HAL open science

A Wald test for the cointegration rank in nonstationary fractional systems

Marco Avarucci, Carlos Velasco

To cite this version:

Marco Avarucci, Carlos Velasco. A Wald test for the cointegration rank in nonstationary fractional systems. Econometrics, 2009, 151 (2), pp.178. 10.1016/j.jeconom.2009.03.007 . hal-00578280

HAL Id: hal-00578280

https://hal.science/hal-00578280v1

Submitted on 19 Mar 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Accepted Manuscript

A Wald test for the cointegration rank in nonstationary fractional systems

Marco Avarucci, Carlos Velasco

PII: S0304-4076(09)00081-5
DOI: 10.1016/j.jeconom.2009.03.007
Reference: ECONOM 3191

To appear in: *Journal of Econometrics*

Received date: 10 November 2008
Accepted date: 13 March 2009

Please cite this article as: Avarucci, M., Velasco, C., A Wald test for the cointegration rank in nonstationary fractional systems. *Journal of Econometrics* (2009), doi:10.1016/j.jeconom.2009.03.007

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

A Wald test for the cointegration rank in nonstationary fractional systems☆

Marco Avarucci[a], Carlos Velasco[*,b]

[a] *Department of Quantitative Economics, Maastricht University*
[b] *Department of Economics, Universidad Carlos III de Madrid.*

Abstract

This paper develops new methods for determining the cointegration rank in a nonstationary fractionally integrated system, extending univariate optimal methods for testing the degree of integration. We propose a simple Wald test based on the singular value decomposition of the unrestricted estimate of the long run multiplier matrix. When the "strength" of the cointegrating relationship is less than 1/2, the test statistic has a standard asymptotic distribution, like Lagrange Multiplier tests exploiting local properties. We consider the behavior of our test under estimation of short run parameters and local alternatives. We compare our procedure with other cointegration tests based on different principles and find that the new method has better properties in a range of situations by using information on the alternative obtained through a preliminary estimate of the cointegration strength.

Key words: Fractional integration, fractional error correction model, singular value decomposition, cointegration test.
JEL: C12, C32

1. Introduction

Fractional cointegration models are increasingly used as a flexible tool for the modeling of long run relationships among economic time series. These models allow observed time series to be integrated of any arbitrary order, being even stationary as in many financial applications, and simultaneously permit any degree of persistence for the equilibrium relationship. Much effort has been dedicated in the last years to the estimation of the cointegrating relationship, including the asymptotic analysis of different variants of ordinary least squares

☆We are grateful to D. Marinucci, S. Johansen, J. Breitung, the co-editor M.A. Delgado and two referees for very helpful comments. Financial support of MIUR, Prin 2006, ref. 2006133899 and Spanish Plan Nacional de I+D+I, ref. SEJ2007-62908/ECON, is gratefully acknowledged.

*Corresponding author: Carlos Velasco, Departamento de Economía, Universidad Carlos III de Madrid C./ Madrid, 126 28903 Getafe (Madrid), Spain. Email: Carlos.Velasco@uc3m.es

ACCEPTED MANUSCRIPT

(OLS), such as narrow band and generalized least squares (GLS) versions, see e.g. Robinson and Marinucci [2003] and Robinson and Hualde [2003]. At the same time, a number of cointegration tests have been developed, most of them built on the null hypothesis of no cointegration versus the alternative of fractional cointegration. If the cointegration vector is known, standard methods for testing the integration degree could be routinely applied, but if this vector, or the level of integration of the original series, has to be estimated, inference methods should adapt to these additional sources of uncertainty. Thus, testing for the cointegration rank in this framework poses further complications in systems with more than two series, even if certain restrictions on the definition of cointegration are imposed.

In a semiparametric frequency domain set up, Marinucci and Robinson [2001] suggested a Hausman-type cointegration test comparing different estimates of the integration orders of the observed series. Recently Robinson [2008] provided rigorous theoretical support to this idea. Marmol and Velasco [2004] proposed a Wald test of the null of spurious relationships against the alternative of a single cointegration relationship among the components of a nonstationary vector process. Their approach relies upon comparing OLS and narrow band GLS-type estimates of the cointegrating vector, with different properties under the competitive hypotheses. A similar idea was used by Hualde and Velasco [2008], employing the GLS estimates of Robinson and Hualde [2003]. The chi-squared distribution of the GLS Wald statistic is inherited by a parallel cointegration test, hence avoiding the nonstandard asymptotic distribution of Marmol and Velasco's (2004) test and allowing for vector series with components of different integration orders.

In a parametric, time domain framework, Breitung and Hassler [2002] proposed a trace test for the cointegration rank based on a generalized eigenvalue problem of the type considered by Johansen [1988, 1991]. The resulting limit distribution of the statistic was found to be chi squared, where the degrees of freedom depend only on the cointegration rank under the null hypothesis. Nielsen [2005] argued that the equivalence of this regression based test and the Lagrange Multiplier (LM) test for integration does not extend to the multivariate case and showed that the actual multivariate LM test is also implicitly a test of the null of no cointegration. Breitung and Hassler [2006] considered the case were the cointegrating vector has to be estimated allowing for only one cointegration relationship. They showed that the limit distribution of the statistic is standard under the null of no cointegration, when employing the residuals from a regression in differences. Gil-Alaña [2004] extended Engle and Granger [1987]'s procedure, testing for the equality of memory parameters of the original series and of regression residuals using Robinson [1994]'s univariate test, while Nielsen [2004a] proposed a residual based LM test of the null hypothesis of cointegration assuming that the integration orders were known a priori.

Despite this effort on cointegration testing, relatively few work has been dedicated to the analysis of cointegration matrices, subspaces, and rank in fractional systems of dimension greater than two, allowing for multiple cointegrating relationships. For stationary series, Robinson and Yajima [2002] analyzed testing

ACCEPTED MANUSCRIPT

procedures based on the eigenvalues of the estimated and normalized spectral density matrix around frequency zero after a preliminary step to partition the vector series into subsets with identical differencing parameters. The restriction imposed by cointegration on the spectral density matrix at zero frequency was also investigated by Nielsen [2004b] and by Nielsen and Shimotsu [2007] using alternative semiparametric memory estimates. A different route was explored by Chen and Hurvich [2003], who proposed to estimate the cointegrating relationships by the eigenvectors corresponding to the smallest eigenvalues of an averaged periodogram matrix of tapered, differenced observations. Then, Chen and Hurvich [2006] developed and justified a test for fractional cointegration and a procedure for consistently determining the number and the dimension of the cointegrating subspaces.

A further line of work has focused on several fractional generalizations of Granger [1986]'s Error Correction Model (ECM), such as Davidson [2002], who applied parametric bootstrap to testing the existence of cointegrating relationships. Lasak [2007] considered Likelihood Ratio (LR) tests in a related framework, extending original Johansen's (1988, 1991) set up to allow explicitly for fractional cointegration alternatives with stationary residuals of unknown memory.

In this paper we focus on fractional cointegration methods inspired on a further test for the integration degree proposed by Lobato and Velasco [2007]. They questioned the choice of the regressor of the fractional Dickey-Fuller test of Dolado, Gonzalo, and Mayoral [2002] for the null hypothesis of unit root against the alternative of fractional unit root, and proposed an efficient version based on a different regression model. In Lobato and Velasco's (2007) basic framework, x_t is a *Type* II $I(d)$ fractional differencing process,

$$\Delta^d x_t 1_{\{t>0\}} = \eta_t, \qquad t = 1, 2, \ldots, \tag{1}$$

where η_t are independent and identically distributed (iid) random variables with zero mean and finite variance, and $1_{\{\cdot\}}$ is the indicator function. The fractional difference filter $\Delta^d = (1-L)^d$ is given by its formal expansion for any $d \neq -1, -2, \ldots,$

$$(1-z)^{-d} = \sum_{j=0}^{\infty} z^j \psi_j(d),$$

with $\psi_j(d) = \Gamma(j+d)/\left(\Gamma(d)\Gamma(j+1)\right)$, where Γ is the gamma function and $\Gamma(d) = \infty$ for $d = 0, -1, -2, \ldots,$ with $\Gamma(0)/\Gamma(0) = 1$. Suppressing the truncation in the notation, for any d, equation (1) can be rewritten as

$$\Delta x_t = \phi(\Delta^{d-1} - 1)\Delta x_t + \eta_t, \qquad t = 1, 2, \ldots, \tag{2}$$

where $\phi = 0$ under the null $H_0 : d = 1$, and $\phi = -1$ under the alternative $H_a : d < 1$. The null hypothesis is tested by means of a simple one-sided t-test for $\phi = 0$ in a rescaled regression, exploiting the fact that $(\Delta^{\delta-1} - 1)\Delta x_t$ is uncorrelated with Δx_t for any value of δ under H_0.

ACCEPTED MANUSCRIPT

We explore in this paper a multivariate extension of Lobato and Velasco's (2007) procedure for testing the cointegrating rank in a nonstationary fractionally integrated system. The univariate regression model (2) is replaced by an appropriate ECM to be estimated by OLS. The idea is to test whether the smallest singular values of the long run multiplier matrix estimates are significantly different from zero, exploiting the approach recently proposed by Kleibergen and Paap [2006]. We derive the limit distribution that is standard in the case of "weak cointegration". It is shown that the estimation of the memory of the residuals does not affect the asymptotic properties of the statistic, neither the estimation of other short run parameters.

The plan of the article is the following. In Section 2 we propose the Wald test to determine the cointegration rank in fractional systems, adapting Kleibergen and Paap's (2006) approach. In Section 3 we show the link between our test statistic and the canonical correlation test statistic and compare it with the trace test proposed by Breitung and Hassler [2002]. Section 4 proposes a generalized model accounting for more complex dynamics. The behavior of the test under local alternatives is analyzed in Section 5. In Section 6 the finite sample properties of the considered test are investigated by means of a small Monte Carlo experiment. Section 7 concludes and proposes some further lines of research. Proofs are collected in the Appendix.

Throughout this paper we shall adopt the following notation: $[A, B]$ indicates the $p \times (q+s)$ matrix obtained by placing side by side the $p \times q$ matrix A and the $p \times s$ matrix B. For $p > q$ we define $A_\perp$ to be a $p \times (p-q)$ matrix of rank $p-q$, for which $A'A_\perp = 0$. $|A|$, $\mathrm{rank}(A)$, $\mathrm{tr}(A)$ denote respectively the determinant, the rank and the trace of the (square) matrix A, $\|\cdot\|$ the Euclidean norm of a matrix such that $\|A\| = \sqrt{\mathrm{tr}(A'A)}$, $\mathrm{vec}(A)$ the vec operator stacking the columns of a matrix one over the other, I_p the p-rowed identity matrix. We write $A > 0$ to indicate that A is positive definite, "$\otimes$" indicates the Kronecker product and ":=" stands for definition.

2. Testing the cointegration rank

Our basic framework is given by the m-dimensional fractional error correction model (FECM)

$$\Delta^d X_t 1_{\{t>0\}} = \alpha\beta'(\Delta^{d-b} - \Delta^d) X_t 1_{\{t>0\}} + \varepsilon_t, \qquad t = 1, 2, \ldots, \tag{3}$$

where $\varepsilon_t \sim \mathrm{iid}(0, \Sigma_\varepsilon)$, $\Sigma_\varepsilon > 0$, and α and β are full rank $m \times r$ matrices, $0 \le r < m$, $0 < b < d$. Rewrite the FECM (3) as

$$\Delta^{d-b}\left(\Delta^b X_t 1_{\{t>0\}} - \alpha\beta'(1-\Delta^b) X_t 1_{\{t>0\}}\right) = \varepsilon_t \tag{4}$$

and define the characteristic polynomial $\Pi(z) = (1-z)^d - \alpha\beta'(1-(1-z)^b)(1-z)^{d-b}$ such that (4) is equivalent to $\Delta^{d-b}\Pi^*(L_b) X_t 1_{\{t>0\}} = \varepsilon_t$, with $L_b = 1 - \Delta^b$, and

$$\Pi^*(u) = (1-u) I_m - \alpha\beta' u, \quad u = u(z) := 1 - (1-z)^b, \tag{5}$$

ACCEPTED MANUSCRIPT

so that $\Pi(z) = (1-z)^{d-b}\Pi^*(1-(1-z)^b)$. Let the set $\mathbb{C}_b$ be the image of the unit disk under the mapping $z \mapsto 1-(1-z)^b$ (see Johansen [2008], Appendix A.3 and Figure 1 for further details). Next lemma characterizes the solutions of the FECM when $r > 0$.

Lemma 1. *Assume that $|\Pi^*(u)| = 0$ implies either $u = 1$ or $u \in \mathbb{C}_b$, that α, β have rank r, $0 < r < m$, and $|\alpha'_\perp \beta_\perp| \neq 0$. Then the solution of $\Pi(L)X_t 1_{\{t>0\}} = \varepsilon_t$ has the representation*

$$X_t = C_{\alpha\beta}\Delta^{-d}\varepsilon_t 1_{\{t>0\}} + \Delta^{b-d}Y_t 1_{\{t>0\}} \tag{6}$$

with $C_{\alpha\beta} = \beta_\perp(\alpha'_\perp\beta_\perp)^{-1}\alpha'_\perp$ and $Y_t = \alpha(\beta'\alpha)^{-1}\beta'\Delta^{d-b}X_t$, where $\Delta^{d-b}\beta'X_t$ is an $I(0)$ process having a positive definite spectral density matrix at the origin.

Then each element $x_{j,t}$ of the vector X_t, and each cointegrating residual $\beta'_i X_t$ are respectively *type*-II $I(d)$ and $I(d-b)$ fractionally differencing processes, for $j = 1, \ldots, m$, $i = 1, \ldots, r$, where β_i indicates the i-*th* column of the matrix β. Note that if $r = 0$, so that $\alpha\beta' = 0$, there is no cointegration and b is not identified by (3), but in this case we set its true value to 0 because any non trivial linear combination of the elements of X_t is $I(d)$. In principle it could be possible that each of the r cointegrating relationships has different degree of integration, allowing for different b. Our methods could be adapted to this possibility, but for simplicity we only consider a unique parameter b, and that all observed series are of the same memory d.

This set up includes triangular systems much used in the literature, see e.g. Robinson and Hualde [2003],

$$\begin{aligned} \beta'\Delta^{d-b}X_t &= U_{1t}1_{\{t>0\}} \\ \gamma'\Delta^d X_t &= U_{2t}1_{\{t>0\}}, \end{aligned} \qquad t = 1, 2, \ldots, \tag{7}$$

where $U_t = [U'_{1t}, U'_{2t}]' \sim \text{iid}(0, \Sigma_u)$, $\Sigma_u > 0$ and β and γ are $m \times r$ and $m \times (m-r)$ matrices respectively, $0 \leq r < m$, such that (β, γ) has full rank m. In this case $b > 0$ is required when $r > 0$ to identify β and γ. Using the identity $\gamma_\perp(\beta'\gamma_\perp)^{-1}\beta' + \beta_\perp(\gamma'\beta_\perp)^{-1}\gamma' = I_m$, the system (7) can then be rewritten as the FECM (3) when $r > 0$, where now ε_t is a linear and invertible transformation of the vector U_t, $\alpha = -\gamma_\perp(\beta'\gamma_\perp)^{-1}$, and $Y_t = -\alpha\beta'\varepsilon_t$.

From now on, for notational convenience, we will suppress the truncation for nonpositive t in (3), assuming implicitly that $X_t = 0$, $t \leq 0$. Moreover, we focus on the case $d = 1$, as assumed in most economic applications. To simplify the exposition, we first assume that b is known (or equal to a fixed value if testing the null hypothesis that $r = 0$). We later consider the case where we use a consistent estimate of b.

Given a sample X_t, $t = 1, 2, \ldots, n$, the basic idea of our procedure is testing the rank of the unrestricted OLS estimation of (3), i.e. a linear regression of ΔX_t on $(\Delta^{1-b} - \Delta)X_t = (\Delta^{-b} - 1)\Delta X_t$. We note that this regressor vanishes for $b = 0$. In order to make the regressor continuous at $b = 0$, following Lobato and Velasco [2007] we employ the rescaled regression model

$$\Delta X_t = \Pi Z^{(b)}_{t-1} + \varepsilon_t, \qquad t = 2, 3, \ldots, n \tag{8}$$

ACCEPTED MANUSCRIPT

where

$$Z_{t-1}^{(b)} := \frac{(\Delta^{-b}-1)\Delta X_t}{b}, \qquad \Pi := b\alpha\beta'. \tag{9}$$

For $b \to 0$, the indetermination $0/0$ in the first equation of (9) is solved using the L'Hôpital's rule, since the ratio $\lim_{b\to 0}((1-z)^{-b}-1)/b$ tends to the derivative of the fractional filter $(1-z)^{-\delta}$ evaluated at $\delta = 0$, that is, to the linear filter $J(z) = -\log(1-z) = \sum_{j=1}^{\infty} j^{-1}z^j$. We therefore define $Z_{t-1}^{(0)} = J(L)\Delta X_t$. The value of b reflects some knowledge on the alternative hypothesis that we wish to use when testing to gain power. If we set $b = 0$ this leads to local tests of LM type, see also Breitung and Hassler's (2002) test, but if under the alternative the true b is positive, local tests are not efficient. If $r = 0$ (so that $b = 0$), ΔX_t is an iid sequence and hence is uncorrelated with $Z_{t-1}^{(\delta)}$ for any value of δ. The hypothesis of no cointegration can be therefore easily nested in our framework and formulated as the null $\Pi = 0$ in the regression model $\Delta X_t = \Pi Z_{t-1}^{(\delta)} + \varepsilon_t$, any $\delta \geq 0$.

We stress that for $b \in [0, 0.5)$ the process $Z_{t-1}^{(b)}$ is asymptotically stationary and therefore standard normal asymptotics apply to the unrestricted OLS estimate of Π in (8),

$$\hat{\Pi} := S_{01}S_{11}^{-1}, \tag{10}$$

where we define the following sample moments based on n observations,

$$S_{01} := \frac{1}{n}\sum_{t=2}^{n} \Delta X_t Z_{t-1}^{(b)\prime}, \quad S_{11} := \frac{1}{n}\sum_{t=2}^{n} Z_{t-1}^{(b)} Z_{t-1}^{(b)\prime}.$$

Extending the results provided by Theorem 1 in Dolado, Gonzalo, and Mayoral [2002] to the multivariate framework, it can be shown that for $b = 0.5$ the distribution remain gaussian but the rate of convergence is equal to $(n\log(n))^{1/2}$. Allowing for $b > 0.5$, the distribution of $n^b \mathrm{vec}(\hat{\Pi} - \Pi)$ is expressed in terms of functionals of fractional Brownian motion. This distribution would lead to a non-pivotal statistic depending on the memory parameter b. The above arguments justify the following assumption allowing only for "weak cointegration", $b < 0.5$. This maintains the economic meaning of long-run equilibrium, implying that deviations from the equilibrium are "mean reverting" but nonstationary.

Assumption 1. We assume that data is generated by the FECM (3), d is known and equal to 1, $b \in [0, 0.5)$ and $\varepsilon_t \sim \mathrm{iid}(0, \Sigma_\varepsilon)$ with finite fourth order moment.

The rank test statistic relies on the following decomposition of a generic square matrix C of dimension m,

$$C = A_s B_s + A_{s,\perp}\Lambda_s B_{s,\perp}, \quad s = 0, 1, \ldots, m-1, \tag{11}$$

related to the singular value decomposition (hereafter SVD) $C = USV'$. A_s, B_s' are $m \times s$ matrices, $A_{s,\perp}, B_{s,\perp}'$ $m \times (m-s)$ matrices, satisfying the relations $A_s'A_{s,\perp} = 0$, $B_{s,\perp}B_s' = 0$, $A_{s,\perp}'A_{s,\perp} = I_{m-s}$, $B_{s,\perp}B_{s,\perp}' = I_{m-s}$. Let S_{m-s}

ACCEPTED MANUSCRIPT

the $m-s$ square submatrix of S, containing in its diagonal the $m-s$ smallest singular values of the matrix C. The $(m-s)\times(m-s)$ matrix Λ_s is defined as a transformation of S_{m-s}. If $\Lambda_s = 0$ and both A_s and B_s have full column rank, the rank of C is equal to s. The exact relation between the decomposition (11) and the SVD depends on the specification of A_s and B_s, obtained by imposing a normalization for A_sB_s. Every normalization leads to the same expressions of Λ_s, $A_{s,\perp}$ and $B_{s,\perp}$. For simplicity, we have followed here the notation of Kleibergen and Paap [2006] as close as possible, despite that it is not standard in the cointegrating literature. Let $\hat{C}$ be a consistent estimator of the unknown matrix C, and let $\hat{\Lambda}_s$ be an $(m-s)\times(m-s)$ matrix, such that there are matrices $\hat{A}_s$, $\hat{B}_s$, $\hat{A}_{s,\perp}$ and $\hat{B}_{s,\perp}$ for which $\hat{C}$ has a decomposition given by $\hat{C} = \hat{A}_s\hat{B}_s + \hat{A}_{s,\perp}\hat{\Lambda}_s\hat{B}_{s,\perp}$. Then, $\hat{\Lambda}_s$ reflects the distance to rank reduction, that is, a test for $\text{rank}(C) = r_0$ will be based on a test for $H_0 : \Lambda_{r_0} = 0$. Since $\text{vec}(\hat{\Lambda}_s)$ is just a rotation of $\text{vec}(\hat{S}_{m-s})$ around the origin, its elements are no longer restricted to be non-negative (as the singular values are), and can be asymptotically normally distributed. See Kleibergen and van Dijk [1998] and Kleibergen and Paap [2006] for details.

If we apply the SVD directly on $\hat{\Pi}$, the resulting testing procedure could be sensitive to scaling of $\hat{\Pi}$. Kleibergen and Paap [2006] suggest to normalize the estimator $\hat{\Pi}$ before we conduct the SVD of it, in order to improve the power properties of the test. Therefore, according to (11), we decompose the matrix

$$\hat{\Theta} := S_{00}^{-\frac{1}{2}}\hat{\Pi}S_{11}^{\frac{1}{2}},$$

with $S_{00} := n^{-1}\sum_{t=2}^{n}\Delta X_t\Delta X_t'$, as

$$\hat{\Theta} = \hat{A}_s\hat{B}_s + \hat{A}_{s,\perp}\hat{\Lambda}_s\hat{B}_{s,\perp}, \quad s = 0,\ldots,m-1. \tag{12}$$

Under the model (3) and Assumption 1, the probability limits of the matrices S_{00} and S_{11}, Ω_{00} and Ω_{11} respectively, have full rank according to next result.

Lemma 2. *Under Assumption 1, the following matrices have full rank,*

$$\Omega_{00} := \lim_{n\to\infty}\frac{1}{n}\sum_{t=2}^{n}\text{E}\left(\Delta X_t\Delta X_t'\right),\ \Omega_{11} := \lim_{n\to\infty}\frac{1}{n}\sum_{t=2}^{n}\text{E}\left(Z_{t-1}^{(b)}Z_{t-1}^{(b)'}\right).$$

Then the decomposition

$$\Theta := \Omega_{00}^{-\frac{1}{2}}\Pi\Omega_{11}^{\frac{1}{2}} = A_sB_s + A_{s,\perp}\Lambda_sB_{s,\perp} \tag{13}$$

is well defined and it follows that Π and Θ have the same rank. We note that up to the constant b, $A_s = \Omega_{00}^{-\frac{1}{2}}\alpha$ and $B_s = \beta'\Omega_{11}^{\frac{1}{2}}$, so that A_s and B_s can still be interpreted respectively as adjustment coefficients and cointegrating matrices α and β'.

We formulate the null hypothesis

$$H_0\left(r_0\right):\ r := \text{rank}(\Theta) = r_0. \tag{14}$$

ACCEPTED MANUSCRIPT

Testing (14) is equivalent to test for $H_0'(r_0) : \Lambda_{r_0} = 0$ if and only if A_{r_0} and B_{r_0} have full rank, as assumed when (3) was derived from (7). If the full rank condition on A_{r_0}, B_{r_0} is not satisfied, the hypothesis $\Lambda_{r_0} = 0$ should be interpreted as the null that the rank of the cointegrating space is at most r_0, against the alternative that it exceeds r_0; see the discussion in Johansen [1996], Chapter 5. The alternative hypothesis $H_a(r_0) : r > r_0$ can be therefore formulated as $\Lambda_{r_0} \neq 0$. The hypothesis of no-cointegration ($r = 0$) corresponds to the case $A_0 B_0 = 0$ and $A_{0,\perp}\Lambda_0 B_{0,\perp} = USV' = \Theta$. In this case we test if all the singular values of $\hat{\Theta}$ are statistically different from zero.

In order to identify $\hat{A}_s\hat{B}_s$ we impose the normalization $\hat{B}_s = [I_s, \hat{B}_{s,2}]$ as suggested by Kleibergen and Paap [2006]. The limiting behavior of the different elements of $\hat{\Theta}$ in (12) is stated in Lemma 3 and Theorem 1.

Lemma 3. *If $\hat{\Theta} \xrightarrow{p} \Theta$, then, under $H_0(r_0)$, $\hat{A}_{r_0}, \hat{B}_{r_0}, \hat{A}_{r_0,\perp}, \hat{B}_{r_0,\perp}$ converge in probability, respectively to $A_{r_0}, B_{r_0}, A_{r_0,\perp}, B_{r_0,\perp}$ and $\hat{\Lambda}_{r_0} \xrightarrow{p} 0$.*

Theorem 1. *Assume that b is known. Under $H_0(r_0)$ and Assumption 1,*

$$\sqrt{n}\hat{\lambda}_{r_0} \xrightarrow{d} \mathcal{N}\left(0, I_{(m-r_0)^2}\right),$$

where $\hat{\lambda}_s = vec(\hat{\Lambda}_s)$ and $\hat{\Lambda}_s = \hat{A}'_{s,\perp}\hat{\Theta}\hat{B}'_{s,\perp}$.

All proofs are in the Appendix. The proof of Lemma 3 is similar to the proof of Theorem 1 in Kleibergen and Paap [2006] and it is reported for completeness. The main difference with the proof of Theorem 1 in Kleibergen and Paap [2006] lies in the derivation of the intermediate results discussed in Remark 1 below. The limiting behaviour of the rank statistic $\mathrm{rk}_n(r_0) := n\hat{\lambda}'_{r_0}\hat{\lambda}_{r_0}$ to test (14) is described in the next corollary, whose proof is straightforward and hence omitted.

Corollary 1. *Under Assumption 1 and $H_0(r_0)$,*

$$rk_n(r_0) := n\hat{\lambda}'_{r_0}\hat{\lambda}_{r_0} \xrightarrow{d} \chi^2_{(m-r_0)^2}.$$

Under the alternative $H_a(r_0)$ the statistic $rk_n(r_0)$ diverges to infinity at rate n.

Remark 1. Kleibergen and Paap [2006] argue in the proof of their Theorem 1 that, provided that $\hat{\Lambda}_{r_0} \xrightarrow{p} \Lambda_{r_0} = 0$, the asymptotic normality of the (rescaled) OLS estimator $\hat{\Theta}$ of Θ,

$$\sqrt{n}\,\mathrm{vec}(\hat{\Theta} - A_{r_0}B_{r_0}) = \sqrt{n}\left[\mathrm{vec}(\hat{A}_{r_0}\hat{B}_{r_0} - A_{r_0}B_{r_0}) - \mathrm{vec}(\hat{A}_{r_0,\perp}\hat{\Lambda}_{r_0}\hat{B}_{r_0,\perp})\right]$$

implies that $\hat{A}_{r_0}\hat{B}_{r_0}$ is a root-n consistent estimator of $A_{r_0}B_{r_0}$. However this needs not to be the case because of possible cancelation between some non negligible components of the two $\mathrm{vec}(\cdot)$ terms inside the square brackets. In our proofs we use an alternative argument that only requires consistency as stated in Lemma 3.

ACCEPTED MANUSCRIPT

In the remainder of this section we consider the model (8) allowing b to be unknown, whereas d remains fixed and equal to one. Assumption 1 is still valid for the "true" b. To perform inference on the rank of Π in the equation (8) we need a consistent estimator $\hat{b}$ of b. We label $\tilde{\Pi}$ the least square estimator of Π obtained plugging in $\hat{b}$ in $Z_{t-1}^{(b)}$,

$$\tilde{\Pi} := S_{02} S_{22}^{-1},$$

where S_{02} and S_{22} are defined as S_{01} and S_{11}, respectively, with $Z_{t-1}^{(b)}$ replaced by $Z_{t-1}^{(\hat{b})}$. As done before, we can define the decomposition (11) on the scaled matrix

$$\tilde{\Theta} := S_{00}^{-\frac{1}{2}} \tilde{\Pi} S_{22}^{\frac{1}{2}} = \tilde{A}_s \tilde{B}_s + \tilde{A}_{s,\perp} \tilde{\Lambda}_s \tilde{B}_{s,\perp}. \tag{15}$$

If $\tilde{\Theta} \xrightarrow{p} \Theta$, then $\tilde{\Lambda}_s \xrightarrow{p} \Lambda_s$ by Lemma 3, where Λ_s is given by (13), so that $\Lambda_{r_0} = 0$ under $H_0(r_0)$. The following theorem shows that the first order asymptotic properties of the proposed test are not affected by the pre-estimation of the memory parameter b.

Theorem 2. *Let Assumption 1 hold and let the input $\hat{b}$ of $Z_{t-1}^{(\hat{b})}$ satisfy*

$$\hat{b} - b = O_p(n^{-\tau}), \quad \textit{with} \quad \tau > 0, \quad \textit{and} \quad \hat{b} \in [0, 0.5). \tag{16}$$

Then, under $H_0(r_0)$

$$\sqrt{n}\tilde{\lambda}_{r_0} \xrightarrow{d} \mathcal{N}\left(0, I_{(m-r_0)^2}\right),$$

where $\tilde{\lambda}_s = vec(\tilde{\Lambda}_s)$ and $\tilde{\Lambda}_s = \tilde{A}'_{s,\perp} \tilde{\Theta} \tilde{B}'_{s,\perp}$ and

$$\widetilde{rk}_n(r_0) := n\tilde{\lambda}'_{r_0} \tilde{\lambda}_{r_0} \xrightarrow{d} \chi^2_{(m-r_0)^2}. \tag{17}$$

Under the alternative $H_a(r_0)$ the statistic $\widetilde{rk}_n(r_0)$ diverges to infinity at rate n.

Remark 2. If the fractional difference parameter of the observed series d is also unknown, a test could be performed by replacing it by some consistent estimator $\hat{d}$. However, using similar arguments to Breitung and Hassler [2006], it can be shown that the estimation of d may affect the limiting distribution of the test, so Theorems 1 and 2 would no longer be valid.

Remark 3. We propose to estimate b from the residuals of the univariate regression in levels $X_{i,t} = \theta' X_t^{[i]} + e_{i,t}$, where $X_t^{[i]}$ is the $(m-1)$-vector resulting from the deletion of the i-th component from X_t. If $r > 1$ and the observables and cointegrating residuals are purely nonstationary, Marmol and Velasco [2005] showed that, in contrast to the standard case with $I(0)$ cointegration errors (see Wooldrige [1991], Johansen [2002]), the OLS estimate of the cointegrating vector $(1, \theta')'$ does not provide a consistent estimate of a suitable linear combination of the cointegrating relations, though remains bounded in probability. Despite of that, in our setting of common error memory of cointegrating residuals, it was

ACCEPTED MANUSCRIPT

shown that the OLS residuals $\hat{e}_t$ still approximate an $I(1-b)$ process as in the single equation set-up. See also Marmol and Velasco [2004] for a discussion.

If d is taken to be known and equal to 1, the condition $\hat{b} \in [0, 0.5)$ can be imposed naturally for implicitly defined memory estimators, such as the Gaussian semiparametric procedure of Robinson [1995] and the related exact local Whittle procedure of Shimotsu and Phillips [2005]. The upper bound implies that $Z_{t-1}^{(\hat{b})}$ is always asymptotically stationary. The lower bound is justified by the fact that order of integration of a linear combination of $I(1)$ systems cannot be greater than one, and in case of no cointegration it should be $I(1)$, being this the reason to fix the true b to 0 in this case. We do not discuss formally in this paper which estimation procedures satisfy the condition of power-rate convergence of the estimator $\hat{b}$ stated in (16). Bias reduction techniques like higher-order kernels suggested by Hualde and Robinson [2006], should be useful to augment the speed of convergence of $\hat{b}$. Further improvements could be obtained employing spectral regression methods for the estimation of θ.

3. A comparison with a related approach

Kleibergen and Paap [2006], Proposition 1, show that, if the covariance matrix of $\tilde{\lambda}_{r_0}$ has a Kronecker structure, $\tilde{\Pi}$ is the least square estimator and the normalization matrices are appropriately specified, the rank statistic $\widetilde{\mathrm{rk}}_n(r_0)$ can be computed as the sum of the smallest $m - r_0$ eigenvalues of $\tilde{\Theta}'\tilde{\Theta}$ multiplied by n. It follows that $\tilde{\lambda}'_{r_0}\tilde{\lambda}_{r_0} = \sum_{j=r_0+1}^{m} \mu_j$, where $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_m$ are the ordered eigenvalues solving the generalized eigenvalue problem, see also Johansen [1991], p. 94,

$$|\mu S_{00} - S_{02} S_{22}^{-1} S_{20}| = 0, \tag{18}$$

where $S_{20} = S'_{02}$.

We note that our approach is not developed in a likelihood inference framework and therefore it cannot be considered an extension to fractional set up of the analysis of Johansen [1988] of nonstationary systems. The rank statistic (17) results directly from the decomposition (12), showing that it can be viewed as a multivariate extension of the Wald statistic proposed by Lobato and Velasco [2007]. However, an alternative proof of Theorems 1-2 could be provided adapting the proof of Theorem 11.1 and 14.4 in Johansen [1996], where likelihood ratio (LR) statistics were considered.

The approach based on the solution of (18) is a useful computational device and allows us to compare our method with the trace statistic proposed by Breitung and Hassler [2002]. They test the null hypothesis that the cointegration rank is equal to r_0, checking if the $m - r_0$ smallest eigenvalues solving

$$\left|\mu S_{00} - M_{01} M_{11}^{-1} M_{10}\right| = 0 \tag{19}$$

are equal to zero, where

$$M_{10} = \frac{1}{n}\sum_{t=2}^{n} X_{t-1}^{*} \Delta X_t', \quad M_{11} = \frac{1}{n}\sum_{t=2}^{n} X_{t-1}^{*} X_{t-1}^{*'}$$

ACCEPTED MANUSCRIPT

with $X^*_{t-1} := \sum_{j=1}^{t-1} j^{-1}\Delta X_{t-j} = J(L)\Delta X_t$ and $M_{01} = M'_{10}$. Breitung and Hassler's trace statistic has the form $\mu(r_0) = n\sum_{j=r_0+1}^{m}\mu_j$ and it was proven that under the null it is asymptotically distributed as a $\chi^2_{(m-r_0)^2}$ random variable. Nielsen [2005] showed that $\mu(r_0)$ is not a regression variant of the multivariate score statistic, as demonstrated by Breitung and Hassler [2002] in the univariate case.

The eigenvectors corresponding to the eigenvalues μ_j, $j = 1, \ldots, m$, solving the generalized eigenvalue problem (19), can always be chosen to be orthogonal respect to S_{00}. Stated differently, given two generalized eigenvectors v_i, v_j, $i, j = 1, \ldots, m$, $v'_i S_{00} v_j = 1_{\{i=j\}}$. It turns out that

$$n\mu_j = n\frac{v'_j M_{01} M_{11}^{-1} M_{10} v_j}{v'_j S_{00} v_j}, \qquad j = 1, \ldots, m,$$

is a test-statistic for $\phi_j = 0$ in the auxiliary regression

$$(v'_j \Delta X_t) = \phi'_j J(L)\Delta X_t + e_t, \qquad t = 1, 2, \ldots, n. \tag{20}$$

The test based on $n\mu_j$ should not reject the null hypothesis that $\phi_j = 0$ in (20) if $v_j \in \text{span}(\alpha_\perp)$, and reject it when $v_j \in \text{span}(\beta)$. Therefore, $\mu(r_0)$ can be regarded as a Wald statistic rather than as a Score statistic.

However, rewriting the triangular system (7) as

$$\begin{aligned} \beta'\Delta X_t &= \beta'(\Delta - \Delta^{1-b})X_t + U_{1t} \\ \gamma'\Delta X_t &= U_{2t}, \end{aligned} \tag{21}$$

we note that the model (20) is misspecified because it does not include the data generating process (DGP) defined by (21) as a particular case under the alternative hypothesis, as pointed out by Lobato and Velasco [2007] in the univariate case. This misspecification can affect the efficiency of the resulting Wald test compared to a statistic based on the regression model

$$(v'_j \Delta X_t) = \varphi'_j Z^{(b)}_{t-1} + e_t, \qquad t = 1, 2, \ldots, n. \tag{22}$$

As noted in Section 2, $\lim_{b\to 0} Z^{(b)}_{t-1} = J(L)\Delta X_t$ so that models (20) and (22) are identical when b approaches to zero. Therefore also the statistics are equivalent if the normalization S_{00} is adopted. From the proof of Theorem 1, is easy to show that, under the null, $A_{r,\perp} S_{00.2}^{\frac{1}{2}} \Delta X_t$ has a smaller variance than $A_{r,\perp} S_{00}^{\frac{1}{2}} \Delta X_t$, where $S_{00.2} = S_{00} - S_{02} S_{22}^{-1} S_{20}$. It follows that the normalization S_{00} leads to better empirical size of the test, but employing the normalization $S_{00.2}$ (see Johansen [1996], p. 94), the asymptotic variance of the vector $\hat{\lambda}_{r_0}$ is the identity matrix both under the null and under the alternative and power may improve.

4. Short run dynamics

The FECM (3) was adequate to illustrate the idea behind the test procedure, but it is undoubtedly very restrictive for empirical applications. In order to allow

ACCEPTED MANUSCRIPT

for a richer dynamics we propose the process

$$\Phi(L)\Delta X_t = \alpha\beta'(\Delta^{-b} - 1)\Phi(L)\Delta X_t + \varepsilon_t, \tag{23}$$

where the roots of the polynomial matrix $\Phi(z) = I_m - \Phi_1 z - \cdots - \Phi_p z^p$ are all outside the unit circle. From Lemma 1 it follows that

$$X_t = \xi(1)C_{\alpha\beta}\Delta^{-1}\varepsilon_t + \Delta^{b-1}\xi(L)Y_t + \xi^*(L)C_{\alpha\beta}\varepsilon_t,$$

where $\xi(z) = \Phi(z)^{-1}$ and we exploit the decomposition $\xi(z) = \xi(1) + (1-z)\,\xi^*(z)$. The cointegrating matrix of X_t defined by (23) is given by $\beta^* = \xi(1)^{-1\prime}\beta = \Phi(1)'\beta$, so the cointegration rank is preserved and depends only on $\Pi = -b(\alpha\beta')$. Now the cointegrating residuals are given by the sum of an $I(1-b)$ process plus an $I(0)$ component. For the general case $d > 1$, $\beta'\Phi(1)X_t$ is a fractionally integrated process of order $\delta^* := \max\{d-1, d-b\}$.

Equation (23) motivates a nonlinear regression model but we propose to consider the rescaled linear regression

$$\Delta X_t = \Pi^* Z_{t-1}^{(b)} + \sum_{j=1}^{p} \Psi_j \Delta Z_{t-j}^{(b)} + \sum_{j=1}^{p} \Phi_j \Delta X_{t-j} + \varepsilon_t, \tag{24}$$

where $\Phi(z) = \Phi(1) + (1-z)\,\Phi^*(z)$, so that $\Psi_j = \Pi\Phi_j^*$ and $\Pi^* = \Pi\Phi(1) = b\alpha\beta^{*\prime}$. We build the rank test on the unrestricted OLS estimate of Π^* in model (24). Our approach ignores the multiplicative structure of the Ψ_j, therefore it can be inefficient but keeps the test procedure simple. Rewrite the model (24) as $\Delta X_t = \Pi^* Z_{t-1}^{(b)} + \Gamma W_{t-1}^{(b)} + \varepsilon_t$, with $\Gamma = [\Psi_1, \ldots, \Psi_p, \Phi_1, \ldots, \Phi_p]$ and $W_{t-1}^{(b)} = \left[\Delta Z_{t-1}^{(b)\prime}, \ldots, \Delta Z_{t-p}^{(b)\prime}, \Delta X_{t-1}', \ldots, \Delta X_{t-p}'\right]'$, and consider the partitioned linear regression model

$$\Delta X_t = \Pi^* Z_{t-1}^{(\hat{b})} + \Gamma W_{t-1}^{(\hat{b})} + e_t, \quad t = 2, \ldots, n. \tag{25}$$

Applying the *Frish-Waugh Theorem* it follows that the OLS estimate of Π^* is $\tilde{\Pi}^* = \tilde{S}_{02}\tilde{S}_{22}^{-1}$ where the matrices $\tilde{S}_{00}$, $\tilde{S}_{02}$ and $\tilde{S}_{22}$ are defined as S_{00}, S_{02} and S_{22} in Section 2, substituting $\Delta X_t, Z_{t-1}^{(\hat{b})}$ with $\Delta\tilde{X}_t, \tilde{Z}_{t-1}^{(\hat{b})}$, obtained after projection on $W_{t-1}^{(\hat{b})}$, e.g.

$$\Delta\tilde{X}_t = \Delta X_t - \sum_{t=2}^{n} \Delta X_t W_{t-1}^{(\hat{b})\prime} \left(\sum_{t=2}^{n} W_{t-1}^{(\hat{b})} W_{t-1}^{(\hat{b})\prime}\right)^{-1} W_{t-1}^{(\hat{b})}.$$

Then we generalize the definition of $\tilde{\Theta}$ to the estimation of the model with p lags, so that $\tilde{\Theta} = \tilde{S}_{00}^{-\frac{1}{2}}\tilde{\Pi}^*\tilde{S}_{22}^{\frac{1}{2}}$, $p = 0, 1, \ldots,$ as well as the parameter matrix $\Theta = \tilde{\Omega}_{00}^{-\frac{1}{2}}\Pi^*\tilde{\Omega}_{11}^{\frac{1}{2}}$ generalizes definition (13) for $p > 0$, where $\tilde{\Omega}_{00}$ and $\tilde{\Omega}_{11}$ are the probability limits of $\tilde{S}_{00}$ and $\tilde{S}_{22}$, respectively. Then $H_0(r_0)$ fixes a value for the rank of Θ and the FECM, for any $p = 0, 1, \ldots$.

We now present the following generalization of Theorem 2 based on the decomposition (15) of $\tilde{\Theta}$. Its proof is simpler than Theorem 4 bellow, and hence omitted.

ACCEPTED MANUSCRIPT

Theorem 3. *Let Assumption 1 hold and $\hat{b}$ satisfy (16). Assume that the DGP is given by (23). Then, under $H_0(r_0)$*

$$\sqrt{n}\tilde{\lambda}_{r_0} \xrightarrow{d} \mathcal{N}\left(0, I_{(m-r_0)^2}\right),$$

and the rank statistic $\widetilde{rk}_n(r_0) \xrightarrow{d} \chi^2_{(m-r_0)^2}$. Under the alternative $H_a(r_0)$ the statistic $\widetilde{rk}_n(r_0)$ diverges to infinity at rate n.

Given the structure of the FECM (23), we can consider an alternative two-step procedure based on a first step preliminary estimation of the filter $\Phi(L)$, obtained e.g. by means of the unrestricted least squares estimation of (24). The second step obtains the least squares estimate $\hat{\Pi}$ in the linear model

$$\Delta X_t = \Pi\hat{\Phi}(L) Z_{t-1}^{(\hat{b})} + \sum_{j=1}^{p} \Phi_j \Delta X_{t-j} + e_t, \tag{26}$$

where $\hat{\Phi}(L) Z_{t-1}^{(\hat{b})} = \left(I - \hat{\Phi}_1 L - \cdots - \hat{\Phi}_p L^p\right) Z_{t-1}^{(\hat{b})}$ is based on $n^{1/2}$-consistent estimates of the matrices Φ_j. Then we apply the SVD to the corresponding scaled least squares estimator, $\bar{\Theta} := \bar{S}_{00}^{-\frac{1}{2}}\hat{\Pi}\bar{S}_{22}^{\frac{1}{2}}$, so that $\bar{\Theta} = \bar{A}_{r_0}\bar{B}_{r_0} + \bar{A}_{r_0,\perp}\bar{\Lambda}_{r_0}\bar{B}_{r_0,\perp}$, where now $\bar{S}_{00}$ and $\bar{S}_{22}$ are moment matrices based on versions of ΔX_t and $\Delta Z_{t-1}^{(\hat{b})}$ only filtered by p lags of ΔX_t setting $W_{t-1}^{(b)} = \left[\Delta X'_{t-1}, \ldots, \Delta X'_{t-p}\right]'$. This method is parallel to the one proposed by Lobato and Velasco (2007) for univariate time series and the asymptotic theory for the corresponding test statistic based on $\bar{\lambda}_{r_0} = \mathrm{vec}(\bar{\Lambda}_{r_0})$ is the same as that in Theorem 3.

Corollary 2. *Under the conditions of Theorem 3 and $H_0(r_0)$,*

$$\sqrt{n}\bar{\lambda}_{r_0} \xrightarrow{d} \mathcal{N}\left(0, I_{(m-r_0)^2}\right),$$

and $\overline{rk}_n(r_0) := n\bar{\lambda}'_{r_0}\bar{\lambda}_{r_0} \xrightarrow{d} \chi^2_{(m-r_0)^2}$. Under the alternative $H_a(r_0)$ the rank statistic $\overline{rk}_n(r_0)$ diverges to infinity at rate n.

Since the estimation of Φ and b in the term $\hat{\Phi}(L) Z_{t-1}^{(\hat{b})}$ does not affect first order asymptotic properties of the test, the rank test statistic $\overline{\mathrm{rk}}_n(r_0)$ resulting of the least squares estimation of (26) is asymptotically equivalent to the LR statistic obtained from solving an eigenvalue problem similar to (18), where now in the sample moments S_{22}, S_{02} and S_{20}, $(\Delta X_t, Z_{t-1}^{(\hat{b})})$ is replaced by $(\Delta X_t, \Phi(L) Z_{t-1}^{(b)})$ previously projected on p lags of ΔX_t. This is true even in absence of knowledge on Φ and b, given in the latter case that $r_0 > 0$ so that b can be estimated consistently under both the null and the alternative that $r > r_0$. Therefore, for rank testing purposes is not necessary to consider the nonlinear model (23), and, following the same line of argument, we can expect the two step procedure to be more efficient than single step methods.

ACCEPTED MANUSCRIPT

5. Local alternatives

In this section we investigate the power properties of our tests against local alternatives. We consider first a general class of alternatives local to the null of $r = r_0 \geq 0$, for a fixed value of b, and later we also reinterpret these hypotheses under the null of no cointegration, $r = 0$, as a weakly cointegrated system with $b \to 0$ with n.

Consider the general model (24), and let h_i, $i = 1, \ldots, m$, be the singular values of the matrix $\Theta = \tilde{\Omega}_{00}^{-\frac{1}{2}} \Pi^* \tilde{\Omega}_{11}^{\frac{1}{2}}$, i.e. the limit of the coefficient $\tilde{\Theta}$ corresponding to the normalized regression (24). Similarly, we might work with procedures based on (26), see Corollary 3. We define the class of local alternatives for $r_0 = 0, 1, \ldots, m-2$ and $q_0 = 1, \ldots, m - r_0 - 1$ in terms of the ordered singular values $\{h_j\}_{j=1}^m$ of Θ,

$$H_{1n}(r_0, q_0): \ \mathrm{diag}\{h_{r_0+1}, \ldots, h_{r_0+q_0}\} = n^{-\frac{1}{2}} K_0,$$

where $K_0 = \mathrm{diag}\{k_1, \ldots, k_{q_0}\}$, $k_i > 0$ for $i = 1, \ldots, q_0$. The first r_0 singular values of Θ, $h_1, \ldots, h_{r_0}$, are positive constants and the remaining $m - (r_0 + q_0)$ ones are equal to zero. The interpretation of this parameterization is that under the local alternative $H_{1n}(r_0, q_0)$ the process has q_0 extra cointegrating vectors, $0 < q_0 < m - r_0$, that are difficult to detect, apart from the r_0 fixed cointegrating relationships included in $H_0(r_0)$. We label as $\Lambda_{r_0}^{(1,n)}$ the $(m - r_0) \times (m - r_0)$ matrix Λ_{r_0} for the r_0-SVD of Θ under the local alternatives $H_{1n}(r_0, q_0)$. In the next theorem we provide the asymptotic behavior of Wald tests under H_{1n}.

Theorem 4. *Let Assumption 1 hold and $\hat{b}$ satisfy (16). Under $H_{1n}(r_0, q_0)$,*

$$\sqrt{n}\tilde{\lambda}_{r_0} \xrightarrow{d} N(\xi_{r_0}, I_{m-r_0} \otimes I_{m-r_0}),$$

where $\xi_{r_0} = \lim_{n\to\infty} \sqrt{n}\lambda_{r_0}^{(1,n)}$ and $\lambda_{r_0}^{(1,n)} = vec(\Lambda_{r_0}^{(1,n)})$. The rank statistic $\widetilde{rk}_n(r_0)$ converges to a noncentral $\chi^2_{(m-r_0)^2}$ with noncentrality parameter $\sum_{j=1}^{q_0} k_j^2$.

Remark 4. Theorem 4 shows that our Wald tests have nontrivial power under local alternatives converging to the null at a parametric rate. This result is in line with the standard asymptotics of the tests under the null. The drift depends naturally on the strength of the extra local cointegrating relationships. However this drift does not depend explicitly on the presence of extra short term lags in the model, since the local hypothesis is established on the singular values of Θ, which is already normalized by sample moments of sequences projected out on these lags. If the local hypothesis is established in terms of the unnormalized coefficient Π^*, then the corresponding drift would depend on the probability limits of $\tilde{S}_{00}$ and $\tilde{S}_{22}$, and the local power of the test will be affected by the usual signal-to-noise ratio tradeoff. Similarly, the larger q_0, i.e. the more extra cointegrating relationships in $H_{1,n}(r_0, q_0)$, the larger is the local power.

Remark 5. The effect of the parameter b on the local power is however twofold. First, we note that $\|\Pi^*\| = O(b)$, for given cointegrating and adjustment matrices. On the other hand, the variance of $Z_{t-1}^{(b)}$ depends also on b, both because

ACCEPTED MANUSCRIPT

normalization and fractional integration, so the overall impact factor of b on the noncentrality parameter ξ_{r_0} is given by $\sum_{j=1}^{\infty}\psi_j(b)^2$, which is increasing with b, and zero for $b=0$. It is also important to note that in this case the pre-estimation of b does not play a role in asymptotics, but that when $r_0>0$ in $H_{1n}(r_0,q_0)$, b has to be strictly positive to give a full meaning to the (fixed) r_0 cointegration relationships.

Changing the focus of the alternative hypotheses when $r_0=0$, we can set up explicitly local alternatives in the cointegrating degree b, assuming that it converges to zero with n as $b=n^{-1/2}b_0$, $b_0>0$, whereas $\Pi/b=\alpha\beta'$ and $\Pi^*/b=\alpha\beta^{*\prime}$ are fixed with n, where α,β and β^* are full rank $m\times q_0$ matrices. Under the local alternatives $H_{1n}(0,q_0)$, $S=\mathrm{diag}\{n^{-\frac{1}{2}}K_0,\mathbf{0}_{m-q_0}\}$, where $\mathbf{0}_{m-q_0}$ is a $m-q_0$ vector of zeros. In this set up, instead of considering the singular values of Θ to establish the local alternatives, we consider the SVD of the unnormalized coefficient Θ/b, $\tilde{\Omega}_{00}^{-\frac{1}{2}}\alpha\beta^{*\prime}\tilde{\Omega}_{11}^{\frac{1}{2}}=US_0V'$. Then S and S_0 are related by the following relation

$$S=n^{-1/2}b_0S_0,\qquad \text{with}\qquad S_0=\mathrm{diag}\left\{\frac{K_0}{b_0},\mathbf{0}_{m-q_0}\right\},$$

where b_0 and S_0 are fixed with n.

Therefore, if $\{h_i\}_{i=1}^m$ are the ordered singular values of $\tilde{\Omega}_{00}^{-\frac{1}{2}}\alpha\beta^{*\prime}\tilde{\Omega}_{11}^{\frac{1}{2}}$, we define the equivalent class of local alternatives

$$H_{1n}^{\dagger}(0,q_0):\ b=n^{-\frac{1}{2}}b_0,$$

with $h_1,\ldots,h_{q_0}$ being positive constants, $h_j=k_j/b_0$, and the remaining $m-q_0$ singular values of S_0 being equal to zero.

Then, under $H_{1n}^{\dagger}(0,q_0)$, $\sqrt{n}\tilde{\lambda}_0$ and the rank statistic $\widetilde{\mathrm{rk}}_n(0)$ have the same asymptotic behavior as under $H_{1n}(0,q_0)$, where the noncentrality parameter can be written as $\sum_{j=1}^{q_0}k_j^2=b_0^2\sum_{j=1}^{q_0}h_j^2$. In this case the relevant regressor is $Z_{t-1}^{(b_0/\sqrt{n})}$ which in the limit, as well as $Z_{t-1}^{(\hat{b})}$ under (16), leads to the local regressor $J(L)\Delta X_t$, the natural one to test for local alternatives around $b=0$, and to the proportionally factor $\sum_{j=1}^{\infty}j^{-2}=\pi^2/6$ in the variance matrix of $Z_{t-1}^{(0)}$ replacing $\sum_{j=1}^{\infty}\psi_j(b)^2$.

Remark 6. Working with the two step procedure based on estimation of (26), we can set up the local alternatives $H_{1n}(r_0,q_0)$ in terms of the singular values of $\bar{\Theta}$, and find that $\overline{\mathrm{rk}}_n(r_0)$ has power against H_{1n}. However, the direct comparison of both procedures, single and two step, is not easy in this framework since they are based on different coefficient matrices, Π^* and Π, and different projection strategies.

6. Simulations

In this section we examine the finite sample performance of the proposed test by means a small Monte Carlo experiment. The data are generated according

ACCEPTED MANUSCRIPT

to the triangular model

$$\begin{cases} (1-L)^{1-b}(X_{1t} - \Xi' X_{2t}) = \varepsilon_{1t} 1_{\{t>0\}} \\ (1-L)X_{2t} = \varepsilon_{2t} 1_{\{t>0\}} \end{cases} \qquad t = 1, \ldots, n,$$

see Nielsen and Shimotsu [2007]. The dimension of the system is set to $m = 4$ and $r_0 = 2$ cointegration relationships are imposed. The cointegrating matrix is given by $\beta = [I_2, -\Xi']'$, with $\Xi = [[1, 0.5]', [0.5, 1]']$ and $[\varepsilon_{1t}', \varepsilon_{2t}']'$ is an iid(0,I_4) gaussian sequence. We also consider a model with short run dynamics following (23) with $p = 1$ and $\Phi_1 = 0.6 \cdot I_4$ by replacing X_t by $\Phi(L) X_t$ in the DGP. Let $X_t = (X_{1t}', X_{2t}')'$ and $x_{i,t}$ the i-th component of X_t, $i = 1, \ldots, 4$. The memory parameter b is estimated from the OLS residuals $\hat{e}_t$ of the auxiliary regression

$$x_{1,t} = \theta' \boldsymbol{z}_t + e_t, \quad \text{with} \quad \boldsymbol{z}_t = [x_{2,t}, x_{3,t}, x_{4,t}]'$$

by means of the Exact Local Whittle (hereafter ELW) estimator by Shimotsu and Phillips [2005]. More precisely we apply the ELW to estimate the memory δ of the series $\hat{e}_t = x_{1t} - \hat{\theta}' \boldsymbol{z}_t$ over the compact set $[0.500001, 1]$, with $\delta = 1 - b$, so that $\hat{b} := 1 - \hat{\delta}$ satisfies condition (16) in Theorem 2.

Our inference procedure consist in testing $H_0'(r_0) : \Lambda_{r_0} = 0$. If $r_0 > r$, then $\text{rank}(A_{r_0} B_{r_0}) < r_0$ and $H_0'(r_0)$ is not equivalent to $H_0(r_0) : \text{rank}(\Theta) = r_0$. To avoid this case, we suggest to perform the test $H_0'(r_0)$, for $r_0 = 0, 1, \ldots, \hat{r}$, where $\hat{r} < m$ is the first value assumed by the index r_0 for which we can not reject $H_0'(r_0)$. In other words, the test rejects $H_0'(r_0)$ for $r_0 = 0, 1, \ldots, \hat{r} - 1$ but not $H_0'(\hat{r})$. Then $\hat{r}$ is a consistent estimator of r. In our experiment we test $H_0'(r_0)$, $r_0 = 0, 1, 2$. The tables report the percentage of rejection of the null hypothesis $H_0'(r_0)$ at the nominal level of 5%. The number of replications is 50,000, the sample lengths n are $100, 200, 500$. The parameter b takes the values $b = \{0.1, 0.2, 0.3, 0.4, 0.499999\}$. When we test for no cointegration ($r_0 = 0$), we also include $b = 0$. The first row of Tables 1 and 2 indicate the memory of the cointegrating residuals $1 - b$, approximated to the first decimal number. We also report results for Johansen's LR (employing the critical values computed by Mackinnon, Haug, and Michelis [1999], while BH indicates Breitung and Hassler [2002]'s trace statistic results. The rank statistics $\widetilde{\text{rk}}_n(r_0)$ and $\overline{\text{rk}}_n(r_0)$ proposed in this paper have been computed employing different bandwidths M for the ELW estimation of b in the auxiliary univariate regression; $[z]$ denotes the largest integer smaller of equal to z. The (unfeasible) statistic $\text{rk}_n(r_0)$ is computed using the "true" b.

Our Wald tests are implemented by solving the generalized eigenvalue problem (18). We have also tried the alternative scaling of the matrix $\tilde{\Pi}$ obtained by pre-multiplying this matrix by $S_{00.2}^{-\frac{1}{2}}$, with $S_{00.2} = S_{00} - S_{02} S_{22}^{-1} S_{20}$. From simulations unreported here, it emerged that building test statistics on the matrix $S_{00.2}^{-\frac{1}{2}} \tilde{\Pi} S_{22}^{\frac{1}{2}}$ the size is upper-biased especially for $T = 100$ and large b, but power increases.

Table 1 contains the results for tests within the model with $p = 0$. The first panel of Table 1 provides the percentage of rejections when the hypothesis

ACCEPTED MANUSCRIPT

$H_0(0)$ of no cointegration is tested. As $b \to 0$ all the eigenvalues of Π approach zero. In this case, corresponding to the last column of the table ($1 - b = 1$), X_t can be considered as equivalent to a multivariate random walk with full rank covariance matrix and therefore the null is true. The empirical rejection frequencies of $\widetilde{\mathrm{rk}}_n(0)$ are above the nominal level, because the estimation of b leads to an increase in the sampling variation of the test statistic. The distortion is higher for small sample sizes and more narrow bandwidths.

In the second panel of Table 1 we consider the rejection frequencies for the hypothesis that the cointegration rank is one. The test shows good power also in small samples for $b \geq 0.3$. For $1 - b = 0.9$ the rank of $A_r B_r$ is near to be zero and the power of the test is very poor for $n \leq 200$. This indicates that large samples are needed for detecting very weak cointegration relationships. Simulations unreported here show that for $n = 1000$ the size for $1 - b = 0.9$ is still around 2%, showing that is very complicated to estimate precisely very weak cointegration relationships with small b. Similar arguments can explain the empirical size of the test, examined in the third panel of Table 1. We also note that in this simple framework the behavior of the proposed test is not greatly affected from the first step estimation of $\hat{b}$, at least for $n \geq 200$. Moreover its power is superior to Johansen's test and the trace statistic by Breitung and Hassler [2002] while the size is comparable.

When $p = 1$ we consider three alternative testing procedures: a first one labeled OLS based on (18) after projecting on one lag of all variables, cf. (25); a two step ($2S$) procedure based on (26) and $\overline{\mathrm{rk}}_n(r_0)$, where the estimate of Φ_1 in $\hat{\Phi}(L)$ is obtained in the initial OLS estimation; and an iterated procedure ($ITER$) based on the same equation, but where $\hat{\Phi}_1$ is updated 10 times from the previous fit of (26). We only report results for $n = 500$, since the increased number of parameters to be estimated made procedures quite erratic for smaller sample sizes. The results reported in Table 2 are comparable to those in Table 1. Here all tests show more overrejection for $H_0'(0)$ when $b \to 0$, OLS being the most distorted, while $2S$ seems to be the most accurate. However in the third panel the empirical size for $H_0(2)$ of OLS and the iterated procedure is closer to the nominal one. Power comparisons are difficult because of these distortions and dependence on the estimated b. Size adjusted power results not reported here seem to confirm that OLS provides a good overall performance, though for some particular cases iterated procedures can be more powerful. In any case, these preliminary findings should be confirmed in a more detailed analysis with more DGPs and parameter values.

7. Conclusion and directions for further research

In this paper we have introduced a simple Wald test for determining the cointegration rank of a nonstationary system, allowing to the cointegrating residuals to be fractionally integrated of unknown order. The test is regression based but can be easily implemented solving a generalized eigenvalue problem of the type proposed by Johansen [1988]. Many directions for further investigation could be suggested. First, we only allow for weak cointegration leading to standard

ACCEPTED MANUSCRIPT

asymptotics. If $b > 1/2$ then the limit distribution of the test is not standard and bootstrap techniques could be employed, following Davidson [2002, 2006]. Second, we recall that we used the ELW estimator by Shimotsu and Phillips [2005] without formally proving that the assumption of the power rate consistency of $\hat{b}$ was satisfied. This issue deserves further investigation. The finite sample performance of these estimates could be improved applying spectral regression techniques to the univariate auxiliary regression. Third, in the presence of short run correlation in the ECM we propose to use linear unrestricted estimation and two-step procedures. However efficiency improvements could be achieved in finite samples by using nonlinear estimates that exploit the multiplicative structure of the matrix coefficients.

Appendix

Proof of Lemma 1. The assumption that $\alpha_\perp'\beta_\perp$ have full rank is equivalent to the condition that $\beta'\alpha$, $(\beta, \alpha_\perp)$ have full rank (see Exercises 3.7 and 4.12 in Johansen [1996]). Premultiply equation (4) by $\alpha_\perp'$ and β'

$$\alpha_\perp'\Delta^d X_t = \alpha_\perp'\varepsilon_t \tag{27}$$

$$\Delta^{d-b}\beta'\Delta^b X_t = \beta'\alpha\Delta^{d-b}(1-\Delta^b)\beta' X_t + \beta'\varepsilon_t \tag{28}$$

We rewrite (28) as $\beta'\Pi(L)X_t = \Delta^{d-b}(I_r - (I_r + \beta'\alpha)L_b)\beta' X_t = \beta'\varepsilon_t$ and define the polynomial characteristic polynomial associated to the process $\beta' X_t$

$$\zeta(z) = (1-z)^{d-b}\left(I_r - (I_r + \beta'\alpha)\left(1 - (1-z)^b\right)\right)$$

and the polynomial $\zeta^*(u) = (I_r - (I_r + \beta'\alpha)u)$. These are related via the transformation $u(z) = 1 - (1-z)^b$, so that $\zeta(z) = (1-z)^{d-b}\zeta^*(1-(1-z)^b)$. Since $\beta'\alpha$ has full rank, then $\text{eig}\{I_r + \beta'\alpha\} < 1$ (see again Exercise 4.12 in Johansen [1996]) and therefore $|\zeta^*(u)| \neq 0$ for $|u| \leq 1$, so that $\zeta^{*^{-1}}(u) = \sum_{j=0}^{\infty}(I_r + \beta'\alpha)^j u^j$ is regular and continuous for $|u| \leq 1$. Following similar arguments to Johansen (2008), Theorem 8, the function $u(z) = 1 - (1-z)^b$ is also regular for $|z| < 1$, and continuous for $|z| \leq 1$, when $b > 0$. Hence the compound function $F(z) = \zeta^{*^{-1}}(1-(1-z)^b)$ is continuous for $|z| \leq 1$ and regular without singularities on $|z| < 1$. Therefore, by Lemma A.6 in Johansen (2008) it has an expansion $F(z) = \sum_{j=0}^{\infty} F_j z^j$, $|z| < 1$, where the coefficients satisfy $\sum_{j=0}^{\infty}\|F_j\|^2 < \infty$. We define the zero mean stationary process $W_t = \sum_{j=0}^{\infty} F_j\beta'\varepsilon_{t-j}$. Using the identity $\beta_\perp(\alpha_\perp'\beta_\perp)^{-1}\alpha_\perp' + \alpha(\beta'\alpha)^{-1}\beta' = I_m$ and the solution of (27), given by $\alpha_\perp' X_t = \Delta^{-d}\alpha_\perp'\varepsilon_t$, we find that

$$X_t = \beta_\perp(\alpha_\perp'\beta_\perp)^{-1}\alpha_\perp'\Delta^{-d}\varepsilon_t + \alpha(\beta'\alpha)^{-1}\Delta^{-d+b}W_t$$

and (6) follows defining $Y_t = \alpha(\beta'\alpha)^{-1}W_t$. Then the spectral density matrix of $\beta'\Delta^{d-b}X_t = W_t$ is given by$f_W(\omega) = (2\pi)^{-1}F(e^{-i\omega})\beta'\Sigma_\varepsilon\beta F(e^{i\omega})'$ where $f_W(0) = (\beta'\alpha)^{-1}\beta'\Sigma_\varepsilon\beta(\beta'\alpha)^{-1'} > 0$. □

ACCEPTED MANUSCRIPT

Proof of Lemma 2. Using that $\mathrm{Cov}(Z_{t-1}^{(b)}, \varepsilon_t) = 0$, we have that $\Omega_{00} = \Pi\Omega_{11}\Pi' + \Sigma_\varepsilon > 0$, because $\Pi\Omega_{11}\Pi'$ is semidefinite positive. The full rank of Ω_{11} follows using the definition of $Z_{t-1}^{(b)}$ and the full rank of Ω_{00}, since $Z_{t-1}^{(b)}$ is obtained from ΔX_t with the scalar filter $\sum_{j=1}^{t} \psi_j(b) L^j$, with limit transfer function $\left|\sum_{j=1}^{\infty} \psi_j(b) e^{i\omega j}\right|^2 > 0$ for all $\omega \in [0, \pi]$ and all $b > 0$. □

Proof of Lemma 3. If $\hat{\Theta}$ is a consistent estimator of Θ and null hypothesis $H_0(r_0)$ holds true, then $\|\hat{\Theta} - A_{r_0}B_{r_0}\| \xrightarrow{p} 0$, where A_{r_0} and B_{r_0} are of full rank r_0. Therefore, pre- and post-multiplying $\hat{\Theta} - \Theta$ by the matrices $(\hat{A}_{r_0}, \hat{A}_{r_0,\perp})'$ and $[\hat{B}'_{r_0}, \hat{B}'_{r_0,\perp}]$ we obtain

$$\boldsymbol{R} = \begin{bmatrix} \hat{A}'_{r_0}(\hat{A}_{r_0}\hat{B}_{r_0} - A_{r_0}B_{r_0})\hat{B}'_{r_0} & -\hat{A}'_{r_0}(A_{r_0}B_{r_0})\hat{B}'_{r_0,\perp} \\ -\hat{A}'_{r_0,\perp}(A_{r_0}B_{r_0})\hat{B}'_{r_0} & \hat{\Lambda}_{r_0} - \hat{A}'_{r_0,\perp}(A_{r_0}B_{r_0})\hat{B}'_{r_0,\perp} \end{bmatrix} \xrightarrow{p} 0. \quad (29)$$

We label $\boldsymbol{R}_{ij}$, $i, j = 1, 2$ the sub-matrices obtained partitioning the matrix $\boldsymbol{R}$ as in (29). The singular values of $\hat{\Theta}$ are equal to $\sqrt{\mathrm{eig}(\hat{\Theta}'\hat{\Theta})}$ and the eigenvalues are continuous function of $\hat{\Theta}$. Therefore Slutsky's theorem implies that the smallest $m-r$ singular values of $\hat{\Theta}$ converge in probability to the $m-r$ null singular values of $\Theta = A_{r_0}B_{r_0}$, that is $\hat{\Lambda}_{r_0} \xrightarrow{p} 0$. It follows that also the second component of the submatrix $\boldsymbol{R}_{22}$, i.e. $\hat{A}'_{r_0,\perp}(A_{r_0}B_{r_0})\hat{B}'_{r_0,\perp}$, converges in probability to the null matrix. Considering jointly $\boldsymbol{R}_{12}$ and $\boldsymbol{R}_{22}$, we have $\hat{A}'_{r_0}(A_{r_0}B_{r_0})\hat{B}'_{r_0,\perp} \xrightarrow{p} 0$, and $\hat{A}'_{r_0,\perp}(A_{r_0}B_{r_0})\hat{B}'_{r_0,\perp} \xrightarrow{p} 0$. It allows us to conclude that $B_{r_0}\hat{B}'_{r_0,\perp} \xrightarrow{p} 0$, since it is not possible that both $\hat{A}'_{r_0}A_r \xrightarrow{p} 0$ and $\hat{A}'_{r_0,\perp}A_{r_0} \xrightarrow{p} 0$ hold true. Then $\hat{B}_{r_0,\perp} \xrightarrow{p} B_{r_0,\perp}$, and this implies, directly from the identity $\hat{B}_{r_0,\perp}\hat{B}'_{r_0} \equiv 0 \Leftrightarrow (\hat{B}_{r_0,\perp} - B_{r_0,\perp})B'_{r_0} \equiv \hat{B}_{r_0,\perp}(B_{r_0} - \hat{B}_r)'$ so that also $\hat{B}_{r_0} \xrightarrow{p} B_{r_0}$.

The consistency of $\hat{A}_{r_0}$ and $\hat{A}_{r_0,\perp}$ can be derived by the same arguments, considering the blocks $\boldsymbol{R}_{21}$ and $\boldsymbol{R}_{22}$ and the identity $\hat{A}'_{r_0}\hat{A}_{r_0,\perp} \equiv 0$. □

Proof of Theorem 1. As discussed in Section 3, the proof of Theorem 1 we provide is not a simple application to fractional cointegration of Kleibergen and Paap's (2006) Theorem 1, which relies on root-n consistency of $\hat{A}_{r_0}$ and $\hat{B}_{r_0}$. We instead only use consistency of $\hat{A}_{r_0}$ and $\hat{B}_{r_0}$ under H_0. Under Assumption 1, ΔX_t and $Z_{t-1}^{(b)}$ are asymptotically stationary and using a standard central limit theorem for martingale difference sequences it follows that $\sqrt{n}\mathrm{vec}(\hat{\Pi} - \Pi) \xrightarrow{d} \mathcal{N}(0, \Omega_{11}^{-1} \otimes \Sigma_\varepsilon)$. Using the decomposition $\hat{\Theta} - \Theta = \hat{A}_{r_0}\hat{B}_{r_0} - A_{r_0}B_{r_0} + \hat{A}_{r_0,\perp}\hat{\Lambda}_{r_0}\hat{B}_{r_0,\perp}$, and recalling that $\Theta = \Omega_{00}^{-\frac{1}{2}}\Pi\Omega_{11}^{\frac{1}{2}}$, we have that under $H_0(r_0)$

$$\sqrt{n}\mathrm{vec}\left(\hat{A}_{r_0}\hat{B}_{r_0} - A_{r_0}B_{r_0}\right) + \sqrt{n}\ \mathrm{vec}\left(\hat{A}_{r_0,\perp}\hat{\Lambda}_{r_0}\hat{B}_{r_0,\perp}\right) \xrightarrow{d} N(0, \mathcal{V}) \quad (30)$$

ACCEPTED MANUSCRIPT

where $\mathcal{V} = (\Omega_{11}^{\frac{1}{2}} \otimes \Omega_{00}^{-\frac{1}{2}})(\Omega_{11}^{-1} \otimes \Sigma_\varepsilon)(\Omega_{11}^{\frac{1}{2}} \otimes \Omega_{00}^{-\frac{1}{2}})' = I_m \otimes \Omega_{00}^{-\frac{1}{2}} \Sigma_\varepsilon \Omega_{00}^{-\frac{1}{2}}$, with

$$\Omega_{00}^{-\frac{1}{2}} \Sigma_\varepsilon \Omega_{00}^{-\frac{1}{2}} = \Omega_{00}^{-\frac{1}{2}} \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \mathrm{E}(\Delta X_t - \Pi Z_{t-1}^{(b)})(\Delta X_t - \Pi Z_{t-1}^{(b)})' \Omega_{00}^{-\frac{1}{2}}$$

$$= \quad I_m - \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \mathrm{E}\left(\Omega_{00}^{-\frac{1}{2}} \Delta X_t Z_{t-1}^{(b)\prime} \Omega_{11}^{-\frac{1}{2}} \Theta' - \Theta \Omega_{11}^{-\frac{1}{2}} Z_{t-1}^{(b)} \Delta X_t' \Omega_{00}^{-\frac{1}{2}} \right) + \Theta\Theta',$$

so that $A_{r_0,\perp}' \Omega_{00}^{-\frac{1}{2}} \Sigma_\varepsilon \Omega_{00}^{-\frac{1}{2}} A_{r_0,\perp} = A_{r_0,\perp}' A_{r_0,\perp} = I_m$ because $A_{r_0,\perp}' \Theta = 0$ under $H_0(r_0)$. Using that $A_{r_0}' A_{r_0,\perp} \equiv 0$, $\hat{B}_{r_0,\perp} \hat{B}_{r_0,\perp}' = 0$, $A_{r_0,\perp}' A_{r_0,\perp} = I_{m-r_0}$, and the consistency of $\hat{A}_{r_0,\perp}$ from (30), we get

$$\sqrt{n}\mathrm{vec}\left\{ A_{r_0,\perp}' (\hat{\Theta} - \Theta) \hat{B}_{r_0,\perp}' \right\} = (1 + o_p(1)) \sqrt{n} \hat{\lambda}_{r_0} \xrightarrow{d} \mathcal{N}(0, I_{m-r_0} \otimes I_{m-r_0}),$$

proving the theorem. □

Proof of Theorem 2. We first show the consistency of the estimator. Consider the regression model

$$\Delta X_t = \Pi Z_{t-1}^{(b)} + \varepsilon_t = \Pi Z_{t-1}^{(\hat{b})} + \Pi(Z_{t-1}^{(b)} - Z_{t-1}^{(\hat{b})}) + \varepsilon_t. \tag{31}$$

In order to show that $\tilde{\Pi}$ is consistent, we rewrite $\tilde{\Pi} - \Pi$ as

$$= \quad \Pi \sum_{t=2}^{n} \left\{ \frac{(\Delta^{-b} - \Delta^{-\hat{b}})}{b} \Delta X_t \right\} Z_{t-1}^{(\hat{b})\prime} \left(\sum_{t=2}^{n} Z_{t-1}^{(\hat{b})} Z_{t-1}^{(\hat{b})\prime} \right)^{-1} + O_p(n^{-\tau}) \tag{32}$$

$$+ \frac{1}{n} \sum_{t=2}^{n} \varepsilon_t Z_{t-1}^{(\hat{b})} \left(\frac{1}{n} \sum_{t=2}^{n} Z_{t-1}^{(\hat{b})} Z_{t-1}^{(\hat{b})\prime} \right)^{-1}. \tag{33}$$

We consider the single terms distinctly, and firstly (32). We first note that

$$b Z_{t-1}^{(b)} - \hat{b} Z_{t-1}^{(\hat{b})} = (\Delta^{-b} - \Delta^{-\hat{b}}) \Delta X_t = \sum_{j=1}^{t-1} \left\{ \psi_j(b) - \psi_j(\hat{b}) \right\} \Delta X_{t-j}. \tag{34}$$

Following Lobato and Velasco [2007] and Robinson and Hualde [2003], Proposition 9, we note that, for $j = 1, 2, \ldots, n$, the expression $\left\{ \psi_j(b) - \psi_j(\hat{b}) \right\}$ in (34) equals

$$\sum_{q=1}^{Q-1} \frac{1}{q!} \left(b - \hat{b} \right)^q \psi_j^{(q)}(b) + \frac{1}{Q} \left(b - \hat{b} \right)^Q \psi_j^{(Q)}(\bar{b}) \tag{35}$$

where, $\bar{b}$ is an intermediate stochastic point between b and $\hat{b}$ and $\psi_j^{(q)}(b) =$

ACCEPTED MANUSCRIPT

$\mathrm{d}^q\psi_j(b)/\mathrm{d}b^q$. It follows that

$$\frac{1}{n}\sum_{t=2}^{n}\left[(\Delta^{-b}-\Delta^{-\hat{b}})\Delta X_t\right]Z_{t-1}^{(\hat{b})'}=\frac{1}{n}\sum_{t=2}^{n}\left(\sum_{j=1}^{t-1}(\psi_j(b)-\psi_j(\hat{b}))\Delta X_{t-j}\right)Z_{t-1}^{(\hat{b})'}$$

$$=\sum_{q=1}^{Q-1}\frac{1}{q!}\left(\hat{b}-b\right)^q\left[\frac{1}{n}\sum_{t=2}^{n}\left(\sum_{j=1}^{t-1}\psi_j^{(q)}(b)\Delta X_{t-j}\right)Z_{t-1}^{(\hat{b})'}\right] \tag{36}$$

$$+\frac{1}{Q!}\left(\hat{b}-b\right)^Q\frac{1}{n}\left[\sum_{t=2}^{n}\left(\sum_{j=1}^{t-1}\psi_j^{(Q)}(\bar{b})\Delta X_{t-j}\right)Z_{t-1}^{(\hat{b})'}\right] \tag{37}$$

Since $|\psi_j^{(q)}(b)|\leq C(1+j)^{b-1}(\log(1+j))^q,\ j=1,2,\ldots n,\ 1\leq q\leq Q$ (see Robinson and Hualde [2003], Lemma D1) the sequence $\psi_j^{(q)}(b)$ is square summable when $b<1/2$, then the term in square brackets in (36) is $O_p(1)$. Moreover, for $\hat{b}<1/2$, $n^{-1}\sum_{t=2}^{n}Z_{t-1}^{(\hat{b})}Z_{t-1}^{(\hat{b})'}=O_p(1)$, then if $\hat{b}-b=O_p(n^{-\tau})$, $\tau>0$, (36) is $O_p(n^{-\tau})$.

In order to analyze (37), we focus on the term:

$$\frac{1}{n}\sum_{t=2}^{n}\sum_{j=1}^{t-1}\psi_j^{(Q)}(\bar{b})\Delta X_{t-j}^{(\kappa)}Z_{t-1}^{(\hat{b},\ell)}\leq\frac{1}{n}\sqrt{\sum_{t=2}^{n}\left|\sum_{j=1}^{t-1}\psi_j^{(Q)}(\bar{b})\Delta X_{t-j}^{(\kappa)}\right|^2}\sqrt{\sum_{t=2}^{n}Z_{t-1}^{(\hat{b},\ell)^2}}$$

where $\Delta X_{t-j}^{(\kappa)}, Z_{t-1}^{(\hat{b},\ell)}$ are, respectively, the κ-th and ℓ-th elements of the vectors ΔX_{t-j} and $Z_{t-1}^{(\hat{b})}$. The first term of the product in the right hand side of the above equation can be bounded by

$$\sup_t\left|\sum_{j=1}^{t-1}\psi_j^{(Q)}(\bar{b})\Delta X_{t-j}^{(\kappa)}\right|\leq\sum_{j=1}^{n}\left|\psi_j^{(Q)}(\bar{b})\right|\left|\Delta X_{t-j}^{(\kappa)}\right| \tag{38}$$

$$\leq\left(\sum_{j=1}^{n}\left|\psi_j^{(Q)}(\bar{b})\right|^2\sum_{j=1}^{n}\left|\Delta X_{t-j}^{(\kappa)}\right|^2\right)^{1/2}.$$

By Lemma D.5 of Robinson and Hualde [2003], for any $\epsilon>0$,

$$\left|\psi_j^{(Q)}(\bar{b})\right|=O_p\left((j+1)^{b+\epsilon-1}(\log(j+1))^Q\right),\quad\text{as}\quad n\to\infty \tag{39}$$

so that $\sum_{j=1}^{n}\left|\psi_j^{(Q)}(\bar{b})\right|^2\leq Cn^{2b+2\epsilon-1}(\log n)^{2Q}$. Noting that $\sum_{j=1}^{n}(\Delta X_{t-j}^{(\kappa)})^2=O_p(n)$, the term (38) is $O_p\left(n^{b+\epsilon}(\log n)^Q\right)$ and taking Q large enough we can make (37) $o_p(n^{-\tau})$, since $n^{-1}\sum_{t=2}^{n}|Z_{t-1}^{(\hat{b},\ell)}|^2$ is $O_p(1)$.

It remains to analyze the asymptotic behavior of the term (33). Since the second factor in this term is $O_p(1)$, we consider the first one scaled by $\sqrt{n}$ and

ACCEPTED MANUSCRIPT

we have that

$$\frac{1}{\sqrt{n}}\left(\sum_{t=2}^{n}\varepsilon_t Z_{t-1}^{(\hat{b})'}-\sum_{t=2}^{n}\varepsilon_t Z_{t-1}^{(b)'}\right)$$

$$= \frac{1}{\sqrt{n}}\sum_{t=2}^{n}\varepsilon_t\sum_{j=1}^{t-1}\left\{\sum_{q=1}^{Q-1}\frac{1}{q!}\left(b-\hat{b}\right)^q\psi_j^{(q)}(b)\right\}\Delta X_{t-j} \tag{40}$$

$$+\frac{1}{\sqrt{n}}\sum_{t=2}^{n}\varepsilon_t\sum_{j=1}^{t-1}\left\{\frac{1}{Q!}\left(b-\hat{b}\right)^Q\psi_j^{(Q)}(\bar{b})\right\}\Delta X_{t-j}, \tag{41}$$

where

$$\frac{1}{\sqrt{n}}\sum_{t=2}^{n}\varepsilon_t\sum_{j=1}^{t-1}\frac{1}{q!}\psi_j^{(q)}(b)\Delta X'_{t-j} = O_p(1), \quad \text{for } q=1,2,\ldots,Q-1$$

$$\frac{1}{\sqrt{n}}\sum_{t=2}^{n}\varepsilon_t\sum_{j=1}^{t-1}\left\{\frac{1}{Q!}\psi_j^{(Q)}(\bar{b})\right\}\Delta X'_{t-j} = O_p(n^{b+\epsilon+1/2}(\log n)^Q).$$

For Q large enough, (41) is $o_p(1)$ and the higher order terms in (40) are $O_p(n^{-\tau})$ Therefore (33) is $o_p(1)$, proving that $\tilde{\Theta}\xrightarrow{p}\Theta$, so that Lemma 2 applies.

From the first order condition for the OLS estimator $\tilde{\Pi}$, we obtain $\sum_{t=2}^{n}(\Delta X_t-\tilde{\Pi}Z_{t-1}^{(\hat{b})'})Z_{t-1}^{(\hat{b})'}=0$, which is equivalent to the equality

$$0=\sum_{t=2}^{n}(S_{00}^{-\frac{1}{2}}\Delta X_t-\tilde{\Theta}S_{22}^{-\frac{1}{2}}Z_{t-1}^{(\hat{b})})Z_{t-1}^{(\hat{b})'}S_{22}^{-\frac{1}{2}}. \tag{42}$$

Given the regression model $S_{00}^{-\frac{1}{2}}\Delta X_t = S_{00}^{-\frac{1}{2}}\Pi S_{22}^{\frac{1}{2}}S_{22}^{-\frac{1}{2}}Z_{t-1}^{(b)}+S_{00}^{-\frac{1}{2}}\varepsilon_t$, (42) can therefore be written as

$$0=\sum_{t=2}^{n}(\Theta^{(n)}S_{22}^{-\frac{1}{2}}Z_{t-1}^{(b)}+S_{00}^{-\frac{1}{2}}\varepsilon_t-\tilde{\Theta}S_{22}^{-\frac{1}{2}}Z_{t-1}^{(\hat{b})})Z_{t-1}^{(\hat{b})'}S_{22}^{-\frac{1}{2}}$$

with $\Theta^{(n)}=S_{00}^{-\frac{1}{2}}\Pi S_{22}^{\frac{1}{2}}$ and $\text{plim}_{n\to\infty}\Theta^{(n)}=\Theta$, and both Θ and $\Theta^{(n)}$ have exactly rank r_0 under the null. Therefore, with the SVD $\Theta^{(n)}=A_{r_0}^{(n)}B_{r_0}^{(n)}$ and using the decomposition (15) we get

$$\tilde{A}_{r_0,\perp}\tilde{\Lambda}_{r_0}\tilde{B}_{r_0,\perp}S_{22}^{-\frac{1}{2}}\left(\frac{1}{n}\sum_{t=2}^{n}Z_{t-1}^{(\hat{b})}Z_{t-1}^{(\hat{b})'}\right)S_{22}^{-\frac{1}{2}}=S_{00}^{-\frac{1}{2}}\left(\frac{1}{n}\sum_{t=2}^{n}\varepsilon_t Z_{t-1}^{(\hat{b})'}\right)S_{22}^{-\frac{1}{2}},$$

$$-\quad\tilde{A}_{r_0}\tilde{B}_{r_0}S_{22}^{-\frac{1}{2}}\left(\frac{1}{n}\sum_{t=2}^{n}Z_{t-1}^{(\hat{b})}Z_{t-1}^{(\hat{b})'}\right)S_{22}^{-\frac{1}{2}}+A_{r_0}^{(n)}B_{r_0}^{(n)}S_{22}^{-\frac{1}{2}}\left(\frac{1}{n}\sum_{t=2}^{n}Z_{t-1}^{(b)}Z_{t-1}^{(\hat{b})'}\right)S_{22}^{-\frac{1}{2}}$$

so that $\tilde{A}_{r_0,\perp}\tilde{\Lambda}_{r_0}\tilde{B}_{r_0,\perp}$ equals

$$A_{r_0}^{(n)}B_{r_0}^{(n)}S_{22}^{-\frac{1}{2}}\left(\frac{1}{n}\sum_{t=2}^{n}Z_{t-1}^{(b)}Z_{t-1}^{(\hat{b})'}\right)S_{22}^{-\frac{1}{2}}+S_{00}\left(\frac{1}{n}\sum_{t=2}^{n}\varepsilon_t Z_{t-1}^{(\hat{b})'}\right)S_{22}^{-\frac{1}{2}}-\tilde{A}_{r_0}\tilde{B}_{r_0}. \tag{43}$$

ACCEPTED MANUSCRIPT

Pre-multiplying the above equation by $A_{r_0,\perp}^{(n)\prime}$, post-multiplying by $\tilde{B}'_{r_0,\perp}$ and noting that $A_{r_0,\perp} = A_{r_0,\perp}^{(n)} + o_p(1) = \tilde{A}_{r_0,\perp} + o_p(1)$ by $\Theta^{(n)} \xrightarrow{p} \Theta$, see Lemma 3, we obtain

$$\sqrt{n}\tilde{\lambda} = (1+o_p(1))^{-1}\left(\tilde{B}_{r_0,\perp}S_{22}^{-\frac{1}{2}} \otimes A'_{r_0,\perp}S_{00}^{-\frac{1}{2}}\right)\frac{1}{\sqrt{n}}\text{vec}\sum_{t=2}^{n}\varepsilon_t Z_{t-1}^{(\hat{b})\prime}.$$

Using (40)-(41) it follows that $n^{-1/2}\text{vec}\sum_{t=2}^{n}\varepsilon_t Z_{t-1}^{(\hat{b})\prime} \xrightarrow{d} N(0, \Omega_{11}\otimes\Sigma_\varepsilon)$, whereas, the proof that $\text{plim}_{n\to\infty} S_{22} = \Omega_{11}$ is based on (35). Proceeding as in the proof of Theorem 1, it is easy to check that, as $n \to \infty$, $\sqrt{n}\tilde{\lambda}_{r_0} \xrightarrow{d} N(0, I_{m-r_0}\otimes I_{m-r_0})$, proving the theorem. □

Proof of Theorem 4 We consider first the case without lags. Under local alternative $H_{1n}(r_0, q_0)$, equation (43) should be substituted by

$$\begin{aligned}\tilde{A}_{r_0,\perp}\tilde{\Lambda}_{r_0}\tilde{B}_{r_0,\perp} &= \left(A_{r_0}B_{r_0} + A_{r_0,\perp}\Lambda_p^{(1,n)}B_{r_0,\perp}\right)S_{22}^{-\frac{1}{2}}\left(\frac{1}{n}\sum_{t=2}^{n}Z_{t-1}^{(b)}Z_{t-1}^{(\hat{b})\prime}\right)S_{22}^{-\frac{1}{2}} \\ &\quad -\tilde{A}_{r_0}\tilde{B}_{r_0} + S_{00}^{-\frac{1}{2}}\left(\frac{1}{n}\sum_{t=2}^{n}\varepsilon_t Z_{t-1}^{(\hat{b})\prime}\right)S_{22}^{-\frac{1}{2}}.\end{aligned}$$

Pre- and post- multiplying the above equation, respectively by $A'_{r_0,\perp}$, by $\tilde{B}'_{r_0,\perp}$ and taking the vec we obtain

$$\begin{aligned}&(1+o_p(1))\sqrt{n}\tilde{\lambda}_{r_0} - \left[\tilde{B}_{r_0,\perp}S_{22}^{-\frac{1}{2}}\left(\frac{1}{n}\sum_{t=2}^{n}Z_{t-1}^{(\hat{b})}Z_{t-1}^{(b)\prime}\right)S_{22}^{-\frac{1}{2}}B'_{r_0,\perp}\otimes I_{m-r_0}\right]\sqrt{n}\lambda_{r_0}^{(1,n)} \\ &= \left(\tilde{B}_{r_0,\perp}S_{22}^{-\frac{1}{2}} \otimes A'_{r_0,\perp}S_{00}^{-\frac{1}{2}}\right)\left(\frac{1}{\sqrt{n}}\text{vec}\sum_{t=2}^{n}\varepsilon_t Z_{t-1}^{(\hat{b})\prime}\right). \qquad (44)\end{aligned}$$

Rewriting $Z_{t-1}^{(b)} = Z_{t-1}^{(\hat{b})} - \left(Z_{t-1}^{(\hat{b})} - Z_{t-1}^{(b)}\right)$, equation (44) can be stated as

$$(1+o_p(1))\sqrt{n}\tilde{\lambda} - \left[\tilde{B}_{r_0,\perp}B'_{r_0,\perp}\otimes I_{m-r_0}\right]\sqrt{n}\lambda_{r_0}^{(1,n)} + \qquad (45)$$

$$+\left\{\tilde{B}_{r_0,\perp}S_{22}^{-\frac{1}{2}}\left[\frac{1}{n}\sum_{t=2}^{n}Z_{t-1}^{(\hat{b})}\left(Z_{t-1}^{(\hat{b})} - Z_{t-1}^{(b)}\right)'\right]S_{22}^{-\frac{1}{2}}B'_{r_0,\perp}\otimes I_{m-r_0}\right\}\sqrt{n}\lambda_{r_0}^{(1,n)}. \qquad (46)$$

Provided that

$$\frac{1}{n}\sum_{t=2}^{n}Z_{t-1}^{(\hat{b})}\left(Z_{t-1}^{(\hat{b})} - Z_{t-1}^{(b)}\right)' = o_p(1) \quad \text{and} \quad \sqrt{n}\lambda_{r_0}^{(1,n)} = O(1) \qquad (47)$$

equation (46) is $o_p(1)$ and the theorem follows because, adapting the proof of Proposition 1 in Kleibergen and Paap [2006], it can be easily shown that

ACCEPTED MANUSCRIPT

$\lambda_{r_0}^{(1,n)'}\lambda_{r_0}^{(1,n)} = \mathrm{tr}(S_{m-r_0}^2) = n^{-1}\sum_{j=1}^{p} k_j^2$. When we have lags, the first order condition is given by,

$$\sum_{t=2}^{n}\left(\tilde{S}_{00}^{-\frac{1}{2}}\Delta X_t - \tilde{\Theta}\tilde{S}_{22}^{-\frac{1}{2}}\tilde{Z}_{t-1}^{(\hat{b})}\right)\tilde{Z}_{t-1}^{(\hat{b})'}\tilde{S}_{22}^{-\frac{1}{2}} = 0 \iff$$
$$\sum_{t=2}^{n}\left(\Theta^{(n)}\tilde{S}_{22}^{-\frac{1}{2}}(\Phi(1)+\Phi^*(L)\Delta)Z_{t-1}^{(b)} - \tilde{S}_{00}^{-\frac{1}{2}}(\Phi(L)-I_m)\Delta X_t\right)\tilde{Z}_{t-1}^{(\hat{b})'}\tilde{S}_{22}^{-\frac{1}{2}}$$
$$+\sum_{t=2}^{n}\left(\tilde{S}_{00}^{-\frac{1}{2}}\varepsilon_t - \tilde{\Theta}\tilde{S}_{22}^{-\frac{1}{2}}\tilde{Z}_{t-1}^{(\hat{b})}\right)\tilde{Z}_{t-1}^{(\hat{b})'}\tilde{S}_{22}^{-\frac{1}{2}} = 0,$$

where $\Theta^{(n)} = \tilde{S}_{00}^{-\frac{1}{2}}\Pi^*\tilde{S}_{22}^{\frac{1}{2}}$, $\sum_{t=2}^{n}\Delta\tilde{X}_t\tilde{Z}_{t-1}^{(\hat{b})'} = \sum_{t=2}^{n}\Delta X_t\tilde{Z}_{t-1}^{(\hat{b})'}$ and the terms ΔX_{t-j} and $\tilde{Z}_{t-1}^{(\hat{b})}$ are orthogonal for $j \leq 1$. Using arguments similar to (44), it can be shown that $\sum_{t=2}^{n}\Delta Z_{t-j}^{(b)}\tilde{Z}_{t-1}^{(\hat{b})'}$, $o_p(1)$ for $j \geq 1$, and

$$\sum_{t=2}^{n}\tilde{S}_{22}^{-\frac{1}{2}}Z_{t-1}^{(b)}\tilde{Z}_{t-1}^{(\hat{b})'}\tilde{S}_{22}^{-\frac{1}{2}} = \sum_{t=2}^{n}\tilde{S}_{22}^{-\frac{1}{2}}Z_{t-1}^{(\hat{b})}\tilde{Z}_{t-1}^{(\hat{b})'}\tilde{S}_{22}^{-\frac{1}{2}} + o_p(1)$$

Recalling that $\sum_{t=2}^{n}Z_{t-1}^{(\hat{b})}\tilde{Z}_{t-1}^{(\hat{b})'} = \sum_{t=2}^{n}\tilde{Z}_{t-1}^{(\hat{b})}\tilde{Z}_{t-1}^{(\hat{b})'}$, this proves that the introduction of the short run dynamics does not change the expression of the drift. □

References

Breitung, J., and U. Hassler (2002): "Inference on the cointegration rank in fractionally integrated process," *Journal of Econometrics*, 110, 167–185.

——— (2006): "A residual-based LM type test against fractional cointegration," *Econometric Theory*, 22, 1091–1111.

Chen, W. W., and C. M. Hurvich (2003): "Semiparametric estimation of multivariate fractional cointegration," *Journal of American Statistical Association*, 98, 629–642.

——— (2006): "Semiparametric estimation of fractional cointegrating subspaces," *Annals of Statistic*, 34, 2939–2979.

Davidson, J. (2002): "A model of fractional cointegration, and test for cointegration using the bootstrap," *Journal of Econometrics*, 110, 187–212.

——— (2006): "Alternative bootstrap procedures for testing cointegration in fractionally integrated processes," *Journal of Econometrics*, 133, 741–777.

Dolado, J. J., J. Gonzalo, and L. Mayoral (2002): "A fractional Dickey-Fuller test for unit roots," *Econometrica*, 70, 1963–2006.

ACCEPTED MANUSCRIPT

		$H_0(0)$						$H_0(1)$					$H_0(2)$				
1-b		**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**
n=100																	
Johansen LR		99.12	86.64	51.87	20.78	8.01	6.16	69.90	32.38	9.74	2.37	0.66	3.50	2.07	0.78	0.25	0.09
BH trace		99.83	96.68	76.73	36.47	10.51	4.91	79.18	49.64	19.29	4.33	0.63	3.23	2.35	1.16	0.32	0.05
	$M = \lceil n^{0.55} \rceil$	99.99	99.20	86.50	46.90	15.94	8.05	93.47	68.11	29.23	6.94	1.29	4.65	3.70	1.94	0.53	0.08
$\widetilde{\mathrm{rk}}_n(r_0)$	$M = \lceil n^{0.65} \rceil$	100	99.15	85.94	45.63	14.88	7.04	93.49	67.84	28.69	6.58	1.16	4.69	3.68	1.87	0.5	0.07
	$M = \lceil n^{0.75} \rceil$	99.99	98.90	84.18	43.17	13.15	6.00	92.93	66.39	27.23	5.95	0.94	4.60	3.50	1.74	0.43	0.05
$\mathrm{rk}_n(r_0)$	(true b)	100	99.27	84.37	40.97	11.26	4.91	94.28	66.90	26.26	5.29	0.71	4.84	3.49	1.62	0.40	0.06
n=200																	
Johansen LR		100	99.57	85.15	38.36	10.19	5.63	97.92	75.30	29.07	5.42	0.85	4.68	3.56	1.75	0.52	0.13
BH trace		100	100	99.69	80.76	21.64	5.08	99.97	98.01	75.28	22.11	1.79	4.23	3.88	3.00	1.18	0.09
	$M = \lceil n^{0.55} \rceil$	100	100	99.94	86.60	27.03	7.04	100	99.74	85.41	28.06	2.48	5.58	5.10	3.97	1.65	0.15
$\widetilde{\mathrm{rk}}_n(r_0)$	$M = \lceil n^{0.65} \rceil$	100	100	99.93	85.91	25.60	6.22	100	99.74	85.00	27.38	2.28	5.58	5.04	3.91	1.57	0.13
	$M = \lceil n^{0.75} \rceil$	100	100	99.90	84.73	24.15	5.64	100	99.67	84.13	26.43	2.10	5.55	4.87	3.70	1.46	0.12
$\mathrm{rk}_n(r_0)$	(true b)	100	100	99.94	84.60	22.71	5.08	100	99.76	84.55	25.62	1.95	5.73	4.86	3.64	1.40	0.11
n=500																	
Johansen LR		100	100	99.11	70.98	16.88	5.44	99.99	98.01	69.44	16.61	1.53	4.89	4.53	2.96	1.06	0.16
BH trace		100	100	100	99.98	62.06	5.08	100	100	99.99	91.26	11.11	4.70	4.70	4.30	3.33	0.58
	$M = \lceil n^{0.55} \rceil$	100	100	100	99.99	65.71	5.83	100	100	100	94.00	12.97	5.44	5.17	4.69	3.68	0.65
$\widetilde{\mathrm{rk}}_n(r_0)$	$M = \lceil n^{0.65} \rceil$	100	100	100	99.99	64.91	5.59	100	100	100	94.12	12.33	5.43	5.03	4.57	3.68	0.64
	$M = \lceil n^{0.75} \rceil$	100	100	100	99.99	64.15	5.41	100	100	100	93.91	12.06	5.41	4.98	4.47	3.60	0.63
$\mathrm{rk}_n(r_0)$	(true b)	100	100	100	99.99	63.49	5.08	100	100	100	94.14	11.72	5.54	5.00	4.47	3.55	0.60

Table 1. The panels $H_0(0)$ and $H_0(1)$ gives the rejection frequencies (in %) respectively of the null hypothesis of no-cointegration and the hypothesis of one cointegration relationship (power of the test) for the case $p = 0$. $1 - b$ is the memory of the cointegrating residuals, so $1 - b = 1$ corresponds to the null hypothesis because $r = 0$, and the displayed frequencies in this case refer to size. In the third panel, the rejection frequencies (in %) of the null hypothesis of two cointegration relationship ($H_0(2)$) is analyzed (size of the test).

		$H_0(0)$						$H_0(1)$					$H_0(2)$				
1-b		**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**
Johansen LR		98.08	73.69	33.84	13.22	6.76	6.65	61.64	21.27	5.23	1.33	0.56	3.63	1.71	0.52	0.28	0.14
BH trace		94.46	74.52	43.35	19.92	9.63	7.17	42.06	18.20	5.76	1.77	0.48	2.29	1.01	0.43	0.08	0.01
OLS	$M = \lceil n^{0.55} \rceil$	97.82	85.27	57.34	29.93	15.01	10.99	58.15	29.04	11.01	3.21	0.98	4.34	2.11	0.68	0.17	0.04
	$M = \lceil n^{0.65} \rceil$	97.28	83.23	53.40	25.60	12.01	9.00	55.77	26.27	8.68	2.29	0.54	3.86	1.83	0.62	0.10	0.04
	$M = \lceil n^{0.75} \rceil$	96.59	80.08	49.74	24.10	11.76	8.97	51.39	22.48	7.61	2.00	0.56	3.03	1.48	0.44	0.09	0.03
	(true b)	97.80	85.69	56.70	28.35	13.19	8.97	58.08	28.88	9.97	2.75	0.69	4.59	2.03	0.63	0.15	0.04
2S	$M = \lceil n^{0.55} \rceil$	91.24	69.71	38.90	17.30	8.38	6.89	38.29	18.43	5.64	1.46	0.49	2.76	1.51	0.54	0.17	0.03
	$M = \lceil n^{0.65} \rceil$	82.12	53.49	25.26	10.63	6.38	6.17	27.06	9.96	2.71	0.59	0.25	1.50	0.65	0.19	0.04	0.01
	$M = \lceil n^{0.75} \rceil$	61.40	31.94	15.09	8.40	5.95	6.09	12.30	3.50	0.99	0.31	0.21	0.68	0.18	0.06	0.02	0.01
	(true b)	95.00	74.45	40.28	15.18	7.11	6.09	47.51	21.57	5.65	0.96	0.32	3.87	1.83	0.52	0.11	0.03
ITER	$M = \lceil n^{0.55} \rceil$	98.39	83.22	48.87	20.48	9.82	7.75	72.86	34.32	9.20	2.08	0.74	4.69	2.15	0.66	0.19	0.02
	$M = \lceil n^{0.65} \rceil$	96.01	73.48	37.28	14.67	7.22	6.43	60.95	23.30	5.63	1.27	0.41	3.51	1.50	0.44	0.11	0.02
	$M = \lceil n^{0.75} \rceil$	86.51	54.73	25.88	11.52	6.80	6.37	36.70	11.21	2.63	0.77	0.33	1.96	0.78	0.19	0.06	0.01
	(true b)	99.48	87.13	47.49	17.60	7.94	6.37	80.45	34.74	7.69	1.55	0.44	5.18	2.46	0.61	0.10	0.01

Table 2. Same hypotheses as in Table 1 tested for the case $p = 1$. OLS indicates tests based on (18); $2S$ tests are based on (26), where the estimate of Φ_1 in $\hat{\Phi}(L)$ is obtained in the initial OLS estimation; $ITER$ is based on the same equation, with $\hat{\Phi}_1$ updated 10 times from the previous fit of (26).

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Engle, R., and C. W. J. Granger (1987): "Cointegration and error correction: representation, estimation and testing," *Econometrica*, 55, 251–276.

Gil-Alaña, L. (2004): "Testing of fractional cointegration in macroeconomic time series," *Oxford Bulletin of Economics and Statistics*, 65, 517–529.

Granger, C. W. J. (1986): "Developments in the study of cointegrated economic variables," *Oxford Bulletin of Economics and Statistics*, 48, 213–228.

Hualde, J., and P. M. Robinson (2006): "Semiparametric estimation of fractional cointegration," Working Paper 07/06, Universidad de Navarra.

Hualde, J., and C. Velasco (2008): "Distribution-Free test of fractional cointegration," *Econometric Theory*, 24, 216–255.

Johansen, S. (1988): "Statistical analysis of cointegration vectors," *Journal of Economic Dynamics and Control*, 12, 231–254.

——— (1991): "Estimation and hypothesis testing of cointegrating vectors in Gaussian vector autoregressive models," *Econometrica*, 59, 1551–1580.

——— (1996): *Likelihood based inference on cointegrated vector autoregressive model*. Oxford University Press.

——— (2002): "Cointegration in partial systems and efficiency of single equation analysis," *Journal of Econometrics*, 52, 389–452.

——— (2008): "A representation theory for a class of vector autoregressive models for fractional processes," *Econometric Theory*, 24, 651–676.

Kleibergen, F., and R. Paap (2006): "Generalized reduced rank tests using singular value decomposition," *Journal of Econometrics*, 133, 97–126.

Kleibergen, F., and H. K. van Dijk (1998): "Bayesian simultaneous equations analysis using reduced rank structures," *Econometric Theory*, 14, 701–743.

Lasak, K. (2007): "Likelihood based testing for fractional cointegration," PhD Thesis, UAB.

Lobato, I., and C. Velasco (2007): "Efficient Wald test for fractional unit roots," *Econometrica*, 75, 575–589.

Mackinnon, J. G., A. A. Haug, and L. Michelis (1999): "Numerical distribution functions of likelikood ratio tests for cointegration," *Journal of Applied Econometrics*, 14, 563–577.

Marinucci, D., and P. M. . Robinson (2001): "Semiparametric fractional cointegration analysis," *Journal of Econometrics*, 105, 225–247.

Marmol, F., and C. Velasco (2004): "Consistent testing of cointegrating relationships," *Econometrica*, 72, 1809–1844.

ACCEPTED MANUSCRIPT

——— (2005): "Consistent testing of cointegration of unknown rank," Universidad Carlos III de Madrid. Preprint.

Nielsen, M. Ø. (2004a): "Optimal residual-based test for fractional cointegration and exchange rate dynamics," *Journal of Business Economic Statistics*, 22, 331–345.

——— (2004b): "Spectral analysis of fractionally cointegrated systems," *Economics Letters*, 83, 225–231.

——— (2005): "Multivariate Lagrange multiplier test for fractional integration," *Journal of Financial Econometrics*, 3, 372–398.

Nielsen, M. Ø., and K. Shimotsu (2007): "Determining the cointegrating rank in nonstationary fractional systems by exact local Whittle approach," *Journal of Econometrics*, 141, 574–596.

Robinson, P. M. (1994): "Efficient tests of nonstationary hypotheses," *Journal of the American Statistical Association*, 89, 1420–1437.

——— (1995): "Gaussian semiparametric estimation of long range dependence," *Annals of Statistic*, 23, 1630 – 1661.

——— (2008): "Diagnostic testing for cointegration," *Journal of Econometrics*, 143, 206–225.

Robinson, P. M., and J. Hualde (2003): "Cointegration in fractional systems with unknown integration orders," *Econometrica*, 71, 1727–1766.

Robinson, P. M., and D. Marinucci (2003): "Semiparametric frequency domain analysis of fractional cointegration," in *Time series with long memory*, ed. by P. M. Robinson, pp. 334–373. Oxford University Press.

Robinson, P. M., and Y. Yajima (2002): "Determination of cointegrating rank in fractional systems," *Journal of Econometrics*, 106, 217–241.

Shimotsu, K., and P. C. B. Phillips (2005): "Exact local Whittle estimation of fractional integration," *The Annals of Statistics*, 33, 1890–1933.

Wooldrige, J. (1991): "Notes on regression with difference-stationary data," Preprint, Michigan State Univeristy.