Likelihood based testing for no fractional cointegration

Katarzyna Łasak*
CREATES & School of Economics and Management
University of Aarhus, Building 1322
DK-8000 Aarhus C, Denmark
email: klasak@creates.au.dk

July 14, 2008

Abstract

We consider two likelihood ratio tests, so-called maximum eigenvalue and trace tests, for the null of no cointegration when fractional cointegration is allowed under the alternative, which is a first step to generalize the so-called Johansen's procedure to the fractional cointegration case. The standard cointegration analysis only considers the assumption that deviations from equilibrium can be integrated of order zero, which is very restrictive in many cases and may imply an important loss of power in the fractional case. We consider the alternative hypotheses with equilibrium deviations that can be mean reverting with order of integration possibly greater than zero. Moreover, the degree of fractional cointegration is not assumed to be known, and the asymptotic null distribution of both tests is found when considering an interval of possible values. The power of the proposed tests under fractional alternatives and size accuracy provided by the asymptotic distribution in finite samples are investigated.

Keywords: Error correction model, Gaussian VAR model, Maximum likelihood estimation, Fractional cointegration, Likelihood ratio tests, fractional Brownian motion.

JEL: C12, C15, C32.

*I am grateful to Marco Avarucci, Juan José Dolado, Jesús Gonzalo, Niels Haldrup, Javier Hualde, Søren Johansen, Carlos Velasco and all other people who committed to this piece of research. I would like to thank the editor and three anonymous referees for helpful comments and suggestions. Annette Andersen's help with proof reading is appreciated. Financial support from CREATES, funded by Danish National Research Foundation is acknowledged. A previous version of this paper has been included in the Ph.D. thesis defended at the Department of Economics and Economic History, UAB, Spain.

1 Introduction

Cointegration is commonly thought of as a stationary relation between nonstationary variables. It has become a standard tool in econometrics since the seminal paper of Granger (1981). Following the initial suggestion of Engle and Granger (1987), when the series of interest are $I(1)$, testing for cointegration in a single-equation framework can be conducted by means of residual-based tests (cf. Phillips and Ouliaris (1990)). Residual-based tests rely on initial regressions among the levels of the relevant time series. They are designed to test the null of no cointegration by testing whether there is a unit root in the residuals against the alternative that the regression errors are $I(0)$.

Alternatively fully parametric inference on $I(1)/I(0)$ cointegrated systems in the framework of Error Correction Mechanism (ECM) representation has been developed by Johansen (1988, 1991, 1995). He suggests a maximum likelihood procedure based on reduced rank regressions. His methodology consists in identifying the number of cointegration vectors within the vector autoregression (VAR) model by performing a sequence of likelihood ratio tests. If the variables are cointegrated, cointegration vectors, the speed of adjustment to the equilibrium coefficients and short-run dynamics are estimated after selecting the rank. Johansen's procedure can be preferred to the residual-based approach because it provides a simple way of testing for the cointegration rank and of making inference on the parameters of complex cointegrated systems.

However, the assumption that deviations from equilibrium are integrated of order zero is far too restrictive. In a general setup it is possible to permit errors with a fractional degree of integration. This is an important generalization, since fractional cointegration has the same economic implications as when the processes are integer-valued cointegrated, in the sense that there exist long-run equilibria among the variables. The only difference is that the rate of convergence to the equilibrium is slower in the fractional than in the standard case. Moreover since an $I(1)/I(0)$ cointegration setup ignores the fractional cointegration parameter, a fractionally integrated equilibrium error results in a misspecified likelihood function, which may imply an important loss of power for fractional cointegration testing.

There is a growing literature on fractional cointegration. A first group of contributions deals with estimation of the cointegrating coefficients in regression models; e.g. Marinucci (2000) and Marinucci and Robinson (2001) study least squares and narrow band frequency domain least squares estimates of cointegrating vector. Davidson (2002) considers methods for testing the existence of cointegrating relationships using parametric bootstrap. Davidson (2006) compares bootstrap tests for different residual-based statistics using alternative bias reduction techniques. Gil-Alaña (2003, 2004) proposes a two-step testing procedure of fractional cointegration in macroeconomic time series, based on Robinson's (1994) test. Velasco (2003) considers semiparametric consistent estimation of the memory parameters of a nonstationary fractionally cointegrated vector time series. Marmol and Velasco (2004) propose tests of the null of cointegration, without information on the degree of integration, based on Wald statistics for OLS coefficients. Hualde and Velasco (2008) employ GLS-type of estimator as in Robinson and Hualde (2003) and obtain a chi-squared distribution for the Wald test under the null of no cointegration.

Other authors have considered (Gaussian) maximum likelihood (ML) techniques. Dueker and Startz (1998) illustrate a cointegration testing methodology based on joint estimates of the fractional orders of integration of a cointegrating vector and its parent series. Breitung and Hassler (2002) propose a variant of efficient score test, which allows us to determine the cointegration rank of possibly fractionally integrated series, where the error correction terms may be fractionally integrated as well. Nielsen (2005) proposes a Lagrange Multiplier (LM) test of the null hypothesis of cointegration assuming that the (possibly) fractional order of integration of the observables and the errors are known a priori. An LM test against fractional alternatives requiring the knowledge of the integration orders of observables has been also proposed by Breitung and Hassler (2006). Semiparametric methods have been used as well to design tests for the cointegration rank in fractionally integrated systems, e.g. Robinson and Yajima (2002), Chen and Hurvich (2006).

Gonzalo and Lee (1998) have found that likelihood ratio (LR) tests based on the standard VECM models find spurious cointegration between independent variables that are not unit root processes. Andersson and Gredenhoff (1999) have shown that the likelihood ratio test of no cointegration has power against fractional alternatives, so using standard likelihood based approach we are likely to find the evidence of $C(1,1)$ cointegration when in reality we have fractional cointegration. At the same time the standard ML approach on fractional cointegrated systems gives severe bias and large mean square errors for the "impact" matrix Π. So it is much more severe to ignore fractional cointegration than to incorporate it when it is not present.

Lyhagen (1998), on the basis of a fractional ECM, has allowed errors to be fractionally integrated and has found the asymptotic distribution of the trace test when the fractional degree of cointegration is assumed to be known. He also has simulated bias and mean square error of the estimators of cointegrating vector and adjustment coefficients vector when the cointegration rank is assumed to be one. However the assumption that the order of cointegration is known is very restrictive and may have unexpected effects on the power of the test in case of misspecification, so this restriction should be relaxed.

We examine the asymptotic distributions of the trace test and maximum eigenvalue test under the null hypothesis of no cointegration, when the order of cointegration is not known and is estimated by maximum likelihood under the alternative. We allow deviations from equilibrium to be mean reverting with order of integration possibly greater than zero. The standard cointegration case is also included in our setup. We find that our tests have more power than the standard procedure when cointegration is fractional, while in the case of $C(1,1)$ both procedures have essentially the same power.

The rest of the paper is organized as follows: Section 2 describes the fractional cointegration framework. Section 3 presents the model considered in the paper and the procedure that allows us to define LR tests for no fractional cointegration that are in fact sup tests, which is demonstrated in this paper. The asymptotic distribution of the sup trace and sup maximum eigenvalue tests is presented in Section 4. Section 5 presents Monte Carlo analysis. We tabulate asymptotic distribution and investigate small sample properties of sup tests. In Section 6 we consider a model that allows the original variables to have an unknown level of persistence. Section 7 concludes. Appendix A contains the proof of Theorem 1 that demonstrates asymptotic distributions of sup tests. Appendix B contains the proof that asymptotic results do not change if we consider the general model for levels integrated of any order and possibly unknown, which is discussed in Section 6.

2 Framework description

First let us define the fractionally integrated process $I(\delta)$, following Marinucci and Robinson (2001).

Definition 1 *We say that a scalar process a_t, $t \in Z$, is an $I(\delta)$ process, $\delta > 0$, if there exists a zero mean scalar process η_t, $t \in Z$, with positive and bounded spectral density at zero, such that*

$$a_t = \Delta^{-\delta} \eta_t 1_{(t>0)}, \qquad t \in Z, \qquad \delta > 0, \tag{1}$$

where $1_{(\cdot)}$ is the indicator function, $\Delta = 1 - L$, L is the lag operator and the fractional difference filter is defined formally by

$$(1-z)^{\delta} = \frac{1}{\Gamma(-\delta)} \sum_{j=0}^{\infty} \frac{\Gamma(j-\delta)\, z^j}{\Gamma(j+1)}, \tag{2}$$

and $\Gamma(\cdot)$ is gamma function.

The process a_t is said to be asymptotically stationary when $\delta < \frac{1}{2}$, since it is nonstationary only due to the truncation on the right-hand side of (1). The truncation is designed to cater for cases $\delta \geq \frac{1}{2}$, because otherwise the right-hand side of (1) does not converge in mean square and hence a_t is not well defined.

Second let us define cointegration, following Granger (1986).

Definition 2 *A set of $I(\delta)$ variables is said to be cointegrated, or $CI(\delta, d)$, if there exists a linear combination that is $I(\delta - d)$ for $d > 0$.*

In the standard cointegration setup $\delta = d = 1$ and we can use ML techniques as in Johansen (1988, 1991, 1995). However if $\delta \neq 1$ or $d < 1$ we have fractional cointegration, which calls for a generalization of the standard cointegration framework that would encompass also the fractional case.

The fractional cointegration setup that we consider in this paper is an extension of the Error Correction Mechanism (ECM) framework. Johansen (1995) considers the following Vector Error

Correction Model (VECM)

$$\Delta X_t = \Pi X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \Phi D_t + \epsilon_t, \tag{3}$$

where X_t is a vector of $I(1)$ series of order $p \times 1$, ϵ_t is a $p \times 1$ vector of Gaussian error with variance-covariance matrix Ω and $\Pi,\ \Gamma_1, ..., \Gamma_{k-1}, \Phi$ are freely varying parameters. D_t is a matrix containing deterministic terms and other exogenous variables. When X_t is cointegrated we have the reduced rank condition $\Pi = \alpha\beta'$, where the constant matrices α and β are $p \times r$, having rank r, representing the error correction and cointegrating coefficients, respectively. So in case of cointegration we can use the restricted form of the model, i.e.

$$\Delta X_t = \alpha\beta' X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \Phi D_t + \epsilon_t.$$

A first generalization of the VECM to the fractional case has been suggested by Granger (1986). Following Johansen (2008) and with the notation used in this paper it can be presented as

$$A^*(L)\Delta^\delta X_t = \left(1 - \Delta^d\right) \Delta^{\delta-d} \alpha\beta' X_{t-1} + d(L)\epsilon_t, \tag{4}$$

where $A^*(L)$ and $d(L)$ are lag polynomials, ϵ_t is independent identically distributed with zero mean and positive definite covariance matrix Ω. Johansen (2008) shows how this type of model could be derived starting from the following representation

$$\begin{aligned} \xi' \Delta^\delta X_t &= u_{1t}, \\ \beta' \Delta^{\delta-d} X_t &= u_{2t}, \end{aligned} \tag{5}$$

where $u_t = (u_{1t}', u_{2t}')'$ is i.i.d. $(0, \Sigma)$ and ξ is $p \times (p-r)$ so that (ξ, β) has rank p. Model (5) is a special case of model (4) with $A^*(z) = 1$ and with no lagged X_t and parameter restriction $\beta'\alpha = -I_r$.

The formulation (5) allows for modelling and estimating both the cointegrating vector β and "common trends" vector ξ and has also been used by Breitung and Hassler (2002).

To make the model more flexible it is a natural idea to add lag structure. Granger proposed to add lags of $\Delta^\delta X_t$, which leads to model (4), while Johansen (2007, 2008) proposed a model that comes from adding the fractional lag operator $L_d = 1 - (1 - L)^d$, through the lag polynomial $A(L_d)$, to model (5) and has the following form

$$A(L_d)\Delta^\delta X_t = \left(1 - \Delta^d\right) \Delta^{\delta-d} \alpha\beta' X_t + \epsilon_t. \tag{6}$$

The model we consider in this paper contains lag structure as in model (4) proposed by Granger (1986). Note that under the null of no cointegration we can solve for X_t both models (4) and (6). However the LR tests based on the model (4) have a nice property that estimating d, $d \in \mathcal{D}$, by ML leads to sup test statistics. The null asymptotic distributions of these tests do not depend on any nuisance parameters other than the interval $\mathcal{D}$, which in fact is fixed for each value of δ. We first consider the case when $\delta = 1$ and in Section 7 we discuss generalization to the case with any value of δ, which can be also unknown.

3 Model and tests for no cointegration

We consider the following model

$$\Delta X_t = \alpha\beta' \left(\Delta^{1-d} - \Delta\right) X_t + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \epsilon_t \tag{7}$$

where X_t, ϵ_t and the parameters α, β, $\Gamma_1, ..., \Gamma_{k-1}, \Phi$ are as described in (3), Δ^d is the fractional differencing operator defined in (2) and d is interpreted as degree of the fractional cointegration if the system in (7) is cointegrated. Note that the assumption of Gaussianity will be used only to derive the test statistics for different alternative hypotheses, but not to derive the asymptotic properties of the tests.

The procedure described below is a version of the so-called Johansen's procedure, see Johansen (1995), adapted to the fractional VECM. First let's define

$$\begin{aligned} Z_{0t} &= \Delta X_t, \\ Z_{1t}(d) &= \left(\Delta^{1-d} - \Delta\right) X_t \end{aligned} \tag{8}$$

and let Z_{2t} be stacked variables $\Delta X_{t-1}, ..., \Delta X_{t-k+1}$. First we prewhiten the original series Z_{0t} and Z_{1t}, i.e. we regress Z_{0t} and Z_{1t} on Z_{2t} and consider the regression's residuals R_{0t} and R_{1t} instead of Z_{0t} and Z_{1t} respectively. The model expressed in these variables becomes

$$R_{0t} = \alpha\beta' R_{1t}(d) + \epsilon_t, \quad t = 1, ..., T.$$

The log-likelihood function apart from a constant is given by

$$L(\alpha, \beta, \Omega, d) = -\frac{1}{2} T \log|\Omega| - \frac{1}{2} \sum_{t=1}^{T} [R_{0t} - \alpha\beta' R_{1t}(d)]' \Omega^{-1} [R_{0t} - \alpha\beta' R_{1t}(d)].$$

Define as well

$$S_{ij}(d) = T^{-1} \sum_{t=1}^{T} R_{it}(d) R_{jt}(d)' \quad i, j = 0, 1$$

For fixed d and β, the parameters α and Ω are estimated by regressing R_{0t} on $\beta' R_{1t}(d)$. Plugging the estimates into the likelihood we get

$$L_{\max}^{-2/T}(\hat{\alpha}(\beta), \beta, \hat{\Omega}(\beta), d) = L_{\max}^{-2/T}(\beta, d) = |S_{00} - S_{01}(d)\beta(\beta' S_{11}(d)\beta)^{-1}\beta' S_{10}(d)|,$$

and finally the maximum of the likelihood is obtained by solving the following eigenvalue problem

$$\left|\lambda(d) S_{11}(d) - S_{10}(d) S_{00}^{-1} S_{01}(d)\right| = 0 \tag{9}$$

for eigenvalues $\lambda_i(d)$ and eigenvectors $\upsilon_i(d)$, such that

$$\lambda_i(d) S_{11}(d) \upsilon_i(d) = S_{10}(d) S_{00}^{-1} S_{01}(d) \upsilon_i(d),$$

and $\upsilon_j'(d) S_{11}(d) \upsilon_i(d) = 1$ if $i = j$ and 0 otherwise. Note that the eigenvectors diagonalize the matrix $S_{10}(d) S_{00}^{-1} S_{01}(d)$ since

$$\upsilon_j'(d) S_{10}(d) S_{00}^{-1} S_{01}(d) \upsilon_i(d) = \lambda_i(d)$$

if $i = j$ and 0 otherwise. Thus by simultaneously diagonalizing the matrices $S_{11}(d)$ and $S_{10}(d)S_{00}^{-1}S_{01}(d)$ we can estimate the $r-$dimensional cointegrating space as the space spanned by the eigenvectors corresponding to the r largest eigenvalues. With this choice of β

$$L_{\max}^{-2/T}(d) = |S_{00}| \prod_{i=1}^{r} \left(1 - \hat{\lambda}_i(d)\right). \tag{10}$$

The so-called Johansen's procedure consists in performing a sequence of LR tests. The likelihood under each hypothesis is maximized with the assumption imposed that $d = 1$. However in a fractional cointegration framework d is unknown and has to be estimated. What we propose in this paper is to estimate d by maximum likelihood, i.e.

$$\hat{d} = \arg\max_{d \in \mathcal{D}} L_{\max}(d) \tag{11}$$

where $L_{\max}$ is the concentrated likelihood function defined in (10) and $\mathcal{D} = [0.5 + \varepsilon, 1]$ with $\varepsilon > 0$ and small. Note that the maximization range has been chosen is such a way that we allow deviations from equilibrium (cointegrating residuals) to be of all possible orders of integration that would be asymptotically stationary.

As the first step of the so-called Johansen's procedure the null hypothesis of no cointegration needs to be tested. However under the null of no cointegration parameter d is not identified, while under the null of cointegration rank r with $r > 0$, d is identified and can be consistently estimated by (11), which has been proved in Łasak (2006). Thus asymptotic inference under the null hypothesis of no cointegration will be different than asymptotic inference under the null of cointegration rank r with $r > 0$.

In this paper we concentrate on testing the null of no cointegration or in other words we present a testing procedure that allows us to answer the question whether there is fractional cointegration in the system, which is a first step that has to be made in order to generalize the so-called Johansen's procedure to fractional cointegration case. Recall that the null hypothesis being tested in this paper is $\Pi = 0$ or $d = 0$, so under the null hypothesis, X_t is an $I(1)$ not cointegrated $VAR(k-1)$, like in the standard case.

We describe two LR tests that we call sup trace and sup maximum eigenvalue tests. The reason we call our tests that way will become clear once we study the asymptotic distribution of the proposed tests. Note that by means of sup trace test we test the null hypothesis

$$H_0 : rank\ (\Pi) = r_0 = 0$$

against the alternative hypothesis

$$H_1 : rank\ (\Pi) = p$$

using the LR statistic defined by

$$\sup \text{trace} = trace(\hat{d}_p) = -2\ln\left[LR\left(0|p\right)\right] = -T\sum_{i=1}^{p} \ln[1 - \hat{\lambda}_i(\hat{d}_p)], \tag{12}$$

where

$$\hat{d}_p = \arg\max_{d \in \mathcal{D}} L_p(d) = \arg\max_{d \in \mathcal{D}}\ trace(d)$$

and $L_p(d)$ denotes the maximized likelihood under the hypothesis of rank p for a given d.

The sup maximum eigenvalue ($\sup \lambda_{\max}$) statistic we use to test cointegrating rank 0 against rank 1, i.e. to test

$$H_0 : rank\ (\Pi) = r_0 = 0$$

against

$$H_1 : rank\ (\Pi) = 1$$

and the $\sup \lambda_{\max}$ statistic is defined by

$$\text{sup lambda max} = \lambda_{\max}(\hat{d}_1) = -2 \ln \left[LR\left(0|1\right)\right] = -T \ln[1 - \lambda_1(\hat{d}_1)] \tag{13}$$

and

$$\hat{d}_1 = \underset{d \in \mathcal{D}}{\arg\max} L_1(d) = \underset{d \in \mathcal{D}}{\arg\max} \lambda_{\max}(d),$$

where $L_1(d)$ denotes the maximized likelihood under the hypothesis of rank 1 for a given d.

Recall that we cannot hope that $\hat{d}_1$ or $\hat{d}_p$ estimate consistently a nonexisting true value of d and because of that our tests can be interpreted as sup LR tests, in the spirit of Davies (1977) and Hansen (1996).

4 Asymptotic distribution

In this section we derive the asymptotic distribution of the likelihood ratio tests that we have proposed in (12) and (13). We discuss the case with no lags and under the null of no cointegration.

First let's state assumptions about the innovations, necessary to derive the asymptotic distributions of our likelihood ratio tests.

Assumption 1 *ε_t are independent and identically distributed vectors with mean zero, positive definite covariance matrix Ω, and $E||\varepsilon_t||^q < \infty,\ q \geq 4,\ q > 2/\left(2d - 1\right).$*

Note that by law of large numbers under H_0

$$S_{00} \xrightarrow{P} \Omega.$$

Further using the methods of Marinucci and Robinson (2000) we obtain that under Assumption 1

$$T^{0.5-d} Z_{1[T\tau]} \xrightarrow{\omega} W_d(\tau), \qquad \text{for } d > 0.5,$$

where $\xrightarrow{\omega}$ means convergence in the Skorohod J_1 topology of $\mathcal{D}[0,1]$, W_d is a fractional Brownian motion called by Marinucci and Robinson (1999) "Type II" fractional Brownian motion and defined as

$$W_d(\tau) = \int_0^\tau \frac{(\tau - s)^{d-1}}{\Gamma\left(d\right)} dW(s),$$

and $W(s)$ is vector Brownian motion with covariance matrix Ω.
Then by the Continuous Mapping Theorem we have the following convergence for each $d > 0.5$

$$T^{1-2d} S_{11}(d) \xrightarrow{d} \int_0^1 W_d(\tau) W_d(\tau)' d\tau \tag{14}$$

and, as in e.g. Robinson and Hualde (2003), Proposition 3,

$$T^{1-d} S_{10}(d) \xrightarrow{d} \int_0^1 W_d(\tau) dW',$$

where $\xrightarrow{d}$ denotes convergence in distribution.

The product moments $T^{1-2d} S_{11}(d)$, $T^{1-d} S_{10}(d)$ are $O_p(1)$ uniformly in d since we can show weak convergence for $d \in \mathcal{D}$ in the space $C(\mathcal{D})$ of continuous functions in $\mathcal{D}$ (see Proof of Theorem 1 in the Appendix A), S_{00} is also $O_p(1)$, so the roots $\hat{\lambda}_i(d)$ of equation (9) converge to zero like T^{-1} under the null of no cointegration. This implies that

$$-T \sum_{i=1}^{p} \ln[1 - \hat{\lambda}_i(d)] = T \sum_{i=1}^{p} \hat{\lambda}_i(d) + o_p(1).$$

The sum of the eigenvalues can be found as follows

$$\left| \lambda(d) S_{11}(d) - S_{10}(d) S_{00}^{-1} S_{01}(d) \right| = 0$$

that is equivalent to solve the equation

$$\left| \lambda(d) I - S_{11}^{-1}(d) S_{10}(d) S_{00}^{-1} S_{01}(d) \right| = 0,$$

which shows that

$$T \sum_{i=1}^{p} \hat{\lambda}_i(d) = T \; tr\{S_{11}^{-1}(d) S_{10}(d) S_{00}^{-1} S_{01}(d)\}.$$

From the above reasoning we find that for each d the product

$$S_{11}^{-1}(d) S_{10}(d) S_{00}^{-1} S_{01}(d)$$

converges in distribution towards

$$\left(\int_0^1 W_d(\tau) W_d(\tau)' d\tau \right)^{-1} \int_0^1 W_d(\tau) dW' \Omega^{-1} \int_0^1 (dW) W_d(\tau)',$$

which we can write as

$$\Omega^{-1/2} \left[\int_0^1 B_d(\tau) B_d(\tau)' d\tau \right]^{-1} \int_0^1 B_d(\tau) dB' \int_0^1 (dB) B_d(\tau)' \left(\Omega^{1/2} \right)', \tag{15}$$

where $B_d(\tau) = \Omega^{-1/2} W_d(\tau)$ is the standard fractional Brownian motion. Then we can see that asymptotic distribution of trace and maximum eigenvalue for a fixed d are respectively the trace and the greatest eigenvalue of (15), i.e.

$$trace(d) \xrightarrow{d} trace \left[\int_0^1 (dB) B_d(\tau)' \left[\int_0^1 B_d(\tau) B_d(\tau)' d\tau \right]^{-1} \int_0^1 B_d(\tau) (dB)' \right]$$

$$\lambda \max(d) \xrightarrow{d} \lambda_1 \left[\int_0^1 (dB) B_d(\tau)' \left[\int_0^1 B_d(\tau) B_d(\tau)' d\tau \right]^{-1} \int_0^1 B_d(\tau) (dB)' \right].$$

In the case when d is estimated the following theorem applies.

Theorem 1 *When d, $d \in \mathcal{D}$, is estimated by the maximum likelihood principle under the model (7) the asymptotic distributions of trace and maximum eigenvalue statistics are given respectively by*

$$\sup trace = trace(\hat{d}_p) \xrightarrow{d} \sup_{d \in \mathcal{D}} trace\left[\pounds(d)\right],$$

and

$$\sup lambda \max = \lambda \max(\hat{d}_1) \xrightarrow{d} \sup_{d \in \mathcal{D}} \lambda_1 \left[\pounds(d)\right],$$

where $\mathcal{D} = [0.5 + \varepsilon, 1]$ is a compact set, and

$$\pounds(d) = \int_0^1 (dB)\, B_d(\tau)' \left[\int_0^1 B_d(\tau)\, B_d(\tau)'\, d\tau\right]^{-1} \int_0^1 B_d(\tau)\,(dB)',$$

where B is a p-dimensional Brownian motion on the unit interval, $B_d(\tau)$ is the standard fractional Brownian motion.

The proof is given in the Appendix A. Note that the same result would hold for any compact subset $\mathcal{D} \subset (0.5, 1]$. However we are interested exactly in the set $\mathcal{D} = [0.5+\varepsilon, 1]$ with $\varepsilon > 0$ and small, in order to allow the deviations from equilibrium to have all possible orders of integration that would be asymptotically stationary. In the theoretical part we consider d that belongs to a set $\mathcal{D} = [0.5+\varepsilon, 1]$, $\varepsilon > 0$ and small, because for these values we have a proof of a non-degenerate asymptotic distribution of our test statistics. But in Monte Carlo we use $\mathcal{D} = [0.5, 1]$, since with very small ε there is no difference between these two sets in practice. Moreover we have checked by simulation that the limiting distribution given in Theorem 1 does not depend on the choice of ε.

Finally let us consider the behavior of our tests under the alternative. Note that if the null hypothesis is not true and we have fractional cointegration, then one of the eigenvalues in (9) will be positive in the limit (see Avarucci (2007)). Then

$$-2\ln\left[LR\left(0|p\right)\right] \geqslant -T\ \ln\left(1 - \hat{\lambda}_1(\hat{d}_p)\right) \xrightarrow{p} \infty$$

and

$$-2\ln\left[LR\left(0|1\right)\right] = -T\ln\left(1 - \hat{\lambda}_1(\hat{d}_1)\right) \xrightarrow{p} \infty.$$

So the asymptotic power of both tests is 1.

5 Monte Carlo

The asymptotic distribution of the sup trace and sup maximum eigenvalue statistics have been simulated using the approximation of fractional Brownian motion by fractionally integrated series based on i.i.d Gaussian noise of length 1000. To maximize the likelihood function, the MaxSQPF procedure has been used and optimization has been done on the interval $\mathcal{D}$ =[0.5; 1].[1] Quantiles of the simulated (with 100,000 repetitions) asymptotic distribution are given in Tables 1-2. All Monte Carlo simulations have been done using OxMetrics 4.02 (see Doornik and Ooms (2007)).

Table 1. **Quantiles of sup trace test**

	1%	2.5%	5%	10%	50%	90%	95%	97.5%	99%
1	0.0006	0.0035	0.012	0.045	0.87	3.71	4.98	6.28	8.07
2	0.42	0.71	1.07	1.65	5.37	10.92	12.84	14.67	16.90
3	2.64	3.50	4.41	5.77	13.44	21.73	24.30	26.67	29.64
4	7.06	8.68	10.25	12.53	25.81	36.72	39.95	42.86	46.52
5	13.92	16.23	18.63	22.09	42.59	55.88	59.76	63.24	67.49
6	23.38	26.62	29.95	35.11	63.67	78.87	83.19	87.14	91.93
7	35.56	40.02	44.60	52.37	88.81	105.87	110.84	115.32	120.89
8	50.74	56.53	63.18	77.58	117.84	136.83	142.47	147.61	153.80
9	69.15	76.69	87.20	124.92	150.68	171.61	177.81	183.49	190.27
10	91.42	102.65	128.49	163.12	187.27	210.32	217.29	223.54	230.94

Table 2. **Quantiles of sup lambdamax test**

	1%	2.5%	5%	10%	50%	90%	95%	97.5%	99%
1	0.0006	0.0035	0.012	0.045	0.87	3.71	4.98	6.28	8.07
2	0.37	0.62	0.93	1.44	4.73	9.86	11.72	13.45	15.67
3	1.87	2.50	3.16	4.15	9.37	15.85	18.01	19.98	22.53
4	4.16	5.07	6.06	7.47	14.26	21.81	24.27	26.49	29.38
5	6.85	8.08	9.35	11.14	19.36	27.72	30.40	32.83	35.91
6	10.00	11.52	13.00	15.11	24.60	33.47	36.28	38.87	41.90
7	13.16	14.92	16.67	19.17	29.89	39.49	42.36	45.06	48.45
8	16.69	18.66	20.64	23.36	35.29	45.29	48.48	51.20	54.62
9	20.41	22.48	24.59	27.83	40.72	51.21	54.35	57.30	60.78
10	24.08	26.30	28.72	32.31	46.19	57.02	60.31	63.43	67.21

[1] We optimize on [0.5; 1] instead of [0.5+ε; 1], since we have checked by simulation that the limiting distribution does not depend on the choice of ε.

To evaluate the finite sample properties of sup tests we have simulated two versions of two equation ($p = 2$) model (7), a basic model with no deterministic terms and no lagged differences (Model A), and a model with first lagged difference (model B). The cointegrating vector β in all cases has been normalized to a unit vector $\beta = [1\ 0]'$, which does not affect the generality of the experiment. The parametric space of α in such case is limited to $a_1 \in (-2, 0]$, $a_2 \in R$.

In our simulation we have considered $a_1 = -1.9, -1.4, -0.9, -0.4, 0$ and $a_2 = 0, 1, 2$, which is sufficient to examine the performance of our tests in the whole parametric space. Note that the case $a_1 = a_2 = 0$ shows the behavior of the tests under the null hypothesis of no cointegration, and $a_1 = 0$ but $a_2 \neq 0$ covers in fact the case of $I(2)$ variables.

Small sample properties of the proposed tests for $d_0 = 0.1, 0.3, 0.6, 0.8, 1$ have been investigated by simulation with $10,000$ repetitions and nominal size of 5%. For model B we have considered $\Gamma_1 = \gamma I_p$ with values of $\gamma = 0.1, 0.5, 0.9$. We have run our simulation experiment for samples of $T = 50, 100, 200, 250$ observations.

In Table 3 we compare percentage of rejections of both sup tests under the null hypothesis of no cointegration for both models A, B and different sample sizes.

Table 3. Percentage of rejections by sup trace and sup maximum eigenvalue tests under the null hypothesis of no cointegration. Nominal size 5%.

sup trace

model	A	B		
$T/\ \gamma$		0.1	0.5	0.9
50	4.9	6.8	8.6	19.4
100	4.9	5.3	6.1	12.5
200	4.8	4.9	5.4	8.3
250	4.7	4.9	5.2	7.8

sup lambdamax

model	A	B		
$T/\ \gamma$		0.1	0.5	0.9
50	5.0	6.9	8.3	18.4
100	4.8	5.1	5.7	12.1
200	4.7	4.9	5.4	8.3
250	4.8	4.9	5.3	7.8

The Monte Carlo simulation shows that in case of model A the size distortions of sup tests are small and close to the nominal size of 5%. We observe that both sup tests are slightly undersized in this case. For model B we observe that both tests are usually oversized. The size distortions increase with γ, but decrease when T increases.

In Tables 4 and 5 we present the percentage of rejections by sup tests for model A and the value of $d_0 = 0.6$ under the alternative hypothesis. These tables illustrate behaviour of percentage of rejection in the whole parametric space of α_1 and α_2. Further we comment on all results obtained in this simulation experiment.

Table 4. Percentage of rejections by sup trace test for model A with $d_0 = 0.6$. Right-bottom cell of each table shows the result under the null. Nominal size 5%.

T=50

α_2/ a_1	-1.9	-1.4	-0.9	-0.4	0
2	100	100	100	100	100
1	100	100	99.8	99.3	99.7
0	100	99.8	85.7	23.4	4.9

T=100

α_2/ a_1	-1.9	-1.4	-0.9	-0.4	0
2	100	100	100	100	100
1	100	100	100	100	100
0	100	100	99.9	58.9	4.9

T=200

α_2/ a_1	-1.9	-1.4	-0.9	-0.4	0
2	100	100	100	100	100
1	100	100	100	100	100
0	100	100	99.9	97.1	4.8

T=250

α_2/ a_1	-1.9	-1.4	-0.9	-0.4	0
2	100	100	100	100	100
1	100	100	100	100	100
0	100	100	99.9	99.5	4.7

Table 5. Percentage of rejections by sup maximum eigenvalue test for model A with $d_0 = 0.6$. Right-bottom cell of each table shows the result under the null. Nominal size 5%.

T=50

α_2/ a_1	-1.9	-1.4	-0.9	-0.4	0
2	100	100	100	100	100
1	100	100	100	99.4	99.8
0	100	99.9	86.9	23.5	5.0

T=100

α_2/ a_1	-1.9	-1.4	-0.9	-0.4	0
2	100	100	100	100	100
1	100	100	100	100	100
0	100	100	100	60.2	4.8

T=200

α_2/ a_1	-1.9	-1.4	-0.9	-0.4	0
2	100	100	100	100	100
1	100	100	100	100	100
0	100	100	100	97.6	4.7

T=250

α_2/ a_1	-1.9	-1.4	-0.9	-0.4	0
2	100	100	100	100	100
1	100	100	100	100	100
0	100	100	100	99.6	4.8

The results of the Monte Carlo experiment shows that the power of the sup trace test and the sup maximum eigenvalue test increases with the value of true order of fractional cointegration d_0 under the alternative hypothesis and with the sample size T. We can also see that the power decreases when Π gets closer to zero, which is not strange since we expect problems in this part of the parameters space. These results hold for all models simulated. We do not observe significant difference in power between sup trace test and sup maximum eigenvalue test for any of the considered models.

We have compared the performance of sup tests to LR tests based on the standard VECM. Sup tests have smaller size distortions and better power to detect cointegration. The difference in power is more significant the smaller d is, which is what we would naturally expect. Note that if $d_0 \neq 1$ and we apply LR tests based on the standard VECM we fit a wrong model to the data, which explains the gain of power using sup tests. For $d_0 = 1$ both procedures perform equally well, which is due to the fact that sup tests are quite powerful and they reach maximum power for $d < 1$ already. Note also that a standard cointegration case is nested in our approach. We have also checked that our tests perform well for the values of d_0 that are not covered by our asymptotic results.

6 Model for a general integration degree

Although in most applications it is assumed that all variables are integrated of an integer order, it is interesting to allow the original series to be integrated of an unknown and possibly fractional order δ, $\delta > 0.5$. Then we may consider a general VAR model for $I(\delta)$ processes, i.e.

$$\Delta^{\delta} X_t = \Pi \left(\Delta^{\delta-d} - \Delta^{\delta}\right) X_t + \sum_{i=1}^{k-1} \Gamma_i \Delta^{\delta} X_{t-i} + \epsilon_t \tag{16}$$

In order to apply sup tests described in this paper we could proceed in the following way. First pre-estimate δ under the null hypothesis of no cointegration by ML or any other method, which provides $\hat{\delta}$, a consistent estimate of δ. Further plug $\hat{\delta}$ into the model (16) and follow the procedure described in Section 3 to define test statistics. In Appendix B we prove that under Assumption 2 asymptotic distributions of sup tests have the same general form as in the basic model (7).

Assumption 2 *We have a pre-estimate $\hat{\delta}$, such that*

$$\hat{\delta} - \delta = O_p\left(T^{-\kappa}\right), \quad \kappa > 0$$

where $|\hat{\delta}| \leq K$ for some finite K,

The estimator of δ ($\hat{\delta}$) can be based on the whole vector X_t or can be obtained using only univariate information, see for example Nielsen (2008), Shimotsu and Phillips (2005) or Robinson and Velasco (2000).

Corollary 2 *Asymptotic distributions of sup trace and sup maximum eigenvalue tests have the same general form as in the basic model (7) and are given by Theorem 1.*

Note that the critical values of our tests depend on the interval $\mathcal{D} = [\delta - 0.5 + \varepsilon;\ \delta]$ of possible values of d, on which we maximize the likelihood. The bounds of interval $\mathcal{D}$ are determined by the fact that we want to allow deviations from equilibrium (cointegrating residuals) to be of all possible orders of integration that would be asymptotically stationary. For practical purposes we can simulate the tables of critical values for each δ. We have checked by Monte Carlo simulation that the critical values would converge to a limit when $\delta \to \infty$ and also if $\varepsilon \to 0$.

7 Conclusions

In this paper we have made a first attempt to generalize standard cointegration methodology based on the Error Correction Mechanism (ECM) framework to fractional cointegration case. We have considered two likelihood ratio tests for absence of cointegration against the alternative hypotheses of fractional cointegration. These tests are more general than other tests considered previously in the literature because of three aspects: original variables can be allowed to have an unknown level of persistence, departures from equilibrium can be fractionally cointegrated and the memory of the errors is estimated, not assumed a priori. A great advantage of the considered tests is that they are simple and natural extensions of existing ones, so they can be easily used by practitioners. By means of Monte Carlo simulation we have demonstrated that proposed tests have very good power to detect cointegration, while size distortions are small. There are many extensions to the setup considered in this paper to be developed in the nearest future. We would like to propose a testing procedure for higher ranks, preferably allowing different cointegrating relations to have different memory. We are also planning to consider the estimation of d and linear parameters in the fractional ECM and the analysis of their asymptotic properties.

8 Appendix A

Proof. **(of Theorem 1).** We provide here the proof for the model with no lagged differences and no deterministic terms. For the general version of the model (7) the proof follows as in Johansen (1995, 2007) after prewhitening original variables Z_{0t} and Z_{1t} on Z_{2t} and considering the regressions residuals R_{0t} and R_{1t} instead of Z_{0t} and Z_{1t} respectively. This step has a negligible effect on the asymptotic distribution of tests because of nonstationarity of $Z_{1t}(d)$ for $d > 0.5$.

First note that for each d we have

$$trace(d) \xrightarrow{d} trace\{\int_0^1 (dB)\, B_d' \left[\int_0^1 B_d B_d' du\right]^{-1} \int_0^1 B_d\, (dB)'\},$$

where $\xrightarrow{d}$ denotes usual standard convergence in distribution, which follows because of the joint convergence of the matrices of sample moments to the corresponding stochastic integrals. Then by the same argument we have convergence for finitely many $d's$.

Second recall that $trace(d)$ is a continuous function in all elements of the matrices involved and the random processes on the right hand side are continuous in d. Then if we check that the process is tight in $d \in \mathcal{D}$, we have that

$$trace(d) \Longrightarrow trace\{\int_0^1 (dB)\, B_d' \left[\int_0^1 B_d B_d' du\right]^{-1} \int_0^1 B_d\, (dB)'\},$$

where $\Longrightarrow$ denotes weak convergence for $d \in \mathcal{D}$ in the space $C(\mathcal{D})$ of continuous functions on $\mathcal{D}$, with the supremum norm.

Third since sup function is well defined and continuous on $C(\mathcal{D})$ and

$$trace(\hat{d_p}) = trace(\arg\max_{d\epsilon\mathcal{D}} trace(d)) = \sup_{d\epsilon\mathcal{D}} trace(d),$$

we get by the Continuous Mapping Theorem that the asymptotic distribution of $\sup_{d\epsilon\mathcal{D}} trace(d)$ is the distribution of the

$$\sup_{d\epsilon\mathcal{D}}\left(trace\{\int_0^1 (dB)\, B_d' \left[\int_0^1 B_d B_d' du\right]^{-1} \int_0^1 B_d\,(dB)'\}\right).$$

So to prove that Theorem 1 holds, it is enough to demonstrate that the elements of the sample moments matrices ($S_{11}(d)$ and $S_{10}(d)$) are tight in d, since $trace(d)$ is a continuous function in all elements of the matrices involved as we stated before. Note that S_{00} does not depend on d and $S_{01}(d) = S_{10}'(d)$. We now give the proof for a typical element of S_{11}. The tightness of $S_{10}(d)$ follows by the same arguments.

Recall that in our case with no lags,

$$S_{11}(d) = T^{-1}\sum_{t=1}^{T} Z_{1t}(d)Z_{1t}(d)',$$

$$Z_{1t}(d) = \sum_{j=1}^{t} \pi_j(d)\,\varepsilon_{t-j}.$$

Since (14) and that we can proceed componentwise, then for tightness, by Billingsley's (1968) Theorem 12.3, it is sufficient to check that

$$E\left|T^{1-2d_a} S_{11}^{v,z}(d_a) - T^{1-2d_b} S_{11}^{v,z}(d_b)\right|^m \le K\left|d_a - d_b\right|^\gamma, \tag{17}$$

for some $m > 0$, $K < \infty$ and $\gamma > 1$, where $S_{11}^{v,z}(d)$ is the (v,z) element of $S_{11}(d)$, K and γ are generic constants that do not depend on T nor on (d_a, d_b). We will demonstrate that (17) holds for $m = 2$ and $\gamma = 2$. Then

$$\begin{aligned} S_{11}^{v,z}(d) &= T^{-1}\sum_{t=1}^{T} Z_{1t}^{v}(d)\, Z_{1t}^{z}(d) \\ &= T^{-1}\sum_{t=1}^{T}\left(\sum_{j=1}^{t}\pi_j(d)\,\varepsilon_{t-j}^{v}\right)\left(\sum_{i=1}^{t}\pi_i(d)\,\varepsilon_{t-i}^{z}\right) \end{aligned}$$

so $E\left|T^{1-2d_a} S_{11}^{v,z}(d_a) - T^{1-2d_b} S_{11}^{v,z}(d_b)\right|^2$ is equal to

$$\sum_{t=1}^{T}\sum_{t'=1}^{T} E\left\{A_t(d_a) - A_t(d_b)\right\}\left\{A_{t'}(d_a) - A_{t'}(d_b)\right\}$$

where

$$A_t(d) = A_t^{v,z}(d) = T^{-2d}\left(\sum_{j=1}^{t}\pi_j(d)\,\varepsilon_{t-j}^{v}\right)\left(\sum_{i=1}^{t}\pi_i(d)\,\varepsilon_{t-i}^{z}\right).$$

First let's calculate the contribution of the expectation of the cross product $A_t(d_a)\, A_{t'}(d_b)$, which is

$$\sigma_{vz}^2 T^{-2d_a-2d_b}\sum_{t=1}^{T}\sum_{t'=1}^{T}\sum_{j=1}^{t}\sum_{j'=1}^{t'}\pi_j(d_a)^2\,\pi_{j'}(d_b)^2$$

$$+\sigma_{vv}\sigma_{zz}T^{-2d_a-2d_b}\sum_{t\geq t'}\sum_{j'=1}^{t\wedge t'}\sum_{i'=1}^{t\wedge t'}\pi_{j'}(d_a)\pi_{i'}(d_a)\pi_{t-t'+j'}(d_b)\pi_{t-t'+i'}(d_b)$$

$$+\sigma_{vv}\sigma_{zz}T^{-2d_a-2d_b}\sum_{t<t'}\sum_{j=1}^{t\wedge t'}\sum_{i=1}^{t\wedge t'}\pi_{t'-t+j}(d_a)\pi_{t'-t+i}(d_a)\pi_j(d_b)\pi_i(d_b),$$

$$+\sigma_{vz}^2T^{-2d_a-2d_b}\sum_{t\geq t'}\sum_{j'=1}^{t\wedge t'}\sum_{i'=1}^{t\wedge t'}\pi_{j'}(d_a)\pi_{i'}(d_a)\pi_{t-t'+i'}(d_b)\pi_{t-t'+j'}(d_b)$$

$$+\sigma_{vz}^2T^{-2d_a-2d_b}\sum_{t<t'}\sum_{j=1}^{t\wedge t'}\sum_{i=1}^{t\wedge t'}\pi_{t'-t+i}(d_a)\pi_{t'-t+j}(d_a)\pi_i(d_b)\pi_j(d_b)$$

$$+\kappa_{vzvz}T^{-2d_a-2d_b}\left(\sum_{t\geq t'}\sum_{j'=1}^{t'}\pi_{j'}(d_a)^2\pi_{t-t'+j'}(d_b)^2+\sum_{t'>t}\sum_{j=1}^{t}\pi_{t'-t+j}(d_a)^2\pi_j(d_b)^2\right).$$

Note that once we have evaluated this cross product we can obtain the contribution of the other terms in exactly the same way, for instance that of $A_t(d_a)A_{t'}(d_a)$ by setting $b=a$ in the previous expression, $A_t(d_b)A_{t'}(d_b)$ by setting $a=b$, and $A_t(d_b)A_{t'}(d_a)$ by interchanging a and b.

Combining all cross-products with the appropriate sign and setting $\pi_j^*(d)=\pi_j(d)T^{-d}$, we get that $E\left|T^{1-2d_a}S_{11}^{v,z}(d_a)-T^{1-2d_b}S_{11}^{v,z}(d_b)\right|^2$ is

$$\begin{aligned}&\sigma_{vz}^2\sum_{t,t'}\sum_{j,i=1}^{t\wedge t'}\left\{\pi_j^*(d_a)^2-\pi_j^*(d_b)^2\right\}\left\{\pi_i^*(d_a)^2-\pi_i^*(d_b)^2\right\}^2 \qquad (18)\\
&+\sigma_{vv}\sigma_{zz}\sum_{t,t'}\sum_{j,i=1}^{t\wedge t'}\left\{\pi_j^*(d_a)\pi_i^*(d_a)-\pi_j^*(d_b)\pi_i^*(d_b)\right\}\\
&\times\left\{\pi^*_{|t-t'|+j}(d_a)\pi^*_{|t-t'|+i}(d_a)-\pi^*_{|t-t'|+j}(d_b)\pi^*_{|t-t'|+i}(d_b)\right\}\\
&+\sigma_{vz}^2\sum_{t,t'}\sum_{j,i=1}^{t\wedge t'}\left\{\pi_j^*(d_a)\pi_i^*(d_a)-\pi_j^*(d_b)\pi_i^*(d_b)\right\}\\
&\times\left\{\pi^*_{|t-t'|+i}(d_a)\pi^*_{|t-t'|+j}(d_a)-\pi^*_{|t-t'|+i}(d_b)\pi^*_{|t-t'|+j}(d_b)\right\}\\
&+\kappa_{vzvz}\sum_{j,i=1}^{t\wedge t'}\left\{\pi_j^*(d_a)^2-\pi_j^*(d_b)^2\right\}\left\{\pi^*_{|t-t'|+j}(d_a)^2-\pi^*_{|t-t'|+j}(d_b)^2\right\}.\end{aligned}$$

From Lemma 3 given below we get that the value of (18) is bounded by $K|d_a-d_b|^2$. Using similar arguments it is possible to demonstrate that the remaining terms can also be bounded by $K|d_a-d_b|^2$ by the monotonicity of $\pi_j^*(d)$ in j. This completes the proof. ■

Lemma 3 *The absolute value of*

$$\sum_{t=1}^{T}\sum_{t'=1}^{T}\sum_{j=1}^{t}\sum_{j'=1}^{t'}\left\{\left(\pi_j^*(d_a)^2-\pi_j^*(d_b)^2\right)\left(\pi_{j'}^*(d_a)^2-\pi_{j'}^*(d_b)^2\right)\right\}$$

is bounded by $K|d_a-d_b|^2$.

Proof. We prove Lemma 3 by applying the Mean Value Theorem to $\pi_j^*(d)$. First, observe that

$$\left|\frac{\Gamma'(j+d)}{\Gamma(j+1)} - \frac{\Gamma(j+d)}{\Gamma(j+1)}\log j\right| \leq Kj^{d-1}, \tag{19}$$

since

$$\begin{aligned}
\frac{\Gamma'(j+d)}{\Gamma(j+1)} - \frac{\Gamma(j+d)}{\Gamma(j+1)}\log j &= \frac{\Gamma'(j+d)}{\Gamma(j+d)}\frac{\Gamma(j+d)}{\Gamma(j+1)} - \frac{\Gamma(j+d)}{\Gamma(j+1)}\log j \\
&= \left\{\frac{\Gamma'(j+d)}{\Gamma(j+d)} - \log j\right\}\frac{\Gamma(j+d)}{\Gamma(j+1)} \\
&= \{\Psi(j+d) - \log j\}\frac{\Gamma(j+d)}{\Gamma(j+1)}
\end{aligned}$$

where $\Gamma(j+d)\Gamma(j+1)^{-1} \sim Kj^{d-1}$ for $j \to \infty$ and $\Psi(z) = (d/dz)\log\Gamma(z)$ is the digamma function, which satisfies

$$\Psi(z) = \log z + \frac{1}{2z} + O\left(z^{-2}\right), \quad z \to \infty,$$

so

$$\begin{aligned}
\Psi(j+d) &= \log(j+d) + O\left(j^{-1}\right) \\
&= \log j + O\left(j^{-1}\right)
\end{aligned}$$

as $j \to \infty$, uniformly for $d \in \mathcal{D}$.

Now consider

$$\begin{aligned}
T^d\left|\frac{\partial}{\partial d}\pi_j^*(d)\right| &= \left|\frac{\partial}{\partial d}\pi_j(d) - \pi_j(d)\log T\right| \\
&= \left|\frac{-\Gamma'(d)\Gamma(j+d) + \Gamma(d)\Gamma'(j+d)}{\Gamma^2(d)\Gamma(j+1)} - \frac{\Gamma(j+d)}{\Gamma(d)\Gamma(j+1)}\log T\right| \\
&= \left|\frac{1}{\Gamma^2(d)\Gamma(j+1)}\{-\Gamma'(d)\Gamma(j+d) + \Gamma(d)\Gamma'(j+d) - \Gamma(d)\Gamma(j+d)\log T\}\right| \\
&\leq \overbrace{\left|\frac{-\Gamma'(d)\Gamma(j+d)}{\Gamma^2(d)\Gamma(j+1)}\right|}^{\sim Kj^{d-1}} \\
&\quad + \frac{1}{\Gamma(d)}\left\{\overbrace{\left|\frac{\Gamma'(j+d)}{\Gamma(j+1)} - \frac{\Gamma(j+d)}{\Gamma(j+1)}\log j\right|}^{\sim Kj^{d-1}} + \overbrace{\left|\frac{\Gamma(j+d)}{\Gamma(j+1)}(\log j - \log T)\right|}^{\sim Kj^{d-1}(\log j - \log T)}\right\} \\
&\leq Kj^{d-1}|\log(j/T)|,
\end{aligned}$$

uniformly for $j = 1, ..., T$ and $d \in \mathcal{D}$, using $\pi_j(d) \sim Kj^{d-1}$ and (19).

Finally we get by the Mean Value Theorem that for some $d^* \in [d_a, d_b]$

$$\sum_{t=1}^{T}\sum_{t'=1}^{T}\sum_{j=1}^{t}\sum_{j'=1}^{t'}\left\{\left(\pi_j^*(-d_a)^2 - \pi_j^*(-d_b)^2\right)\left(\pi_{j'}^*(-d_a)^2 - \pi_{j'}^*(-d_b)^2\right)\right\}$$

$$\leq KT^{-4d^*}|d_a - d_b|^2\left\{\sum_{t=1}^{T}\sum_{j=1}^{t}j^{2d^*-2}|\log(j/T)|\right\}^2$$

$$\leq KT^{-4d^*}|d_a-d_b|^2\left\{\sum_{t=1}^{T}t^{2d^*-1}\underbrace{t^{-1}\sum_{j=1}^{t}(j/t)^{2d^*-2}|\log(j/t)|}_{\sim\int_0^1 x^{2d^*-2}|\log x|dx=c}+\sum_{t=1}^{T}|\log(t/T)|\sum_{j=1}^{t}j^{2d^*-2}\right\}^2$$

$$\leq K|d_a-d_b|^2$$

since $d_b > d_a > 0.5$ (the case $d_a = d_b$ is trivial). ■

9 Appendix B

Proof. Here we prove that Theorem 1 holds for model (16) with a general degree of integration estimated under Assumption (2). Again we only provide the proof for model with no lagged differences and no deterministic terms. For a general version of the model (16) the proof follows by the same argument as in the Proof of Theorem 1.

Consider the simple version of model (16)

$$\Delta^\delta X_t = \Pi(\Delta^{\delta-d}-\Delta^\delta)X_t+\epsilon_t.$$

Define $Z_0(\delta)=\Delta^\delta X_t$, $Z_1(\delta,d)=(\Delta^{\delta-d}-\Delta^\delta)X_t$ and note that the sample moment matrices $S_{ij}(\delta_0,\hat{\delta},d)$ for $i,j=0,1$ are calculated by plugging in the estimator $\hat{\delta}$, when the true value is δ_0, while $S_{ij}(\delta_0,\delta_0,d)$ are calculated using the true value δ_0.

It is clear that for true δ (δ_0) the asymptotic distributions of our tests do not change. For estimated δ ($\hat{\delta}$) it does not change if $S_{ij}(\delta_0,\hat{\delta},d)$ is close to $S_{ij}(\delta_0,\delta_0,d)$ for $i,j=0,1$ as $T\to\infty$, uniformly in $d\in\mathcal{D}$.
Let's first demonstrate that it does hold for a properly normalized S_{11} and $\hat{\delta}$ that satisfies Assumption 2. We follow mainly arguments in Robinson and Hualde (2003) Lemma C.2 and Lemma C.4.

Set $A_{11}(d)=T^{1-2d}\left\{S_{11}(\delta_0,\hat{\delta},d)-S_{11}(\delta_0,\delta_0,d)\right\}$, so

$$\begin{aligned}A_{11}(d) &= T^{-2d}\left[\sum_{t=1}^{T}Z_{1t}(\delta_0,\hat{\delta},d)Z_{1t}(\delta_0,\hat{\delta},d)'-\sum_{t=1}^{T}Z_{1t}(\delta_0,\delta_0,d)Z_{1t}(\delta_0,\delta_0,d)'\right] \qquad (20)\\ &= T^{-2d}\sum_{t=1}^{T}\left\{\left(\sum_{j=1}^{t}\pi_j\left(\delta_0-\hat{\delta}+d\right)\varepsilon_{t-j}\right)\left(\sum_{i=1}^{t}\left[\pi_i\left(\delta_0-\hat{\delta}+d\right)-\pi_i(d)\right]\varepsilon'_{t-i}\right)\right.\\ &\quad\left.+\left(\sum_{j=1}^{t}\left[\pi_j\left(\delta_0-\hat{\delta}+d\right)-\pi_j(d)\right]\varepsilon_{t-j}\right)\left(\sum_{i=1}^{t}\pi_i(d)\varepsilon'_{t-i}\right)\right\}.\end{aligned}$$

Define

$$g(x)\equiv\sum_{j=1}^{t}\pi_j(x)\varepsilon_{t-j}$$

and note that r-th derivate of g

$$g^{(r)}(x) = \sum_{j=1}^{t} \pi_j^{(r)}(x)\, \varepsilon_{t-j}$$

Using Taylor expansion of order R, to be chosen later, for $g\left(\delta_0 - \hat{\delta} + d\right)$ we can

$$\begin{aligned} g\left(\delta_0 - \hat{\delta} + d\right) &= g(d) + \sum_{r=1}^{R-1} \frac{\left(\hat{\delta} - \delta_0\right)^r}{r!} g^{(r)}(d) + \frac{\left(\hat{\delta} - \delta_0\right)^R}{R!} g^{(r)}(d^*) \\ &= \sum_{j=1}^{t} \left\{ \pi_j(d) + \sum_{r=1}^{R-1} \frac{\left(\hat{\delta} - \delta_0\right)^r}{r!} \pi_j^{(r)}(d) + \frac{\left(\hat{\delta} - \delta_0\right)^R}{R!} \pi_j^{(R)}(d^*) \right\} \varepsilon_{t-j}, \end{aligned}$$

where d^* is some intermediate point between d and $d + \delta_0 - \hat{\delta}$, to get

$$A_{11}(d) = T^{-2d} \sum_{t=1}^{T} \sum_{i=1}^{t} \sum_{j=1}^{t} \left[\begin{array}{c} \sum_{r,s=0}^{R-1} \frac{\left(\hat{\delta}-\delta_0\right)^{r+s}}{r!s!} \pi_i^{(r)}(d)\, \pi_j^{(s)}(d) + 2 \sum_{r=0}^{R-1} \frac{\left(\hat{\delta}-\delta_0\right)^{r+R}}{r!R!} \pi_i^{(r)}(d)\, \pi_j^{(R)}(d^*) \\ + \frac{\left(\hat{\delta}-\delta_0\right)^{2R}}{R!^2} \pi_i^{(R)}(d^*)\, \pi_j^{(R)}(d^*) \end{array} \right] \varepsilon_{t-j} \varepsilon_{t-i}',$$

where the first summation requires that $s + r > 0$.
Note that $\pi_j(d) \sim j^{d-1}$, $\pi_j^{(1)}(d) \sim j^{d-1} \log j$, $\pi_j^{(2)}(d) \sim j^{d-1} \log^2 j$ as $j \to \infty$, cf. Delgado and Velasco (2005) and Robinson and Hualde (2003) and that we have in $A_{11}(d)$ terms of the following three types,

$$\begin{aligned} A_1(r,s) &= T^{-2d} \sum_{t=1}^{T} \sum_{i=1}^{t} \sum_{j=1}^{t} \pi_i^{(r)}(d)\, \pi_j^{(s)}(d)\, \varepsilon_{t-j} \varepsilon_{t-i}', \text{ where } r,s = 0,1,2,...,R-1,\ s+r>0 \\ A_2(r) &= T^{-2d} \sum_{t=1}^{T} \sum_{i=1}^{t} \sum_{j=1}^{t} \pi_i^{(r)}(d)\, \pi_j^{(R)}(d^*)\, \varepsilon_{t-j} \varepsilon_{t-i}', \text{ where } r = 0,1,2,...,R-1 \\ A_3 &= T^{-2d} \sum_{t=1}^{T} \sum_{i=1}^{t} \sum_{j=1}^{t} \pi_i^{(R)}(d^*)\, \pi_j^{(R)}(d^*)\, \varepsilon_{t-j} \varepsilon_{t-i}' \end{aligned}$$

where we are interested in bounding $\left(\hat{\delta} - \delta_0\right)^{r+s} A_1(r,s)$, $\left(\hat{\delta} - \delta_0\right)^{R+r} A_2(r)$ and $\left(\hat{\delta} - \delta_0\right)^{2R} A_3$ uniformly in $d \in \mathcal{D}$.

Let us work first with the term $A_1 = A_1(r,s)$. Recall that $A = O_p\left(E(A^2)^{\frac{1}{2}}\right)$ and apply the expectation to a typical element (v,z) of A_1,

$$E\left(T^{4d} A_1^2\right) = \sum_{t=1}^{T} \sum_{t'=1}^{T} \left[\sum_{i=1}^{t} \sum_{j=1}^{t} \pi_i^{(r)}(d)\, \pi_j^{(s)}(d)\, \varepsilon_{t-j} \varepsilon_{t-i} \sum_{i'=1}^{t'} \sum_{j'=1}^{t'} \pi_{i'}^{(r)}(d)\, \pi_{j'}^{(s)}(d)\, \varepsilon_{t'-j'} \varepsilon_{t'-i'} \right]$$

$$\begin{aligned} &= \sigma_{vz}^2 \sum_{t=1}^{T} \sum_{t'=1}^{T} \sum_{i=1}^{t} \sum_{i'=1}^{t'} \pi_i^{(r)}(d)\, \pi_i^{(s)}(d)\, \pi_{i'}^{(r)}(d)\, \pi_{i'}^{(s)}(d) \\ &\quad + \sigma_{vv}\sigma_{zz} \sum_{t=1}^{T} \sum_{t'=1}^{T} \sum_{i=1}^{\min\{t,t'\}} \sum_{j=1}^{\min\{t,t'\}} \pi_i^{(r)}(d)\, \pi_j^{(s)}(d)\, \pi_{i+|t'-t|}^{(r)}(d)\, \pi_{j+|t'-t|}^{(s)}(d) \\ &\quad + \sigma_{vz}^2 \sum_{t=1}^{T} \sum_{t'=1}^{T} \sum_{i=1}^{\min\{t,t'\}} \sum_{j=1}^{\min\{t,t'\}} \pi_i^{(r)}(d)\, \pi_j^{(s)}(d)\, \pi_{j+|t'-t|}^{(r)}(d)\, \pi_{i+|t'-t|}^{(s)}(d) \end{aligned}$$

$$+\kappa_{vzvz}\sum_{t=1}^{T}\sum_{t'=1}^{T}\sum_{i=1}^{\min\{t,t'\}}\pi_i^{(r)}(d)\,\pi_i^{(s)}(d)\,\pi_{i+|t'-t|}^{(r)}(d)\,\pi_{i+|t'-t|}^{(s)}(d),$$

where κ_{vzvz} is the fourth cumulant of ε_t. So finally we obtain that

$$(\log T)^{-2R}A_1(r,s)=O_p\left((\log T)^{r+s-2R}\right)=o_p(1)$$

for $r,s<R$, because

$$\sum_{t=1}^{T}\sum_{t'=1}^{T}\left[\sum_{i=1}^{t}\sum_{i'=1}^{t'}\pi_i^{(r)}(d)\,\pi_i^{(s)}(d)\,\pi_{i'}^{(r)}(d)\,\pi_{i'}^{(s)}(d)\right]=$$

$$\begin{aligned}
&= O\left(\sum_{t=1}^{T}\sum_{t'=1}^{T}\left[\sum_{i=1}^{t}\sum_{i'=1}^{t'}i^{2d-2}\left(i'\right)^{2d-2}(\log T)^{2(r+s)}\right]\right)\\
&= O\left(\sum_{t=1}^{T}\sum_{t'=1}^{T}\left[t^{2d-1}\left(t'\right)^{2d-1}(\log T)^{2(r+s)}\right]\right)\\
&= O\left(T^{4d}(\log T)^{2(r+s)}\right).
\end{aligned}$$

This shows the convergence to zero of the finite dimensional distributions of $(\log T)^{-2R}A_1(r,s)$ for fixed d and each r and s. Using the argument in the proof of Theorem 1, we can show the tightness of $(\log T)^{-2R}A_1(r,s)$ for $d\in\mathcal{D}$, and therefore we conclude that $\sup_{d\in\mathcal{D}}|(\log T)^{-2R}A_1(d)|=o_p(1)$. Then, for $\kappa>0$ and $r+s>0$,

$$\sup_{d\in\mathcal{D}}\left(\hat{\delta}-\delta_0\right)^{r+s}A_1(r,s)=o_p\left(T^{-\kappa(r+s)}(\log T)^{2R}\right)=o_p(1).$$

Let's work now with

$$A_2(r)=T^{-2d}\sum_{t=1}^{T}\sum_{i=1}^{t}\sum_{j=1}^{t}\pi_i^{(r)}(d)\,\pi_j^{(R)}(d^*)\,\varepsilon_{t-j}\varepsilon'_{t-i},\ \text{ where } r=0,1,2,...,R-1$$

Note that by monotonicity of $\pi_i^{(r)}(d)$ in d for $i=1,2,\ldots,$

$$\sup_{t=1,\ldots,T}\sup_{d\in\mathcal{D}}|g_t^{(r)}(d)\,T^{-d}|\le\sum_{i=1}^{T}\sup_{d\in\mathcal{D}}|\pi_i^{(r)}(d)\,T^{-d}||\varepsilon'_{t-i}|\le KT^{-\frac{1}{2}}\sum_{i=1}^{T}i^{-\frac{1}{2}}(\log i)^r|\varepsilon'_{t-i}|=O_p(\log T)^r$$

while for each $\epsilon>0$,

$$\begin{aligned}
\sup_{t=1,\ldots,T}\sup_{d\in\mathcal{D}}|g_t^{(R)}(d^*)\,T^{-d}| &\le \sum_{i=1}^{T}\sup_{d\in\mathcal{D}}|\pi_i^{(R)}(d^*)\,T^{-d}||\varepsilon_{t-i}|\\
&\le K\left(\sum_{i=1}^{T}\left(i^{\epsilon-\frac{1}{2}}T^{-\frac{1}{2}}\right)^2\right)^{1/2}\left(\sum_{i=1}^{T}\varepsilon_{t-i}^2\right)^{1/2}\\
&= O_p\left(\left(T^{2\epsilon-1}\right)^{1/2}T^{1/2}\right)\\
&= O_p(T^{\epsilon}).
\end{aligned}$$

Then, for any $\epsilon>0$,

$$\sup_{d\in\mathcal{D}}A_2(r)=O_P\left(T^{1+\epsilon}\right)$$

and

$$\left(\hat{\delta}-\delta_0\right)^{R+r}\sup_{d\in\mathcal{D}}A_2\left(r\right)=O_P\left(T^{-\kappa(R+r)}T^{1+\epsilon}\right)=o_p\left(1\right)$$

if $R>1/\kappa$ and $\epsilon>0$ small enough.

Let's work with the term A_3. Using that

$$\sup_{t=1,\ldots,T}\sup_{d\in\mathcal{D}}|g_t^{(R)}\left(d^*\right)T^{-d}|=O_P(T^{\epsilon})$$

for any $\epsilon>0$ we obtain that $\sup_{d\in\mathcal{D}}A_3=O_p(T^{1+2\epsilon})$ and

$$\left(\hat{\delta}-\delta_0\right)^{2R}\sup_{d\in\mathcal{D}}A_3=\left(\hat{\delta}-\delta_0\right)^{2R}O_p(T^{1+2\epsilon})=O_p(T^{-2\kappa R}T^{1+2\epsilon})=o_p\left(1\right)$$

if $R>1/\kappa$. Then the proof follows for this choice of R.

For the proof for $S_{10}(d)$, we set $A_{10}\left(d\right)=T^{1-d}\left\{S_{10}(\delta_0,\hat{\delta},d)-S_{10}(\delta_0,\delta_0,d)\right\}$, and proceed in a similar way. However, we now need to consider terms like

$$T^{-d}\sum_{t=1}^{T}\left(\sum_{j=1}^{t}\pi_j\left(\delta_0-\hat{\delta}+d\right)\varepsilon_{t-j}\right)\left(\sum_{i=1}^{t}\pi_i\left(\delta_0-\hat{\delta}\right)\varepsilon_{t-i}'\right),$$

which after Taylor expansion lead us to consider

$$\begin{aligned}
A_1^*\left(r,s\right) &= T^{-2d}\sum_{t=1}^{T}\sum_{i=1}^{t}\sum_{j=1}^{t}\pi_i^{(r)}\left(0\right)\pi_j^{(s)}\left(d\right)\varepsilon_{t-j}\varepsilon_{t-i}',\ \text{where } s=0,1,2,...,R-1,\ r=1,2,...,R-1\\
A_2^*\left(r\right) &= T^{-2d}\sum_{t=1}^{T}\sum_{i=1}^{t}\sum_{j=1}^{t}\pi_i^{(r)}\left(0\right)\pi_j^{(R)}\left(d^*\right)\varepsilon_{t-j}\varepsilon_{t-i}',\ \text{where } r=1,2,...,R-1\\
A_2^{**}\left(s\right) &= T^{-2d}\sum_{t=1}^{T}\sum_{i=1}^{t}\sum_{j=1}^{t}\pi_i^{(s)}\left(d\right)\pi_j^{(R)}\left(d^{**}\right)\varepsilon_{t-j}\varepsilon_{t-i}',\ \text{where } s=0,1,2,...,R-1\\
A_3^* &= T^{-2d}\sum_{t=1}^{T}\sum_{i=1}^{t}\sum_{j=1}^{t}\pi_i^{(R)}\left(d^{**}\right)\pi_j^{(R)}\left(d^*\right)\varepsilon_{t-j}\varepsilon_{t-i}'
\end{aligned}$$

where d^{**} is middle point, so that $|d^{**}|\leq|\delta_0-\hat{\delta}|$. ■

References

[1] Andersson, M. K., Gredenhoff M. P. (1999), On the maximum likelihood cointegration procedure under a fractional equilibrium error, Economics Letters, 65, 143-147.

[2] Avarucci, M. (2007), Three Essays on Fractional Cointegration, Ph.D. Thesis, University of Rome "Tor Vergata".

[3] Banerjee, A., Dolado, J., Galbraith, J. W., Hendry, D.F. (1993), Co-integration, Error Correction and the Econometric Analysis of Non-stationary Data, Oxford University Press.

[4] Billingsley, P. (1968), Convergence of Probability Measures, Wiley, New York.

[5] Breitung, J., Hassler U. (2002), Inference on the Cointegration Rank in Fractionally Integrated Processes, Journal of Econometrics, 110, 167-185.

[6] Breitung, J., Hassler U. (2006), A residual-based LM type test against fractional cointegration, Econometric Theory, 22, 1091-1111.

[7] Chen, W., Hurvich, C. (2006), Semiparametric Estimation of Fractional Cointegrating Subspaces, Annals of Statistics, 34, 2939-2979.

[8] Davidson, J. (2002), A model of fractional cointegration, and tests for cointegration using the bootstrap, Journal of Econometrics, 110, 187-212.

[9] Davidson, J. (2006), Alternative bootstrap procedures for testing cointegration in fractionally integrated processes, Journal of Econometrics, 133, 741-777.

[10] Davies, R.B. (1977), Hypothesis Testing when a Nuisance is Present Only under the Alternative, Biometrika, 64, 247-254.

[11] Delgado, M. A., Velasco, C. (2005), Sign tests for long-memory time series, Journal of Econometrics, 127, 215-251.

[12] Dittmann, I. (2004), Error correction models for fractionally cointegrated time series, Journal of Time Series Analysis, 25, 27-32.

[13] Doornik, J. A. (2002), Object-Oriented Matrix Programming Using Ox, 3rd ed. London, Timberlake Consultants Press and Oxford, www.doornik.com.

[14] Doornik, J.A. and Ooms, M. (2007), Introduction to Ox: An Object-Oriented Matrix Language, London: Timberlake Consultants Press.

[15] Dueker, M., Startz, R. (1998), Maximum-likelihood estimation of fractional cointegration with an application to U.S. and Canadian bond rates. Review of Economics and Statistics, 80, 420-426.

[16] Engle, R.F., Granger C.W.J. (1987), Co-integration and Error Correction: Representation, Estimation and Testing, Econometrica, 55, 251-276.

[17] Gil-Alaña, L. A. (2003), Testing of Fractional Cointegration in Macroeconomic Time Series, Oxford Bulletin of Economics and Statistics, 65, 517-529.

[18] Gil-Alaña, L. A. (2004), A Joint Test of Fractional Integration and Structural Breaks at a Known Period of Time, Journal of Time Series Analysis, 25, 691-700.

[19] Giraitis, L., Leipus, R., Philippe, A. (2006), The test for stationarity versus trends and unit roots for a wide class of dependent errors, Econometric Theory, 22, 989-1029.

[20] Gonzalo, J., Lee, H. T. (1998), Pitfalls in testing for long run relationships, Journal of Econometrics, 86, 129-145.

[21] Granger, C.W.J (1981), Some Properties of Time Series Data and their Use in Econometric Model Specification, Journal of Econometrics, 16, 121-130.

[22] Granger, C.W.J (1986), Developments in the study of cointegrated economic variables, Oxford Bulletin of Economics and Statistics, 48, 213-228.

[23] Hansen, B.E. (1996), Inference when a Nuisance Parameter is not Identified under the Null Hypothesis, Econometrica, 64, 413-430.

[24] Hualde, J. Velasco, C. (2008), Distribution-Free test of fractional cointegration, Econometric Theory, 24, 216-255,.

[25] Jang K., Ogaki M. (2001), User guide for Johansen's Method, Department of Economics, The Ohio State University.

[26] Johansen, S. (1988), Statistical Analysis of Cointegration Vectors, Journal of Economic Dynamics and Control, 12, 231-254.

[27] Johansen, S. (1991), Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models, Econometrica, 59, 1551-1580.

[28] Johansen, S. (1995), Likelihood-based inference in cointegrated Vector Auto-Regressive Models. Oxford University Press, Oxford.

[29] Johansen, S. (2007), Representation of cointegrated autoregressive processes with application to fractional processes, forthcoming in Econometric Reviews.

[30] Johansen, S. (2008), A representation theory for a class of vector autoregressive models for fractional processes, Econometric Theory, 24, 651-676.

[31] Łasak, K. (2006), Maximum likelihood estimation of fractionally cointegrated systems, http://www.econ.au.dk/Creates/seminar%20papers/KL%20paper22.pdf

[32] Lobato, I., Velasco, C. (2006) ,Optimal Fractional Dickey-Fuller tests, Econometrics Journal, 9, 492-510.

[33] Lyhagen J. (1998), Maximum likelihood estimation of the multivariate fractional cointegrating model, Working Paper Series in Economics and Finance, 233, Stockholm School of Economics.

[34] MacKinnon, J. G., Haug, A. A., Michelis, L. (1999), Numerical Distributions Functions of Likelihood Ratio Tests for Cointegration, Journal of Applied Econometrics, 14, 563-577.

[35] Marinucci D. (2000), Spectral Regression For Cointegrated Time Series With Long-Memory Innovations, Journal of Time Series Analysis, 21, 685-705.

[36] Marinucci, D., Robinson, P.M. (1999), Alternative Forms of Fractional Brownian Motion, Journal of Statistical Planning and Inference, 80, 111-122.

[37] Marinucci, D., Robinson, P.M. (2000), Weak Convergence of Multivariate Fractional Processes, Stochastic Processes and their Applications, 86, 103-120.

[38] Marinucci, D., Robinson, P.M. (2001), Semiparametric fractional cointegration analysis, Journal of Econometrics, 105, 225-247.

[39] Marmol, F., Velasco, C. (2004), Consistent testing of cointegrating relationships, Econometrica, 72, 1809-1844.

[40] Nielsen, F. S (2008), Local Polynomial Whittle Estimation Covering Non-Stationary Fractional Processes, CREATES-RP 2008-28.

[41] Nielsen, M. Ø. (2005), Multivariate Lagrange Multiplier Tests for Fractional Integration, Journal of Financial Econometrics, 3, 372-398.

[42] Phillips, P.C.B., Ouliaris, S. (1990), Asymptotic Properties of Residual Based Tests for Cointegration, Econometrica, 58, 165-193.

[43] Robinson, P.M. (1994), Efficient tests of nonstationary hypotheses, Journal of the American Statistical Association 89, 1420-1072.

[44] Robinson, P.M., Hualde, J. (2003), Cointegration in Fractional Systems with Unknown Integration Orders, Econometrica, 71, 1727-1766.

[45] Robinson, P.M., Yajima, Y. (2002), Determination of cointegrating rank in fractional systems, Journal of Econometrics, 106, 217-241.

[46] Robinson, P.M., Marinucci, D. (1998), Semiparametric Frequency Domain Analysis of Fractional Cointegration, in P. M. Robinson, Time Series with Long Memory, Oxford University Press.

[47] Robinson, P.M., Velasco, C. (2000), Whittle Pseudo-Maximum Likelihood Estimation for Nonstationary Time Series, Journal of the American Statistical Association, 95, 1229-1243.

[48] Shimotsu, K., Phillips, P.C.B. (2005), Exact Local Whittle Estimation of Fractional Integration, Annals of Statistics, 33, 1890-1933.

[49] Velasco, C. (2003), Gaussian Semi-parametric Estimation of Fractional Cointegration, Journal of Time Series Analysis, 24, 345-378.