 | −0.0661 |
| $\phi = 0.00$ | $AIV(k_{2,n})$ | −0.5111 | −0.7382 | −0.1531 | $\phi = 0.00$ | −0.1733 | −0.4813 | 0.4665 |
| | $AIV(k_{3,n})$ | −1.1233 | −0.5608 | 0.0731 | | 0.0803 | 0.3889 | 0.2871 |
| | I-OLS | −0.0780 | −0.0442 | −0.0231 | | −0.1332 | −0.0727 | −0.0442 |
| | IM-OLS | −0.0777 | −0.0445 | −0.0235 | | −0.1311 | −0.0726 | −0.0443 |
| $\rho_{13} = 0.50$ | $AIV(k_{1,n})$ | −0.1754 | −0.1213 | −0.2634 | $\rho_{13} = 0.50$ | −0.1066 | 0.1924 | −0.4829 |
| $\phi = 0.60$ | $AIV(k_{2,n})$ | −0.4312 | −0.1667 | −0.0937 | $\phi = 0.60$ | 1.3685 | 0.2596 | −0.2551 |
| | $AIV(k_{3,n})$ | 0.2278 | −0.6713 | 0.2967 | | −0.1473 | −0.1248 | −0.2577 |
| | I-OLS | −0.2408 | −0.1569 | −0.1038 | | −0.3300 | −0.2404 | −0.1712 |
| | IM-OLS | −0.2454 | −0.1601 | −0.1059 | | −0.3337 | −0.2435 | −0.1730 |

**Table 2**

ISSN-Print: 2007-1582- ISSN-On line: 2007-3682
ECORFAN® All rights reserved.