Center for Research in Econometric Analysis of Time Series

CREATES Research Paper 2008-53

Maximum likelihood estimation of fractionally cointegrated systems

Katarzyna Lasak

School of Economics and Management
University of Aarhus
Building 1322, DK-8000 Aarhus C
Denmark

Aarhus School of Business
University of Aarhus

Handelshøjskolen
Aarhus Universitet

UNIVERSITY OF COPENHAGEN

Maximum likelihood estimation of fractionally cointegrated systems

Katarzyna Łasak*

Revised 12 May 2010.

Abstract

In this paper we consider a fractionally cointegrated error correction model and investigate asymptotic properties of the maximum likelihood (ML) estimators of the matrix of the cointegration relations, the degree of fractional cointegration, the matrix of the speed of adjustment to the equilibrium parameters, and the variance-covariance matrix of the error term. We show that by using ML principles to estimate jointly all parameters of the fractionally cointegrated system, consistent estimates are obtained. Their asymptotic distributions are provided. The cointegration matrix is asymptotically mixed normal distributed, while the degree of fracional cointegration and the speed of adjustment to the equilibrium matrix have a joint normal distribution, which proves the intuition that the memory of the cointegrating residuals affects the speed of convergence to the long-run equilibrium, but does not have any influence on the long-run relationship. The rate of convergence of the estimators of the long-run relationships depends on the cointegration degree but it is optimal for the strong cointegration case considered. We also prove that misspecification of the degree of fractional cointegation does not affect the consistency of the estimators of the cointegration relationships, although usual inference rules are not valid. The findings are illustrated for finite samples by Monte Carlo analysis. We also apply the developed methodology in a study of the term structure of interest rates.

Keywords: Error correction model, Gaussian VAR model, Maximum likelihood estimation, Fractional cointegration, Term structure of interest rates

JEL: C13, C32.

*CREATES & School of Economics and Management, Aarhus University, Building 1322, Bartholins Allé 10, DK-8000 Aarhus C, Denmark, email: klasak@creates.au.dk

1 Introduction

Cointegration is thought of as a stationary relation amongst nonstationary variables. It has become a standard tool in econometrics since the seminal paper of Granger (1981). One of the most commonly used procedures in econometric practice, the fully parametric inference on $I(1)/I(0)$ cointegrated systems in the framework of the Vector Error Correction Model (VECM), has been developed by Johansen (1988, 1991, 1995). He suggests a maximum likelihood (ML) procedure based on reduced rank regressions. The methodology consists in identifying the number of cointegration vectors within the Vector AutoRegressive (VAR) model by means of performing a sequence of likelihood ratio (LR) tests. If the variables are cointegrated, after selecting the rank, cointegration vectors and the adjustment coefficients are estimated.

However, the assumption that deviations from equilibrium are integrated of order zero is far too restrictive. In a general set up it is possible to permit errors with 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 exists a long-run equilibrium amongst 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.

It has been studied what happens if we use standard VECM models to make inference in fractionally cointegrated systems. Gonzalo and Lee (1998) found that likelihood ratio tests based on the standard models tend to find spurious cointegration between independent variables that are not unit root processes. Further, Andersson and Gredenhoff (1999) have shown by simulation that trace test of no cointegration based on the standard model has power against fractional alternatives, so using ML techniques we are likely to find the evidence of $C(1,1)$ cointegration when in reality we have fractional cointegration. At the same time the ML approach based on standard models gives the estimates of the "impact" matrix $\Pi = \alpha\beta'$ that are severely biased and have large mean square errors if the variables are fractionally cointegrated. So it can be much more severe to ignore fractional cointegration than to incorporate it when in fact it is not present. Moreover, the fractional framework that we consider nests the standard case.

In Łasak (2010) we have developed an asymptotic theory for LR tests based on the fractional VECM. The procedure that leads to construction of LR tests simultaneously produces ML estimates of all the parameters of the fractional VECM, the fractional cointegration degree, the cointegrating vectors, speed of adjustment to the equilibrium coefficients matrix, short run correlation parameters and the variance-covariance matrix of the error term. Knowledge of the properties of those estimators would allow us to propose more complex and complete analysis of fractionally cointegrated systems in line of the analysis in Johansen (1988, 1991, 1995). Therefore in this paper we examine the properties of ML estimators of the fractional VECM.

The list of other work treating inference problems of cointegrated systems in a fractional context, without pretension of completeness, includes the following papers. Robinson (1994) has established the consistency for frequency domain narrow-band estimates of the fractional cointegrating relationship in the stationary bivariate case. In a nonstationary framework the properties of this estimator have been studied by Marinucci and Robinson (2001) and Robinson and Marinucci (2001, 2003). Robinson and Hualde (2003) considered estimation of the cointegrating relationship using the GLS estimator which is asymptotically mixed normal and leads to a Wald test statistic with a standard χ^2 distribution under the null. Their model assumes "strong cointegration", similarly to the model considered in this paper. The asymptotic properties of the cointegrating vector in the "weak cointegration" case have been established in Hualde and Robinson (2007).

Other works allow for deterministic components whose presence implies a competition between stochastic and deterministic trends as discussed in Marinucci and Robinson (2000). Robinson and Iacone (2005) have developed an asymptotic theory for the cointegrating vector in systems generated by polynomial trends and processes that may be fractionally integrated. Chen and Hurvich (2003a) have derived an asymptotic distribution of a tapered narrow-band least squares estimator of the cointegrating parameter. Finally, Hassler, Marmol and Velasco (2008) have examined bivariate regressions of nonstationary variables dominated by linear time trends.

Cointegration amongst stationary long memory processes is especially of interest in financial applications. Christensen and Nielsen (2006) have found that the asymptotic distribution of narrow band least squares (NBLS) is normal if regressors and errors obey the condition that their collective memory is less than 0.5 and their coherency is zero at the origin. Nielsen and Frederiksen (2007) have shown that if the zero coherence assumption is not satisfied then a bias term appears in the mean of the asymptotic distribution. They have also proposed a fully modified NBLS estimator in the spirit of Phillips and Hansen (1990) that does not have this drawback. Nielsen (2007) has shown consistency of joint local Whittle quasi ML estimators of integration orders of the regressors, errors and the cointegration vector.

Our work is in line with current research on fractional models that has been developed in a fully parametric framework. Johansen (2008, 2009) has found a representation of the solution of Fractional Vector Error Correction Model (FVECM) that we analyze in this paper. In Johansen and Nielsen (2010 a,b) the likelihood analysis of the fractional model has been developed for the univariate and multivariate case, respectively. Franchi (2009) has extended the representation theory of Johansen (2008) for polynomial cofractional relations. Further Rossi and Santucci de Magistris (2009) apply fractional VECM to analyze the joint dynamics of futures and spot volatilities.

In this paper we consider a fractionally cointegrated system and investigate the asymptotic properties of the ML estimators of the cointegration relations, the degree of fractional cointegration, speed of adjustment to the equilibrium parameters and the variance-covariance matrix of the error term. We demonstrate that using ML principles to estimate jointly all the parameters of the fractionally cointegrated model we obtain consistent estimates of all of them with known asymptotic distribution. The cointegration matrix estimate results to be asymptotically mixed normal distributed, while the degree of fractional cointegration and the speed of adjustment to the equilibrium matrix have a joint normal distribution. This proves the intuition that the memory of the cointegrating residuals affects the speed of convergence to the long-run equilibrium, but does not have any influence on the long-run relationship. However, the rate of convergence of the estimators of the long-run relationships depends on the cointegration degree. We also demonstrate that misspecification of the degree of fractional cointegration does not affect the consistency of the cointegration relationships estimators, although usual inference rules are no longer valid.

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 gives us estimates of the fractionally cointegrated systems. In Section 4 we describe the main results regarding joint consistency and the asymptotic distribution of all the estimators of the system. Section 5 discusses a model with short run dynamics. Section 6 presents Monte Carlo simulation. Section 7 includes empirical application. Section 8 concludes. Appendix A contains all the lemmas. In Appendix B and C proofs of main results of this paper are given under different assumptions.

2 A framework description

We use the following definition of fractionally integrated process $I(\delta)$, see Marinucci and Robinson (2001).

Definition 1 *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:

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

where $\pi_j(\delta) = \frac{\Gamma(j-\delta)}{\Gamma(-\delta)\Gamma(j+1)}$ and $\Gamma(\cdot)$ is the 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.

We follow with the definition of cointegration, see for example Nielsen (2010):

Definition 2 *The p-vector time series X_t is cointegrated if $X_t \in I(\delta)$, but there exists a full rank $p \times r$ matrix β such that $\beta' X_t \in I(\delta - d)$ for $d > 0$. The number r is the cointegration rank and the space spanned by the columns of β is the cointegration space.*

In the standard cointegration setup $\delta = d = 1$ and we can use ML techniques as in Johansen (1988, 1991, 1995). However if $d < 1$ we have fractional cointegration, which calls for a generalization of the standard cointegration framework, since inference based on a standard VECM may not be valid.

Johansen (2008) has shown how fractional VECM representations could be derived. Assume that X_t is a $p \times 1$ vector fractionally integrated of order δ and there are r linear combinations β of order $\delta - d$, and

$$\begin{aligned} \xi' \Delta^{\delta} X_t &= u_{1t}, \\ \beta' \Delta^{\delta - d} X_t &= u_{2t}, \end{aligned} \tag{3}$$

where $u_t = (u_{1t}', u_{2t}')'$ is i.i.d. $(0, \Sigma)$, ξ is $p \times (p - r)$ so that (ξ, β) has rank p and $X_t = 0$ for $t \leq 0$. Then using the identity[1]

$$\xi_{\perp}(\beta' \xi_{\perp})^{-1} \beta' + \beta_{\perp}(\xi' \beta_{\perp})^{-1} \xi' = I_p$$

we can show that

$$\begin{aligned} \Delta^{\delta} X_t &= \beta_{\perp}(\xi' \beta_{\perp})^{-1} u_{1t} + \xi_{\perp}(\beta' \xi_{\perp})^{-1} \Delta^{d} u_{2t} \\ &= \beta_{\perp}(\xi' \beta_{\perp})^{-1} u_{1t} + \xi_{\perp}(\beta' \xi_{\perp})^{-1} u_{2t} - \xi_{\perp}(\beta' \xi_{\perp})^{-1} \left(1 - \Delta^{d}\right) u_{2t} \\ &= \left(1 - \Delta^{d}\right) \alpha \beta' \Delta^{\delta - d} X_t + \varepsilon_t \end{aligned} \tag{4}$$

where $\varepsilon_t = \beta_{\perp}(\xi' \beta_{\perp})^{-1} u_{1t} + \xi_{\perp}(\beta' \xi_{\perp})^{-1} u_{2t}$ is i.i.d. Recall that α is a $p \times r$ matrix of speed of adjustment to the equilibrium coefficients, $\alpha = -\xi_{\perp}(\beta' \xi_{\perp})^{-1}$ and α satisfies $\beta' \alpha = -I_r$, r is a cointegration rank. The formulation (3) allows for modelling and estimating both the cointegrating vectors β and "common trends" vectors ξ and has also been used by Breitung and Hassler (2002).

To make a model more flexible it is a natural idea to add a lag structure to the model (4). Granger (1986) has included lags of $\Delta^{\delta} X_t$ and has proposed a model that can be presented as

$$A^*(L) \Delta^{\delta} X_t = \left(1 - \Delta^{d}\right) \Delta^{\delta - d} \alpha \beta' X_{t-1} + d(L) \varepsilon_t, \tag{5}$$

where $A^*(L)$ and $d(L)$ are usual lag polynomials.

Johansen (2008, 2009) has proposed another model that comes from adding the fractional lag operator $L_d = 1 - (1 - L)^d$ to model (4) and has the following form

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

[1] Recall for a $p \times m$ matrix a we define the orthogonal complement $a_{\perp}$ to be a $p \times (p - m)$ matrix of rank $p - m$, for which $a' a_{\perp} = 0$.

An alternative model that allows for short run correlation in both the fractional cointegration relationship and in the levels has been proposed in Avarucci (2007)

$$\Delta^{\delta} X_t = \alpha\beta'(\Delta^{-d} - 1) A(L) \Delta^{\delta} X_t + (I - A(L)) \Delta^{\delta} X_t + \varepsilon_t, \tag{7}$$

with a usual lag polynomial $A(L)$, of order k, that can be also expressed as

$$\Delta^{\delta} X_t = \alpha\beta'(\Delta^{-d} - 1)\Delta^{\delta} X_t + \sum_{j=1}^{k} L^j B_j \left\{(\Delta^{-d} - 1)\Delta^{\delta} X_t\right\} + \sum_{j=1}^{k} L^j A_j \Delta^{\delta} X_t + \varepsilon_t, \tag{8}$$

with the restriction $B_j = -A_j \Pi$. We use this model in Section 5.

3 Model and ML estimation

In this paper as a first natural research step we consider the simplest version of the fractional VECM model without lagged differences, which is obviously a special case of models (5), (6) and (7). Moreover we assume that δ is known and we fix $\delta = 1$ to ease the notation. However, δ can take a value different from 1. We use the VECM representation

$$\Delta X_t = \alpha\beta' \left(\Delta^{1-d} - \Delta\right) X_t + \varepsilon_t \tag{9}$$

together with the representation (3). Note that it implies that we impose the restriction $\beta'\alpha = -I_r$ in the model (9), since only under this condition models (3) and (9) are equivalent. We assume Gaussianity of the errors only to define the likelihood function.

To estimate the parameters of model (9) we follow the procedure described in Johansen (1995), but adjusted for the case of fractional VECM that has been already presented in Łasak (2010). Let us define $Z_{0t} = \Delta X_t,\ Z_{1t}(d) = \left(\Delta^{1-d} - \Delta\right) X_t$. The model expressed in these variables becomes

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

The log-likelihood function apart from a constant for model (9) is given by

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

Define as well

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

where $S_{11}(d) = S_{11}(d, d)$ and note that S_{ij} do not depend on d when $i = j = 0$.
For fixed d and β, parameters α and Ω are estimated by regressing Z_{0t} on $\beta' Z_{1t}(d)$ and

$$\hat{\alpha}(\beta) = S_{01}(d)\beta(\beta' S_{11}(d)\beta)^{-1} \tag{10}$$

while

$$\hat{\Omega}(\beta) = S_{00} - S_{01}(d)\beta(\beta' S_{11}(d)\beta)^{-1}\beta' S_{10}(d) = S_{00} - \hat{\alpha}(\beta)(\beta' S_{11}(d)\beta)\hat{\alpha}(\beta)'. \tag{11}$$

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 eigenvalue problem

$$\left|\lambda(d)S_{11}(d) - S_{10}(d)S_{00}^{-1}S_{01}(d)\right| = 0 \tag{12}$$

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 we can estimate d by maximizing the log-likelihood, i.e.

$$\tilde{d} = \arg\max_{d} \; L_{\max}(d), \tag{13}$$

where

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

Note that we assume that the cointegration rank is known already, or alternatively we can establish it using for example the sequence of the tests considered in Łasak and Velasco (2010).

4 Consistency and asymptotic distribution

First let us make the following assumption.

Assumption 1 *ε_t are independent and identically distributed vectors with mean zero, positive definite covariance matrix Ω, and $E||\varepsilon_t||^q < \infty,\ q \geq 4,\ q > 2/\left(2d_0 - 1\right),\ d_0 > \frac{1}{2}$, where d_0 denotes the true value of d.*

The moment condition on ε_t is needed to obtain weak convergence of partial sums to fractional Brownian motion.

Then we define for $d \in (0.5, 1]$ and omitting the dependence on the true value of d, d_0,

$$\lim_{t\to\infty} Var \begin{bmatrix} Z_{0t} \\ \beta' Z_{1t}(d) \\ \beta' Z_{1t}^{(1)}(d) \end{bmatrix} = \begin{bmatrix} \Sigma_{00} & \Sigma_{0\beta}(d) & \dot{\Sigma}_{0\beta}(d) \\ \Sigma_{\beta 0}(d) & \Sigma_{\beta\beta}(d) & \dot{\Sigma}_{\beta\beta}(d) \\ \dot{\Sigma}_{\beta 0}(d) & \dot{\Sigma}_{\beta\beta}(d) & \ddot{\Sigma}_{\beta\beta}(d) \end{bmatrix} \tag{14}$$

where

$$Z_{1t}^{(1)}(d) := \frac{\partial}{\partial d} Z_{1t}(d).$$

Note that

$$\begin{aligned}
\Sigma_{0\beta}(d_0) &= \alpha \Sigma_{\beta\beta}(d_0) \\
\Sigma_{00} &= \alpha \Sigma_{\beta\beta}(d_0) \alpha' + \Omega
\end{aligned}$$

and using Lemma 7 in Appendix A calculate

$$\begin{aligned}
\Sigma_{\beta\beta}(d_0) &= a_0 \cdot \bar{\Sigma}_{\beta\beta} \\
\dot{\Sigma}_{\beta\beta}(d_0) &= c_0 \cdot \bar{\Sigma}_{\beta\beta} \\
\ddot{\Sigma}_{\beta\beta}(d_0) &= \frac{\pi^2}{6} \cdot \bar{\Sigma}_{\beta\beta}
\end{aligned}$$

where $\bar{\Sigma}_{\beta\beta} = \beta'\Omega\beta$ and $a_0 = \sum_1^\infty \pi_j(d_0)^2$ and $c_0 = -\sum_1^\infty j^{-1}\pi_j(d_0)$.

To derive theoretical results we also use the following normalization of $\hat{\alpha}$ and $\hat{\beta}$, as in Johansen (1995). We choose the coordinate system (β, γ) and expand

$$\hat{\beta} = \beta\bar{\beta}'\hat{\beta} + \bar{\gamma}\gamma'\hat{\beta},$$

where $\bar{\beta} = \beta\left(\beta'\beta\right)^{-1}$etc. and define the estimator

$$\tilde{\beta} = \hat{\beta}\left(\bar{\beta}'\hat{\beta}\right)^{-1} = \beta + \bar{\gamma}\gamma'\tilde{\beta} = \beta + \bar{\gamma}U_T$$

where $U_T = \gamma'\tilde{\beta}$. This way of normalizing is convenient for the analysis, since it has the property that $\tilde{\beta} - \beta$ is contained in the space spanned by γ and hence orthogonal to β. Note that since $\tilde{\beta}$ is just a linear transformation of the columns of $\hat{\beta}$ it also maximizes the likelihood function and hence $\tilde{\beta}$ satisfies the likelihood equations. The normalization depends on β, so for practice it is not so useful, but it is convenient in the analysis. We define $\tilde{\alpha} = \hat{\alpha}\hat{\beta}'\bar{\beta}$ so that $\tilde{\alpha}\tilde{\beta}' = \hat{\alpha}\hat{\beta}'$.

For $d_0 \in \mathcal{D} \subset (0.5, 1]$, where $\mathcal{D}$ is a closed set, we demonstrate following Johansen (1995) that the following theorem holds.

Theorem 1 *The estimators* $\tilde{d}$, $\tilde{\beta} = \hat{\beta}\left(\bar{\beta}'\hat{\beta}\right)^{-1}$, $\tilde{\alpha} = \hat{\alpha}\hat{\beta}'\bar{\beta}$, $\hat{\Omega}$ *are consistent. Moreover* $\tilde{\beta} - \beta = o_P(T^{\frac{1}{2}-d_0})$.

Note that Theorem 1 gives the consistency of all the parameters of the fractional VECM we have proposed to estimate jointly by ML.

Theorem 2 *For any fixed* d, $d \neq d_0$, $d > 0.5$ *so that* $q > 2/(2d-1)$ *in Assumption 1 the estimator* $\tilde{\beta} = \hat{\beta}\left(\bar{\beta}'\hat{\beta}\right)^{-1}$ *remains consistent with a rate* $\tilde{\beta} - \beta = o_P(T^{\frac{1}{2}-d})$, *but* $\tilde{\alpha}$ *and* $\hat{\Omega}$ *are not consistent anymore.*

Theorem 2 tells us that if instead of estimating d we plug in any fixed d, $d > 0.5$ to estimate other parameters of fractional VECM we will still obtain a consistent estimate of β, but not of α and Ω, which might suggest that the bias and large mean square errors of the estimator of the impact matrix $\Pi = \alpha\beta'$ found by Andersson and Gredenhoff (1999) came from the estimation of α rather than β.

Theorem 3 *Under Assumption 1 and for $d_0 \in Int\mathcal{D} \subset (0.5, 1]$ the asymptotic distribution of $\tilde{\beta}$ is mixed Gaussian and given by*

$$T^{d_0} U_T = T^{d_0} \gamma' \left(\tilde{\beta} - \beta\right) \to_d \left[\gamma' C \int_0^1 W_{d_0}(\tau) W_{d_0}(\tau)' \, d\tau C' \gamma\right]^{-1} \gamma' C \int_0^1 W_{d_0}(\tau) \, dV_\alpha',$$

where $W_{d_0}(\tau)$ is p-dimensional standard fractional Brownian motion with parameter $d_0 \in (0.5, 1]$

$$W_{d_0}(\tau) = \Gamma^{-1}(d_0) \int_0^\tau (\tau - z)^{d_0 - 1} \, dW(z),$$

$W_{d_0}(\tau)$ and $dV_\alpha(\tau)$ are independent and $dV_\alpha(\tau) = \left(\alpha'\Omega^{-1}\alpha\right)^{-1} \alpha'\Omega^{-1} dW(\tau)$ with W a Brownian motion with covariance matrix Ω.
The conditional variance of the limit distribution is given by

$$\left[C \int_0^1 W_{d_0}(\tau) W_{d_0}(\tau)' \, d\tau C'\right]^{-1} \otimes \left(\alpha'\Omega^{-1}\alpha\right)^{-1}$$

and $C = \beta_\perp \left(\alpha_\perp' \beta_\perp\right)^{-1} \alpha_\perp'$.

We can observe that the distribution of $\tilde{\beta}$ given by Theorem 3 is similar to the distribution found in Johansen (1995) for $d_0 = 1$ fixed. It is also equal to the distribution that Robinson and Hualde (2003) found for their GLS estimator when $r = 1$. The convergence rate of $\tilde{\beta}$ is optimal, hence $\tilde{\beta} - \beta \in O_P\left(T^{-d_0}\right)$.

We would like to emphasize the fact that the estimator $\tilde{\beta}$ is asymptotically independent of the estimators of $\tilde{\alpha}$ and $\tilde{d}$, which means that estimation of other parameters of the system do not affect the estimate of the long run-relationship.

Note that since the asymptotic distribution of $\tilde{\beta}$ remains mixed normal, we can test for the values of the cointegration vectors using Wald test being χ^2 distributed. Thus, following Johansen (1991) we state Theorem 4.

Theorem 4 *If only one cointegrating vector β is present $(r = 1)$ and we want to test the hypothesis $K'\beta = 0$, then the test statistic $T(K'\tilde{\beta}(\tilde{\beta}S_{11}\tilde{\beta})^{-1}\tilde{\beta}'K)((\hat{\lambda}_1^{-1} - 1)(K'\hat{v}(\hat{v}S_{11}\hat{v})^{-1}\hat{v}'K))^{-1}$ is asymptotically χ^2 with one degree of freedom. Here $\hat{\lambda}_1$ is the maximum eigenvalue and $\tilde{\beta}$ the corresponding eigenvector of the equation (12). The remaining eigenvectors form $\hat{v}$.*

In Section 6 we perform a simulation of the Wald test and check that it has proper size and good power to test the values of the cointegration vector in finite samples.

Theorem 5 *The joint asymptotic distribution of* $\tilde{\alpha}$ *and* $\tilde{d}$ *is given by*

$$\left[\begin{array}{c} T^{\frac{1}{2}}(\tilde{d}-d_0) \\ T^{\frac{1}{2}}vec(\tilde{\alpha}-\alpha) \end{array}\right] \rightarrow_d N\left(0,\Psi\right),$$

where

$$\Psi = \left[\begin{array}{cc} \omega^{-1} & c_0\omega^{-1}vec\left(\alpha\right)' \\ c_0\omega^{-1}vec\left(\alpha\right) & \frac{1}{a_0}\left(\bar{\Sigma}_{\beta\beta}^{-1}\otimes\Omega\right)+\frac{c_0^2}{\omega a_0^2}vec\left(\alpha\right)vec\left(\alpha\right)' \end{array}\right]$$

and

$$\omega = \frac{\pi^2}{6}\left(1-\rho_0^2\right)tr\left(\bar{\Sigma}_{\beta\beta}\alpha'\Omega^{-1}\alpha\right),^2$$

$$\rho_0^2 = \frac{c_0^2}{a_0\pi^2/6}.$$

The asymptotic distribution of $\tilde{\alpha}$ is root-T consistent and we can observe that it is related with the asymptotic distribution of $\tilde{d}$. Therefore, estimation of the degree of the fractional cointegration d affects the speed of the adjustment to the equilibrium coefficients, which agrees with common intuition about the speed of the convergence to the long run equilibrium. The asymptotic variance is the usual result when $d_0 = 1$ is known with the extra multiplicative term a_0 and the contribution from estimation of d equal to $\left(c_0^2/\omega a_0^2\right)vec\left(\alpha\right)vec\left(\alpha\right)'$.

The cointegration degree estimator $\tilde{d}$ is also root-T consistent and has an asymptotic normal distribution. The asymptotic variance includes the factor $\left(1-\rho_0^2\right)^{-1} > 1$ due to estimation of α, the factor $tr\left(\bar{\Sigma}_{\beta\beta}\alpha'\Omega^{-1}\alpha\right)^{-1}$ due to estimation inside the ECM and finally, the factor $\left(\pi^2/6\right)^{-1}$ is the usual asymptotic variance for ML estimators of memory parameters in univariate ARFIMA$(0,d,0)$. Note that $\frac{\pi^2}{6}\left(1-\rho_0^2\right)\bar{\Sigma}_{\beta\beta} = \ddot{\Sigma}_{\beta\beta}\left(d_0\right)-\dot{\Sigma}_{\beta\beta}\left(d_0\right)\Sigma_{\beta\beta}^{-1}\left(d_0\right)\dot{\Sigma}_{\beta\beta}\left(d_0\right)$.

We present proofs of Theorem 1, 2, 3 and 5 in Appendix B. In fact the same conclusions can be obtained using standard results on the existence of a consistent sequence of solutions to stochastic optimization problems, such as Lemma 1 in Andrews and Sun (2004). In Appendix C this is investigated under the assumption that Ω is known and $r = 1$ in order to simplify the presentation.

5 Short run dynamics

The results obtained in the previous section could be extended to a general model allowing for short run dynamics. The likelihood analysis of model (6) has been developed for the univariate case in Johansen and Nielsen (2010a) and for the multivariate case in Johansen and Nielsen (2010b). We show how to extend our results for the model (7)-(8) proposed in Avarucci (2007). His model allows for short run correlation in both the fractional cointegration relationship and

[2] or simpler $\omega = \left(\frac{\pi^2}{6}-\frac{c_0^2}{a_0}\right)tr\left(\bar{\Sigma}_{\beta\beta}\alpha'\Omega^{-1}\alpha\right)$

in the levels. Note that this model can be shown to encompass triangular models used in the literature (cf. Robinson and Hualde (2003)) and has nice representations if the roots of the equation $|A(z)| = 0$ are out of the unit circle, $\delta > d$. Basically, this model implies that there is fractional cointegration amongst the prewhitened series $X_t^{\dagger} = A(L) X_t$. It can also be seen as a multivariate extension of Hualde and Robinson's (2007) bivariate cointegrated model.

The model (7) is nonlinear in Π and $A_1, \ldots, A_k$, but we propose to estimate the unrestricted linear model (8) without imposing $B_j = -A_j \Pi$. Then the estimation procedure runs as in Johansen (1995), but with an initial step to prewhiten the main series $\Delta^{\delta} X_t$ and $\Pi(\Delta^{-d} - 1)\Delta^{\delta} X_t$ on k-lags of both $\{(\Delta^{-d} - 1)\Delta^{\delta} X_t\}$ and $\Delta^{\delta} X_t$ as in equation (8). This estimate is inefficient compared with the ML estimator, but is much simpler to compute and analyze.

Let us maintain the assumption that δ is known and $\delta = 1$ to ease the notation. We are interested in the asymptotic distributions of $\tilde{\beta}$, $\tilde{d}$ and the linear parameter estimates $\left(\tilde{\alpha}, \tilde{A}_1, \ldots, \tilde{A}_k\right)$. If we employ unrestricted estimation, then we could investigate the properties of
$\left(\tilde{\alpha}, \tilde{A}_1, \ldots, \tilde{A}_k, \tilde{B}_1, \ldots, \tilde{B}_k\right)$, though B_j are redundant parameters. We can derive all asymptotic results in a similar way to the case with no lag estimation, but obviously the distributions are affected by lag correction compared to those of Theorem 3.

To derive the asymptotic results we should make appropriate changes in the formulas in Appendix B. For instance replace $\Sigma_{\beta\beta}(d)$ by the limit variance of the residuals of the projection of $(\Delta^{-d} - 1)\Delta\beta' X_t$ on k lags of $\{(\Delta^{-d} - 1)\Delta X_t\}$ and ΔX_t. However, the nice covariance structure in terms of constants a_0 and c_0 need not be kept now. The asymptotic properties of $\tilde{d}$ can be deduced from the expansion (30), where now $\ddot{\Sigma}_{\beta\beta}(d_0)$, $\dot{\Sigma}_{\beta\beta}(d_0)$ and $\Sigma_{\beta\beta}(d_0)$ have to be replaced by the limit variance and covariances of $(\beta' Z_{1t}(d_0), \beta' Z_{1t}^{(1)}(d_0))$, $\Sigma_{\beta\beta}^{+}(d_0)$, etc., when projected on k lags of $\{(\Delta^{-d} - 1)\Delta X_t\}$ and ΔX_t. Then the following theorem holds.

Theorem 6 *In model (7) the estimator $\tilde{\beta}$ has the same properties as in Theorems 1 and 3, and the estimators $\tilde{d}$, $\tilde{\alpha}$, $\tilde{A}_1, \ldots, \tilde{A}_k$ have an asymptotic normal joint distribution.*

The asymptotic distribution of $\tilde{d}$ is

$$T^{1/2}(\tilde{d} - d_0) \to_d N\left(0, \bar{\omega}^{-1}\right)$$

where

$$\bar{\omega} = tr\left\{\left[\ddot{\Sigma}_{\beta\beta}^{+}(d_0) - \dot{\Sigma}_{\beta\beta}^{+}(d_0)\,\Sigma_{\beta\beta}^{+-1}(d_0)\,\dot{\Sigma}_{\beta\beta}^{+}(d_0)\right]\alpha'\Omega^{-1}\alpha\right\}.$$

For example $\Sigma_{\beta\beta}^{+}(d_0)$, can be estimated consistently by

$$\frac{1}{T}\sum_{t=1}^{T} \tilde{\beta}' Z_{1t}^{+}\left(\tilde{d}\right) Z_{1t}^{+\prime}\left(\tilde{d}\right)\tilde{\beta}$$

where $Z_{1t}^{+}\left(\tilde{d}\right)$ are the OLS residuals of projecting $Z_{1t}\left(\tilde{d}\right)$ against k lags of $\left\{(1 - \Delta^{-\tilde{d}})\Delta X_t\right\}$ and ΔX_t, $t = 1, \ldots, T$ and $\tilde{\beta}$ and $\tilde{d}$ are ML estimates of β and d.

For $\tilde{\alpha}$ we could obtain a similar expression to (31), in terms of the projected series, and for $\tilde{A}_j$ a parallel result as in Johansen (1995), Theorem 13.5, but corrected for the d estimation increment as in Theorem 3.

6 Some Monte Carlo evidence

To evaluate the small sample properties of the ML estimators of the cointegrated fractional VECM model we have designed the following Monte Carlo experiment. We have generated the two equation model (see Engle, Granger (1987), Banerjee et al. (1993), p.137 or Lyhagen (1998))

$$x_t + by_t = u_t \tag{15}$$

$$x_t + ay_t = e_t$$

where $\Delta^{1-d_0}u_t = \varepsilon_{1t}$, $\Delta e_t = \varepsilon_{2t}$ and ε_{1t}, ε_{2t} are both independently and standard bivariate normally distributed with expectation zero. d_0 is the true cointegration degree and we have considered $d_0 \in (0.5, 1]$. Note that if $d_0 = 1$ then we are in the special case of Johansen's unit root framework. $\beta = [1\ b]'$ is the cointegrating vector, $\alpha = [1\ -a]'$ is the vector of the speed of the adjustment to the equilibrium coefficients. In all simulations we used the parameters a and b equal to 1 and 2 respectively. Note that model (15) is a special case of the model (3), with $\alpha = \xi\perp$.

All Monte Carlo simulations were done using OxMetrix 6.01 (see Doornik and Ooms (2006) and Doornik (2002)). To maximize the likelihood function we used the MaxSQPF procedure. For all simulations we have made 10,000 iterations. We have calculated bias and standard error of the estimators $\tilde{d}$, $\tilde{\beta}$ and $\tilde{\alpha}$ of the parameters of the system (15) simulated with the values of the true d, $d_0 = 0.55, 0.65, 0.75, 0.85, 0.95, 1$ and sample sizes of $T = 50$, 100, 200 and 500 observations. We report results for $T = 200$ in Table 1 and comment on all of them below.

Table 1. Bias and standard error of estimators $\tilde{d}$, $\tilde{\beta}$, $\tilde{\alpha}$ for $T = 200$ observations

d_0	0.55	0.65	0.75	0.85	0.95	1
bias $\tilde{d}$	0.068	0.060	0.058	0.053	0.049	0.047
std $\tilde{d}$	0.166	0.143	0.128	0.117	0.108	0.104
bias$\tilde{\beta}$	0.003	0.002	0.001	0.000	0.000	0.000
std$\tilde{\beta}$	0.062	0.043	0.030	0.021	0.015	0.012
bias $\tilde{\alpha}_1$	0.002	-0.027	-0.036	-0.028	-0.027	-0.024
std $\tilde{\alpha}_1$	0.785	0.351	0.264	0.231	0.198	0.183
bias $\tilde{\alpha}_2$	0.000	0.028	0.035	0.029	0.027	0.025
std $\tilde{\alpha}_2$	0.703	0.300	0.213	0.180	0.150	0.138

Bias and standard errors of $\tilde{d}$, $\tilde{\beta}$ and $\tilde{\alpha}$ are all decreasing with d_0 and with sample size T. For β we obtain very good estimates already for moderate values of d_0 in larger samples. We can estimate β much better than α even for small values of d_0 where $\tilde{\beta}$ has convergence rate close to $T^{\frac{1}{2}}$.

We have also compared the small sample properties of the estimators $\tilde{\beta}$ and $\tilde{\alpha}$ to the small sample properties of ML estimators $\hat{\beta}^J$ and $\hat{\alpha}^J$ obtained based on the standard VECM with $d = 1$. Estimates $\tilde{\beta}$ and $\hat{\beta}^J$ do not differ significantly, while estimates $\tilde{\alpha}$ have in general smaller bias than estimates $\hat{\alpha}^J$ and bigger standard deviation. The significance of the difference in standard deviation seems to be decreasing with the value of d_0.

Further, we have simulated the size and the power of the Wald test given in Theorem 4. We have used again the system described by (15). To check the size we have tested the true linear restriction $K'\beta = 0$ with $K = [-2, 1]$, while to check the power we have tested the false linear restriction $K'\beta = 0$ with $K = [-3, 1]$. We have compared the performance of the Wald test based on the fractional VECM with Wald test based on standard VECM. In Table 2 we report size of both tests. Power is not reported to save space.

Table 2. Percentage of rejections by Wald test under the null. Nominal size 5%.

estimated $\tilde{d}$

T / d	0.55	0.65	0.75	0.85	0.95	1
50	36	32	30	29	28	27
100	23	21	21	19	19	19
200	15	15	14	13	14	14
500	10	10	10	10	9	10

fixed $d = 1$

T / d	0.55	0.65	0.75	0.85	0.95	1
50	39	33	25	18	12	9
100	42	37	30	20	10	7
200	50	44	34	22	10	6
500	60	54	43	28	11	6

We can see that the Wald test has quite distorted size for all values of d_0 in small samples. However size is getting closer to its nominal value when sample size T increases. The bigger d_0 is the faster we get to the nominal size when the sample size T increases. It is due to the fact that the convergence rate of β is T^{-d_0}, see Theorem 3. We can easily observe that if we base Wald test on standard VECM model in case when we have fractionally cointegrated system then the size distortions are bigger for small values of d_0 and in fact they seem to diverge, while for values of d_0 relatively close to 1 the standard test has less distorted size. Power properties of both versions of the Wald test are comparable and very good.

We have also constructed t-tests, to test for the values of d, based on the results of Theorem 5. Results are presented in Tables 3-5.

Table 3. Percentage of rejections by two-sided t-test under the null. Nominal size 5%.

T / d	0.55	0.65	0.75	0.85	0.95	1
50	28	27	27	26	25	24
100	18	18	18	17	17	17
200	13	12	12	12	12	12
500	9	9	9	9	8	9
1000	8	8	7	7	7	7

Table 4. Percentage of rejections by one-sided (against the alternative $d > d_0$) t-test under the null. Nominal size 5%.

T / d	0.55	0.65	0.75	0.85	0.95	1
50	26	28	28	28	29	28
100	19	20	20	20	21	21
200	15	15	15	15	16	15
500	11	11	11	11	11	11
1000	9	9	9	9	9	9

Table 5. Percentage of rejections by one-sided (against the alternative $d < d_0$) t-test under the null. Nominal size 5%.

T / d	0.55	0.65	0.75	0.85	0.95	1
50	12	9	8	7	6	5
100	8	7	6	5	4	4
200	7	5	5	4	3	4
500	5	5	4	4	3	4
1000	5	4	4	4	4	4

We can observe that t-tests also have quite distorted size in small samples. The size distortions seem to be decreasing with the value of d_0 and the sample size T. One-sided test (against the alternative $d < d_0$) is performing better than other two versions of the test in the sense that it has less distorted size. It seems that the size distortions of t-tests are caused by bias of $\tilde{d}$. We have computed the following feasible estimate of variance ω_T^{-1} of the asymptotic distribution of the estimator $\tilde{d}$, where

$$\omega_T\left(\tilde{d}\right) = \frac{\pi^2}{6}\left(1 - \rho_{0T}^2\left(\tilde{d}\right)\right) tr\left(\tilde{\beta}'\hat{\Omega}\tilde{\beta}\ \tilde{\alpha}'\hat{\Omega}^{-1}\tilde{\alpha}\right),$$

$$\rho_{0T}^2\left(\tilde{d}\right) = \frac{c_{0T}^2\left(\tilde{d}\right)}{a_{0T}\left(\tilde{d}\right)\pi^2/6},\ a_{0T}\left(\tilde{d}\right) = \sum_1^T\left\{\pi_j\left(\tilde{d}\right)\right\}^2, c_{0T}\left(\tilde{d}\right) = -\sum_1^T j^{-1}\pi_j\left(\tilde{d}\right)$$

and have compared average value of the standard deviation obtained throughout the iterations with the corresponding true value of ω calculated for given sample size T and true value of cointegration degree d_0. Results are presented in Tables 6 and 7 and seem to be reasonably close, so the estimated standard deviation should not be the cause of the problem.

Table 6. Asymptotic standard deviation of $\tilde{d}$, $(sqrt(T * \omega))^{-1}$

T / d	0.55	0.65	0.75	0.85	0.95	1
100	0.18	0.16	0.15	0.14	0.13	0.13
200	0.13	0.11	0.10	0.10	0.09	0.09
500	0.08	0.07	0.07	0.06	0.06	0.06

Table 7. Average standard error of $\tilde{d}$, $(sqrt(T * \omega_T(\tilde{d})))^{-1}$ for 10,000 replications

T / d	0.55	0.65	0.75	0.85	0.95	1
100	0.19	0.17	0.15	0.14	0.13	0.13
200	0.13	0.12	0.11	0.10	0.09	0.09
500	0.08	0.07	0.07	0.06	0.06	0.06

Further we have checked how the misspecification of d affects the estimation of $\tilde{\beta}$. Note the result in Theorem 3, that estimator of β remains consistent for any value of $d > 0.5$. We simulated model (15) with d_0 = 0.55, 0.75, 0.95 and sample sizes T = 100, 200, 500 and estimated β using fixed values of d, $d = 0.55$, 0.65, 0.75, 0.85, 0.95. We do not report results to save space, but we observed that the bias and standard deviation of $\tilde{\beta}$ were decreasing with the sample size for each d fixed and that the values of the bias and standard deviation of $\tilde{\beta}$ corresponding to different $d's$ did not seem to differ significantly.

Finally, we have examined what happens if $\tilde{d}$ is restricted to belong to the interval $\mathcal{D}$ =[0.5; 1] in our estimation. Note that all the results in this paper are developed for $d > 0.5$ and we assume that δ is known and in our Monte Carlo experiment we know that it is equal to 1, so we could use this extra piece of information in the estimation. We have checked how imposing such restriction on $\tilde{d}$ affects all the results obtained in this section. We have also drawn estimated densities and histograms of small sample distributions of our estimates of d, β and α for $d_0 = 0.55$, 0.75, 0.95 and sample sizes $T = 50$, 100, 200, 500 to compare their shapes when d is estimated with and without restrictions.

In bigger samples the effect of imposing such restriction disappears, however in small samples we can observe the following consequences. In general, the standard deviation of all estimators is decreased. The effect on the bias is not completely clear, but it seems to be decreased in several cases as well. Estimates of α seem to be the most influenced ones, while the estimate of β is the least influenced. The Wald test seems to have smaller size distortions, while t-tests demonstrate very irregular behavior. We have observed that when we impose restriction $d \in \mathcal{D}$ then in small samples and when d_0 is close to the boundary of $\mathcal{D}$, the distribution of $\tilde{d}$ has a camel-shape. But when we let d to be estimated freely then we get a normal distribution in small samples. The shape of the distribution of β and α_1, α_2 does not seem to be affected by imposing the restriction on d.

7 An empirical example of the term structure of interest rates

To illustrate the usefulness and properties of the described methodology we apply it to the problem of the modelling of the term structure of interest rates. There has been a lot of interest in this issue in the current literature, see for example Chen and Hurvich (2003b) and Nielsen (2010). For comparision purposes we follow quite closely the description of the problem and the analysis done in a recent paper by Iacone (2009).

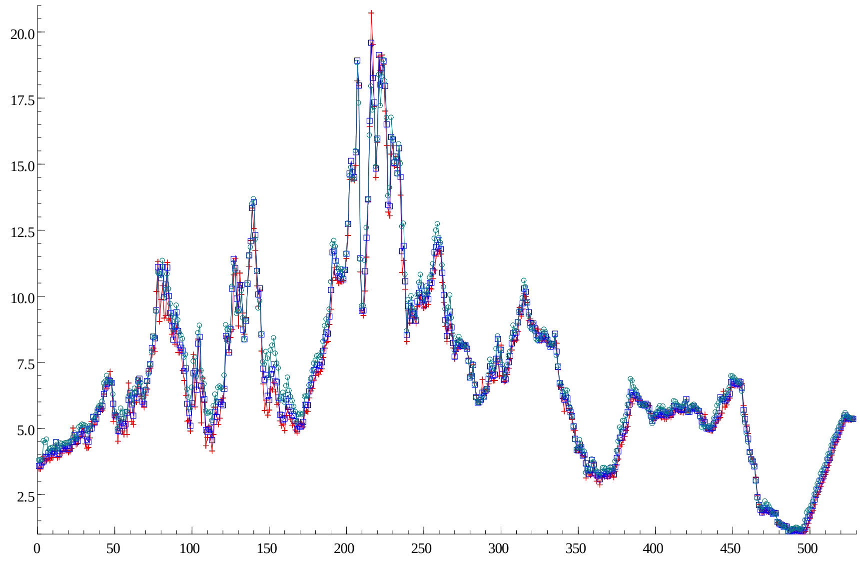
Note that the problem of modelling the behavior of interest rates with different maturities is indeed quite interesting, since as Iacone (2009) argues such a model is necessary both to measure the effects of monetary policy and to price financial assets. Moreover it is an important tool for policy evaluation since the Federal Reserve operates only in one market, the market characterized by contracts with very short maturity, so it is necessary to model the conduction of the monetary policy impulses to the rates of contracts with longer maturities, as a part of a model of the transmission of monetary policy to the final goals. Therefore a VAR model seems to be an ideal setup to model the interactions between interest rates with different maturities in order to take into account the transmission of the monetary policy impulses.

It has been noted by Iacone (2009) as well that modelling the interactions across rates is also important for the economic agents who would like to forecast the effects of future monetary policy decisions on the price of financial assets. Practical example of how to extract the market's expectations on future policy rates from a given term structure and how to use them to price financial instruments was discussed by Soderlind and Svensson (1997).

A theoretical model for the term structure of interest rates was discussed by Fisher (1896) and is known as the "Expectations Hypothesis". It implies that given market efficiency and rational expectations, the interest rates of contracts which only differ in maturity, should be linked by a no-arbitrage relation. Therefore, the return from investing in a contract with maturity over multiple periods should be equivalent to the expected return from investing in multiple consecutive contracts, provided that these span jointly the same time. If the Fisher equation holds, central banks may also find the information in the term structure of interest rates valuable because long term rates include the market's expectations of future inflation. However, there has been a lot of evidence in the literature that the Expectation Hypothesis does not hold and the same conclusion has also been obtained in Iacone (2009).

Iacone (2009) has developed a semiparametric analysis of the US$ interest rates with maturities of 1, 3 and 6 months. The offer rate LIBOR has been used over the period 01/1963-04/2006 of the London interbank deposit, as LIBOR is not affected by any regulation imposed by the central bank, and thus it is a typical measure of the cost of funds in US$. He has found evidence that 3 considered series share the same order of integration, so it is possible to perform a similar analysis in a VAR framework. The order of the integration has been estimated to be 0.88 and the 3 series have been found to cointegrate with cointegration rank 2. However the 3-variate analysis in a VAR framework would impose the assumption that both cointegration relations share the same degree of memory, so we choose to perform the analysis of 3 bivariate systems to avoid the problem. We use the fact that $\hat{\delta}$ equals 0.88. The plot of the raw interest rates series are depicted in Figure 1.

Figure 1. Plot of the raw data



We consider the same data set in order to be able to compare not only the conclusions, but also the values of the parameter estimates. We estimate the basic version of the model presented in section 3, as it seems to be a right choice looking at PACF of the process. We also tested the existence of the break in levels of considered series using the test of Sibbertsen and Kruse (2009) and it indicated no break in the series. So we conclude it is quite reasonable to use the model without short run dynamics. The results of the model estimation are presented in Table 8.

Table 8. Estimation results for the model $\Delta^{\hat{\delta}} X_t = \alpha\beta' \left(\Delta^{\hat{\delta}-d} - \Delta^{\hat{\delta}}\right) X_t + \varepsilon_t$.

model	1,3	1,6	3,6
$\tilde{d}$	0.68	0.59	0.88
$\tilde{d}_s$	0.20	0.29	-
$\tilde{\beta}$	1 -0.98	1 -0.98	1 -0.98
$\tilde{\alpha}$	-0.76 0.19	-0.47 0.16	-0.20 -0.04
trace test	191.5	89.2	29.3
$\lambda_{\max}$ test	190.8	85.8	31.3

The results indicate that we find cointegration between each pair of interest rates. The cointegration vector is very close to $[1, -1]$ in each case and the cointegration residuals are always asymptotically stationary. The estimated order of integration of the spreads $s_t^{(j)} = i_t^{(j)} - i_t^{(1)}$ is denoted by $\tilde{d}_s$. We can see that transmission is slower the longer the distance (in maturity) from the market where FED is present.

Overall we reach the same conclusions and the results we obtained do not differ significantly from Iacone (2009)'s results. We find evidence of cointegration, which is an important result because it means that transmission of impulses along the term structure is still fast enough to let the central bank conduct an active monetary policy. However, the spreads are more persistent than they should be in order for the Expectation Hypothesis to hold.

Possible critics of the above empirical example would be that estimating the equations pairwise defies the multivariate model presented in the paper. However, in order to estimate jointly the 3-variate model with 2 cointegrating relations we would need to impose assumption that both cointegrating relations share the same degree of fractional cointegration, which we considered too restrictive in the given empirical example. Please note that up to date no representation theorem has been proved for any model that would allow for fractional cointegration with cointegrating relations with different memory.

8 Conclusions

In this paper we consider a fractional generalization of the likelihood-based analysis of the cointegrated systems. We describe the estimation procedure based on the reduced rank regressions, which is an adapted to the fractional case version of the procedure proposed in Johansen (1995). We estimate all the parameters of fractionally cointegrated system by ML under the assumption that we know cointegration rank. The novelty of fractional cointegration analysis with respect to standard likelihood-based analysis is the inclusion of an additional parameter, the fractional cointegration degree. The novelty of likelihood-based approach with respect to other fractional methods is that our analysis is fully parametric and allows model based inference.

We investigate asymptotic properties of the ML estimators of a fractional VECM model and we prove that all parameters can be estimated consistently. We show that the asymptotic distribution of the estimator of the cointegration matrix β is independent of other estimates and remains mixed normal as in the standard case, hence we can test for the values of cointegration vectors using Wald test. The asymptotic distributions of the estimators of the speed of the adjustment to the equilibrium coefficients α and cointegration degree d are joint normal, which proves the intuition that the memory of the cointegrating residuals affects the speed of convergence to the long-run equilibrium, but does not have any influence on the long-run relationship. The rate of convergence of the estimators of the long-run relationships depends on the cointegration degree but it is optimal for the strong cointegration case considered.

We also prove that misspecification of the degree of fractional cointegation does not affect the consistency of the estimators of the cointegration relationships, although usual inference rules are not valid. We investigate small sample properties of our estimators by Monte Carlo experiment. We also illustrate applicability of developed methodology by means of real data example, a study of the term structure of interest rates.

There are many possible extensions of the described methodology. First of all we could extend the analysis of short run dynamics and include the estimation of the levels of persistence of the original series. Second it would be very interesting for the empirical analysis to include deterministic terms in the considered models and to allow for structural breaks and different memory of the cointegration relationships. Further we could extend the analysis to the "weak cointegration" case, where the gap between orders of integration of the observables and the cointegration errors is smaller than 1/2, and to the case where original variables are stationary. This would be an interesting extension to be used in financial applications. However, the "strong cointegration case" considered in this paper seems to be the most important case for macroeconomic and econometric practice.

Acknowledgements. *I am grateful to Niels Haldrup, Javier Hualde, Fabrizio Iacone, Søren Johansen, Carlos Velasco and all other people who committed to this piece of research. Solveig Nygaard Sørensen'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 PhD thesis defended at the Department of Economics and Economic History, UAB, Spain.*

9 Appendix A

Lemma 7 *Under the triangular model (3), so that* $\beta'\Delta X_t = \beta'\Delta^{d_0}\varepsilon_t$*, we have*

$$\begin{aligned}\Sigma_{\beta 0}(d) &= \beta'\Omega\beta\alpha' \overbrace{\sum_1^\infty \{\pi_j(d_0) - \pi_j(d_0 - d)\}\pi_j(d_0)}^{=b(d,d_0)} := \bar{\Sigma}_{\beta\beta}\ \alpha' \cdot b(d, d_0) \\ \Sigma_{\beta\beta}(d) &= \beta'\Omega\beta \underbrace{\sum_1^\infty \{\pi_j(d_0) - \pi_j(d_0 - d)\}^2}_{=a(d,d_0)} := \bar{\Sigma}_{\beta\beta} \cdot a(d, d_0)\end{aligned}$$

where $\bar{\Sigma}_{\beta\beta} = \beta'\Omega\beta$*. Denote* $a_0 := a(d_0, d_0) = b(d_0, d_0)$.

Proof. Let us demonstrate the result for $\Sigma_{\beta\beta}(d) = \lim_{t\to\infty} Var(\beta' Z_{1t}(d))$.

$$\begin{aligned}Var(\beta' Z_{1t}(d)) &= E\left\{\frac{1}{T}\sum_{t=1}^T \beta' Z_{1t}(d) Z_{1t}'(d)\beta\right\} = \frac{1}{T}\sum_{t=1}^T E\left\{\left(\Delta^{-d} - 1\right)\beta'\Delta X_t\right\}\{\Delta X_t'\beta\left(\Delta^{-d} - 1\right)\} \\ &= \frac{1}{T}\sum_{t=1}^T E\left\{\left(\Delta^{d_0-d} - \Delta^{d_0}\right)\beta'\varepsilon_t\right\}\{\varepsilon_t'\beta\left(\Delta^{d_0-d} - \Delta^{d_0}\right)\}\end{aligned}$$

which converges to

$$\beta'\Omega\beta\sum_1^\infty \{\pi_j(d_0 - d) - \pi_j(d_0)\}^2.$$

Other elements of (14) could be calculated in a similar way noting for example that

$$\begin{aligned}\beta' Z_{1t}^{(1)}(d) &= \beta'\frac{\partial}{\partial d} Z_{1t}(d) = \frac{\partial}{\partial d}\{\left(\Delta^{-d} - 1\right)\beta'\Delta X_t\} \\ &= \frac{\partial}{\partial d}\{\Delta^{-d} - 1\}\Delta^{d_0}\beta'\varepsilon_t = -\log\Delta(\Delta^{d_0-d}\beta'\varepsilon_t)\end{aligned}$$

■

Define

$$S_{10}^{(i)}(d) = T^{-1}\sum_{t=1}^T \left\{(\partial/\partial d)^i Z_{1t}(d)\right\} Z_{0t}'$$

and

$$S_{11}^{(i,j)}(d_a, d_b) = T^{-1}\sum_{t=1}^T \left\{(\partial/\partial d)^i Z_{1t}(d_a)\right\}\left\{(\partial/\partial d)^j Z_{1t}(d_b)\right\}'.$$

Lemma 8 *Under the triangular model (3), so that $\beta'\Delta X_t = \beta'\Delta^{d_0}\varepsilon_t$, we have that, uniformly in $d \in D \subset (0.5, 1]$,*

$$\begin{array}{rrcll}
(a) & \beta' S_{11}(d)\beta & \to & {}_p\Sigma_{\beta\beta}(d) := a(d,d_0)\bar{\Sigma}_{\beta\beta} & \\
(b) & \beta' S_{1\varepsilon}^{(i)}(d) & = & O_p\left(T^{-1/2}\right), & i = 0,1,2. \\
(c) & T^{1/2-d}\beta' S_{11}^{(i,j)}(d)\bar{\gamma} & \to & {}_p0, & i,j = 0,1,2. \\
(d) & T^{1/2-d}\bar{\gamma}' S_{1\varepsilon}^{(i)}(d) & = & O_p\left(T^{-1/2}\right), & i = 0,1,2.
\end{array}$$

Proof. We first give the proof for (a). We have that $Z_{1t}(d) = \left(\Delta^{-d} - 1\right)\Delta X_t$, so that

$$\beta' Z_{1t}(d) = \left(\Delta^{-d} - 1\right)\beta'\Delta X_t = \left(\Delta^{-d} - 1\right)\beta'\Delta^{d_0}\varepsilon_t = \left(\Delta^{d_0-d} - \Delta^{d_0}\right)\beta'\varepsilon_t = \sum_{j=1}^{t-1}\phi_j(d)\beta'\varepsilon_{t-j},$$

where $\phi_j(d) = \pi_j(d_0 - d) - \pi_j(d_0)$. Then

$$\begin{aligned}
E\left[\beta' S_{11}(d)\beta\right] &= \beta'\Omega\beta\frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{t-1}\left\{\pi_j(d_0-d) - \pi_j(d_0)\right\}^2 \\
&= a(d,d_0)\beta'\Omega\beta + o(1).
\end{aligned}$$

We can write

$$\begin{aligned}
\beta' S_{11}(d)\beta - a(d,d_0)\beta'\Omega\beta &= \beta' S_{11}(d)\beta - E\left[\beta' S_{11}(d)\beta\right] \\
&\quad + E\left[\beta' S_{11}(d)\beta\right] - a(d,d_0)\beta'\Omega\beta,
\end{aligned}$$

where the second line converges uniformly in d to 0, and writing

$$B_T(d) = \beta' S_{11}(d)\beta - E\left[\beta' S_{11}(d)\beta\right]$$

it is easy to show that $B_T(d) = o_p(1)$ for each fixed d. Now we show tightness in d of $B_T(d)$ implying that $\sup_d |B_T(d)| = o_p(1)$. For a typical element of $B_T(d)$ and $d_a, d_b \in D$, we have that

$$\begin{aligned}
& E\left(B_T^{r,s}(d_a) - B_T^{r,s}(d_b)\right)^2 \\
\leq\ & \left(\left[\beta'\Omega\beta\right]^{r,s}\right)^2\left[\frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{t-1}\left(\phi_j(d_a)^2 - \phi_j(d_b)^2\right)\right]^2 \\
& + \frac{1}{T^2}\sum_{t=1}^{T}\sum_{j=1}^{t-1}\sum_{i=1}^{t-1}\sum_{t'=1}^{T}\sum_{j'=1}^{t'-1}\sum_{i'=1}^{t'-1} E\left[u_{t-j}v_{t-i}u_{t'-j'}v_{t'-i'}\right] \\
& \times\left\{\phi_j(d_a)\phi_i(d_a) - \phi_j(d_b)\phi_i(d_b)\right\}\left\{\phi_{j'}(d_a)\phi_{i'}(d_a) - \phi_{j'}(d_b)\phi_{i'}(d_b)\right\}, \qquad (16)
\end{aligned}$$

where $u_t = \left[\beta'\varepsilon_t\right]^r$, $v_t = \left[\beta'\varepsilon_t\right]^s$.

Now note that

$$\begin{aligned}
& \frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{t-1}\left(\phi_j\left(d_a\right)^2-\phi_j\left(d_b\right)^2\right) \\
= \quad & \frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{t-1}\left(\begin{array}{c}\pi_j^2\left(d_0-d_a\right)-\pi_j^2\left(d_0-d_b\right) \\ -2\pi_j\left(d_0\right)\left\{\pi_j\left(d_0-d_a\right)-\pi_j\left(d_0-d_b\right)\right\}\end{array}\right),
\end{aligned}$$

for an intermediate point d^* between d_a and d_b and $\dot{\pi}_j=\left(\partial/\partial x\right)\pi_j\left(x\right)$, is in absolute value no larger than

$$\frac{K}{T}|d_a-d_b|\sum_{t=1}^{T}\sum_{j=1}^{t-1}\left(\begin{array}{c}\dot{\pi}_j\left(d_0-d^*\right)\pi_j\left(d_0-d^*\right) \\ -2\pi_j\left(d_0\right)\dot{\pi}_j\left(d_0-d^*\right)\end{array}\right)\leq K|d_a-d_b|$$

uniformly in T, because $\dot{\pi}_j\left(d_0-d^*\right)\pi_j\left(d_0-d^*\right)$ and $\pi_j\left(d_0\right)\dot{\pi}_j\left(d_0-d^*\right)$ are square summable and can be bounded by $Kj^{-1-\epsilon}$, for some $\epsilon>0$, since $d_0,d_a,d_b\in(0.5,1]$ and therefore $|d_0-d^*|<0.5$.

On the other hand (16) has terms with four typical forms, cf. proof of Theorem 1 in Łasak (2010). The difference with respect to this case is that the weight functions $\phi_j\left(d\right)$ are now square summable for any combination of parameters and they can be bounded by $Kj^{-\eta-1/2}$, for some $\eta>0$, while the differences $|\phi_j\left(d_a\right)-\phi_j\left(d_b\right)|$ can be bounded by $|d_a-d_b|Kj^{-\eta-1/2}$ for some $\eta>0$, uniformly in j. Then the contribution of (16) is of order of magnitude

$$\left(\frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{t-1}|d_a-d_b|\left(Kj^{-\eta-1/2}\right)^2\right)^2\leq K|d_a-d_b|^2,$$

which shows the tightness of B_T and the uniformity of (a).

For the proof of (b) we note that $E\left[\beta' S_{1\varepsilon}^{(i)}\left(d\right)\right]=0$, while the variance of a typical element of $S_{1\varepsilon}^{(0)}\left(d\right)$ is

$$\mathrm{Var}\left[\left\{\beta' S_{1\varepsilon}\left(d\right)\right\}_r\right]\leq\frac{K}{T^2}\sum_t\sum_j\phi_j\left(d\right)^2=O\left(T^{-1}\right),$$

and the uniformity in d for any i can be shown using similar techniques. For terms involving derivatives, note that the asymptotic approximations for the derivatives of $\pi_j\left(\cdot\right)$ for large j are like those for $\pi_j\left(\cdot\right)$ up to logarithmic terms.

For the proof of (c) we note that

$$\bar{\gamma}' Z_{1t}\left(d\right)=\left(\Delta^{-d}-1\right)\bar{\gamma}'\Delta X_t=\left(\Delta^{-d}-1\right)\bar{\gamma}'\varepsilon_t=\sum_{j=1}^{t-1}\pi_j\left(-d\right)\bar{\gamma}'\varepsilon_{t-j},$$

so that for $i,j=0$,

$$T^{1/2-d}\beta' S_{11}\left(d\right)\bar{\gamma}=T^{-1/2-d}\beta'\sum_{t=1}^{T}\sum_{j=1}^{t}\sum_{i=1}^{t}\phi_j\left(d\right)\pi_i\left(-d\right)\varepsilon_{t-j}\varepsilon_{t-i}'\bar{\gamma}.$$

Further note that

$$\begin{aligned}
E\left[T^{1/2-d}\beta' S_{11}(d)\bar{\gamma}\right] &= T^{-1/2-d}\beta' \sum_{t=1}^{T}\sum_{j=1}^{t}\phi_j(d)\pi_j(-d)\Omega\bar{\gamma} \\
&= O\left(T^{-1/2-d}\sum_{t=1}^{T}\sum_{j=1}^{t} j^{-3/2+d-\epsilon}\right) \\
&= O\left(T^{-\epsilon}\right) = o(1)
\end{aligned}$$

for some $\epsilon > 0$, and similarly we can show that for each d, $Var\left[T^{1/2-d}\beta' S_{11}(d)\bar{\gamma}\right] = o(1)$ as $T \to \infty$. Then tightness follows as in the proof of Theorem 1 in Łasak (2010) and thus $\sup_d \left|T^{1/2-d}\beta' S_{11}(d)\bar{\gamma}\right| = o_p(1)$. The argument for other values of i and j is similar.

The proof of (d) follows combining ideas of the proofs of (b) and (c). ■

Lemma 9 *Under the triangular model (3), so that $\beta'\Delta X_t = \beta'\Delta^{d_0}\varepsilon_t$, we have that, uniformly in d such that $|d - d_0| \leq T^{-\kappa}$, for some $\kappa > 0$, and for all $\eta > 0$,*

$$\begin{aligned}
(a) \qquad \beta' S_{11}(d)\beta &\to {}_p\Sigma_{\beta\beta}(d_0) = a_0\bar{\Sigma}_{\beta\beta} \\
\beta' S_{11}^{(1,0)}(d)\beta &\to {}_p\dot{\Sigma}_{\beta\beta}(d_0) = c_0\bar{\Sigma}_{\beta\beta} \\
\beta' S_{11}^{(1,1)}(d)\beta &\to {}_p\ddot{\Sigma}_{\beta\beta}(d_0) = \frac{\pi^2}{6}\bar{\Sigma}_{\beta\beta}
\end{aligned}$$

$$\begin{aligned}
(b) \qquad \beta' S_{1\varepsilon}^{(i)}(d) &= O_p\left(T^{-1/2}\right), \quad i = 0,1,2. \\
(c) \qquad \beta' S_{11}^{(i,j)}(d)\bar{\gamma} &= O_p\left(T^{d_0-1+\eta}\right), \quad i,j = 0,1,2. \\
(d) \qquad \bar{\gamma}' S_{1\varepsilon}^{(i)}(d) &= O_p\left(T^{d_0-1+\eta}\right), \quad i = 0,1,2. \\
(e) \qquad \beta'\left\{S_{1\varepsilon}^{(i)}(d_0) - S_{1\varepsilon}^{(i)}(d)\right\} &= o_p\left(T^{-1/2}\right), \quad i = 0,1. \\
\bar{\gamma}'\left\{S_{1\varepsilon}(d_0) - S_{1\varepsilon}(d)\right\} &= o_p\left(T^{d_0-1}\right). \\
\bar{\gamma}'\left\{S_{11}(d_0) - S_{11}(d)\right\}\bar{\gamma} &= o_p\left(T^{2d_0-1}\right).
\end{aligned}$$

When $d_0 = 1$ we can set $\eta = 0$.

Proof. Omitted, the proofs of $(a) - (b)$ being similar to Lemma 8. For the proof of (e) follow the methods of the proof in Appendix B in Łasak (2010). ■

Lemma 10 *Let the process X_t be given by (3), choose γ orthogonal to β such that (β, γ) has full rank p. Then for any $d \in (0.5, 1]$ as $T \to \infty$*

$$T^{1-d}\bar{\gamma}' S_{1\epsilon}(d) \xrightarrow{d} \bar{\gamma}' C \int_0^1 W_d(\tau)\, dW(\tau)'$$

$$T^{1-2d}\bar{\gamma}' S_{11}(d)\bar{\gamma} \xrightarrow{d} \bar{\gamma}' C \int_0^1 W_d(\tau)\, W_d(\tau)'\, d\tau C'\bar{\gamma}$$

where $C = \beta_\perp \left(\alpha'_\perp \beta_\perp\right)^{-1} \alpha'_\perp$.

Proof. The result follows by similar arguments as in Theorem B.13 of Johansen (1995) and weak convergence follows from Marinucci and Robinson (2000). ■

Lemma 11

$$T^{1/2} tr \left\{ \alpha \left[\dot{\Sigma}_{\beta\beta} (d_0) \Sigma_{\beta\beta}^{-1} (d_0) \beta' S_{1\varepsilon} (d_0) - \beta' S_{1\varepsilon}^{(1)} (d_0) \right] \Omega^{-1} \right\} \to_d N \left(0, \omega^2\right).$$

Proof. Use the martingale Central Limit Theorem and that

$$\begin{aligned}
& \lim_{T\to\infty} Var \left(T^{1/2} tr \left\{ \alpha \left[\dot{\Sigma}_{\beta\beta} (d_0) \Sigma_{\beta\beta}^{-1} (d_0) \beta' S_{1\varepsilon} (d_0) - \beta' S_{1\varepsilon}^{(1)} (d_0) \right] \Omega^{-1} \right\} \right) \\
= & \lim_{T\to\infty} Var \left(\frac{1}{T^{1/2}} \sum_{t=1}^{T} \left[\dot{\Sigma}_{\beta\beta} (d_0) \Sigma_{\beta\beta}^{-1} (d_0) \beta' Z_t (d_0) - \beta' Z_t^{(1)} (d_0) \right]' \alpha \Omega^{-1} \varepsilon_t \right) \\
= & \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} trE \left\{ \begin{array}{c} \left[\dot{\Sigma}_{\beta\beta} (d_0) \Sigma_{\beta\beta}^{-1} (d_0) \beta' Z_t (d_0) - \beta' Z_t^{(1)} (d_0) \right] \\ \times \left[\dot{\Sigma}_{\beta\beta} (d_0) \Sigma_{\beta\beta}^{-1} (d_0) \beta' Z_t (d_0) - \beta' Z_t^{(1)} (d_0) \right]' \alpha \Omega^{-1} \alpha' \end{array} \right\}
\end{aligned}$$

$$\begin{aligned}
&= tr \left[\left\{ \ddot{\Sigma}_{\beta\beta} (d_0) - \dot{\Sigma}_{\beta\beta} (d_0) \Sigma_{\beta\beta}^{-1} (d_0) \dot{\Sigma}_{\beta\beta} (d_0) \right\} \alpha' \Omega^{-1} \alpha \right] \\
&= tr \left[\left\{ \frac{\pi^2}{6} \bar{\Sigma}_{\beta\beta} - \frac{c_0^2}{a_0} \bar{\Sigma}_{\beta\beta} \right\} \alpha' \Omega^{-1} \alpha \right] = \omega.
\end{aligned}$$

■

Lemma 12 *The solutions* $\lambda(d)$ *of*

$$|\lambda(d) \Sigma_{\beta\beta} (d) - \Sigma_{\beta 0} (d) \Sigma_{00}^{-1} \Sigma_{0\beta} (d)| = 0, \tag{17}$$

are maximized with respect to d *when* $d = d_0$.

Proof. Note that equation (17) can be expressed as

$$|\lambda(d) \Sigma_{\beta\beta} (d_0) \frac{a(d, d_0)}{a_0} - \Sigma_{\beta 0} (d_0) \Sigma_{00}^{-1} \Sigma_{0\beta} (d_0) \left[\frac{b(d, d_0)}{a_0} \right]^2 | = 0, \tag{18}$$

or equivalently

$$|\lambda(d) \Sigma_{\beta\beta} (d_0) - \Sigma_{\beta 0} (d_0) \Sigma_{00}^{-1} \Sigma_{0\beta} (d_0) \frac{1}{a_0} \left[\frac{b^2 (d, d_0)}{a(d, d_0)} \right] | = 0. \tag{19}$$

So the solutions $\lambda(d)$ are maximized with respect to d when the scalar factor

$$\begin{aligned}
\frac{b^2(d,d_0)}{a(d,d_0)} &= \frac{\left(\sum_1^\infty \{\pi_j(d_0)-\pi_j(d_0-d)\}\pi_j(d_0)\right)^2}{\sum_1^\infty \{\pi_j(d_0)-\pi_j(d_0-d)\}^2} \\
&= \frac{\left(\sum_1^\infty \{\pi_j(d_0)-\pi_j(\bar{\delta})\}\pi_j(d_0)\right)^2}{\sum_1^\infty \{\pi_j(d_0)-\pi_j(\bar{\delta})\}^2} \\
&: = c(\bar{\delta}) := \frac{b^2(\bar{\delta})}{a(\bar{\delta})}
\end{aligned}$$

is maximized**,** where $\bar{\delta} = d_0 - d$, and we omit the dependence on d_0. Note also that

$$c(\bar{\delta}) = \frac{b^2(d,d_0)}{a(d,d_0)} \leq \sum_1^\infty \pi_j(d_0)^2 \quad \text{for all } \bar{\delta},$$

using Cauchy-Swartz inequality, therefore we conclude that

$$c(\bar{\delta}) < \sum_1^\infty \pi_j(d_0)^2 \quad \text{for all } \ \bar{\delta} \neq 0 \Leftrightarrow d \neq d_0$$

because there will be no perfect correlation amongst the sequences $\pi_j(d_0)$ and $\pi_j(d_0) - \pi_j(\bar{\delta})$ if $\pi_j(\bar{\delta}) \neq 0$, which is the only case were $\pi_j(\bar{\delta})$ is constant for all $j > 0$, $|\bar{\delta}| < 1$ whereas

$$\begin{aligned}
c(0) &= \frac{b^2(d_0,d_0)}{a(d_0,d_0)} = \frac{\left(\sum_1^\infty \{\pi_j(d_0)-\pi_j(0)\}\pi_j(d_0)\right)^2}{\sum_1^\infty \{\pi_j(d_0)-\pi_j(0)\}^2} \\
&= \frac{\left(\sum_1^\infty \pi_j^2(d_0)\right)^2}{\sum_1^\infty \pi_j(d_0)^2} = \sum_1^\infty \pi_j(d_0)^2 \quad \text{for } \bar{\delta} = 0.
\end{aligned}$$

■

Note that results of Lemma 12 imply that the likelihood function is maximized at $d = d_0$ in the limit.

10 Appendix B

Proof. (of Theorem 1) Define the matrix $A_T(d) = \left(\beta, T^{\frac{1}{2}-d}\bar{\gamma}\right)$. By Lemmas 8 and 10, for any value of d, $d > 0.5$ the ordered eigenvalues of

$$|\lambda(d)A_T'(d)S_{11}(d)A_T(d) - A_T'(d)S_{10}(d)S_{00}^{-1}S_{01}(d)A_T(d)| = 0 \tag{20}$$

converge uniformly to those of

$$|\lambda(d)\Sigma_{\beta\beta}(d) - \Sigma_{\beta 0}(d)\Sigma_{00}^{-1}\Sigma_{0\beta}(d)|\ |\lambda(d)\int_0^1 W_d(\tau)W_d(\tau)'\,du| = 0 \tag{21}$$

and the space spanned by the r first eigenvectors of (20) converges to the space spanned by the first unit vectors or equivalently to the space spanned by vectors with zeros in the last

$p-r$ coordinates. The space spanned by the first r eigenvectors of (20) is $sp(A_T^{-1}(d)\hat{\beta}) = sp(A_T^{-1}(d)\tilde{\beta})$, where $A_T^{-1}\tilde{\beta} = \left(\bar{\beta}, T^{-\frac{1}{2}+d}\gamma\right)'\tilde{\beta} = (I, T^{-\frac{1}{2}+d}U_T')'$. Thus we find that $T^{-\frac{1}{2}+d}U_T \xrightarrow{P} 0$. This shows consistency of $\tilde{\beta}$ and moreover that $\tilde{\beta}-\beta = o_P(T^{\frac{1}{2}-d})$. Note that (21) has $p-r$ zero roots and r positive roots given by the solutions of

$$|\lambda(d)\Sigma_{\beta\beta}(d) - \Sigma_{\beta 0}(d)\Sigma_{00}^{-1}\Sigma_{0\beta}(d)| = 0, \tag{22}$$

which can be expressed as

$$|\lambda(d)\Sigma_{\beta\beta}(d_0)\frac{a(d,d_0)}{a_0} - \Sigma_{\beta 0}(d_0)\Sigma_{00}^{-1}\Sigma_{0\beta}(d_0)\left[\frac{b(d,d_0)}{a_0}\right]^2| = 0, \tag{23}$$

so following Lemma 12 we get consistency of $\tilde{d}$.

Moreover, if $d=\tilde{d}$, $\tilde{d}$ is a consistent estimate of d, then (22) converges to

$$|\lambda(d_0)\Sigma_{\beta\beta}(d_0) - \Sigma_{\beta 0}(d_0)\Sigma_{00}^{-1}\Sigma_{0\beta}(d_0)| = 0.$$

Next recall $\tilde{\beta}-\beta = \bar{\gamma}U_T$, so

$$\begin{aligned}
\tilde{\beta}'S_{11}\left(\tilde{d}\right)\tilde{\beta} &= (\beta+\bar{\gamma}U_T)'S_{11}\left(\tilde{d}\right)(\beta+\bar{\gamma}U_T) \\
&= \beta'S_{11}\left(\tilde{d}\right)\beta + \beta'S_{11}\left(\tilde{d}\right)\bar{\gamma}U_T + (\bar{\gamma}U_T)'S_{11}\left(\tilde{d}\right)\beta + (\bar{\gamma}U_T)'S_{11}\left(\tilde{d}\right)(\bar{\gamma}U_T).
\end{aligned}$$

Since $U_T = o_P(T^{\frac{1}{2}-d_0})$ for this case, by consistency of $\tilde{d}$ and, by Lemma 9, we have that for all $\eta>0$

$$\begin{aligned}
\tilde{\beta}'S_{11}\left(\tilde{d}\right)\tilde{\beta} &= \beta'S_{11}\left(\tilde{d}\right)\beta + O_P(T^{d_0-1+\eta})o_P(T^{\frac{1}{2}-d_0}) + o_P(T^{1-2d_0})O_P(T^{2d_0-2+\eta}) \\
&= \beta'S_{11}\left(\tilde{d}\right)\beta + o_P(T^{\eta-1/2}) + o_P(T^{\eta-1}) \\
&= \beta'S_{11}(d_0)\beta + o_P(1) \xrightarrow{P} \Sigma_{\beta\beta}(d_0)
\end{aligned}$$

and also

$$\tilde{\beta}'S_{10}\left(\tilde{d}\right) = (\beta+\bar{\gamma}U_T)'S_{10}\left(\tilde{d}\right) = \beta'S_{10}\left(\tilde{d}\right) + o_P(T^{\frac{1}{2}-d_0}) \xrightarrow{P} \Sigma_{\beta 0}(d_0).$$

Further consider $\tilde{\alpha} = S_{01}\left(\tilde{d}\right)\tilde{\beta}\left(\tilde{\beta}'S_{11}\left(\tilde{d}\right)\tilde{\beta}\right)^{-1}$, which converges towards

$$\Sigma_{0\beta}(d_0)\Sigma_{\beta\beta}^{-1}(d_0) = \alpha\bar{\Sigma}_{\beta\beta}a_0\left[\bar{\Sigma}_{\beta\beta}a_0\right]^{-1} = \alpha$$

and

$$\hat{\Omega} = S_{00} - S_{01}\left(\tilde{d}\right)\tilde{\beta}\left(\tilde{\beta}'S_{11}\left(\tilde{d}\right)\tilde{\beta}\right)^{-1}\tilde{\beta}'S_{10}\left(\tilde{d}\right),$$

which converges towards

$$\begin{aligned}
\Sigma_{00} - \Sigma_{0\beta}(d_0)\Sigma_{\beta\beta}^{-1}(d_0)\Sigma_{\beta 0}(d_0) &= \Sigma_{00} - \alpha\bar{\Sigma}_{\beta\beta}a_0\left[\bar{\Sigma}_{\beta\beta}a_0\right]^{-1}\bar{\Sigma}_{\beta\beta}\alpha'a_0 = \\
&= \Sigma_{00} - \alpha\bar{\Sigma}_{\beta\beta}\alpha'a_0 = \Sigma_{00} - \alpha\Sigma_{\beta\beta}(d_0)\alpha' = \Omega.
\end{aligned}$$

■

Proof. (of Theorem 2) Using again Lemmas 8 and 10 we have consistency of $\tilde{\beta}$ and $\tilde{\beta}-\beta = o_P(T^{\frac{1}{2}-d})$ for any d, $d>0.5$, however, we do not have consistency of estimators of α and Ω for a fixed $d \neq \tilde{d}$, because $\tilde{\alpha}$ converges towards

$$\Sigma_{0\beta}(d)\,\Sigma_{\beta\beta}^{-1}(d) = \alpha\bar{\Sigma}_{\beta\beta}b(d,d_0)\left[\bar{\Sigma}_{\beta\beta}a(d,d_0)\right]^{-1} = \frac{b(d,d_0)}{a(d,d_0)}\alpha,$$

while $\hat{\Omega}$ converges towards

$$\begin{aligned}\Sigma_{00}-\Sigma_{0\beta}(d)\,\Sigma_{\beta\beta}^{-1}(d)\Sigma_{\beta 0}(d) &= \Sigma_{00}-\alpha\bar{\Sigma}_{\beta\beta}a_0\left[\bar{\Sigma}_{\beta\beta}a(d,d_0)\right]^{-1}\bar{\Sigma}_{\beta\beta}\alpha' a_0\\ &= \Sigma_{00}-\alpha a_0\left[a(d,d_0)\right]^{-1}\bar{\Sigma}_{\beta\beta}\alpha' a_0 = \Sigma_{00}-\alpha\frac{a_0}{a(d,d_0)}\Sigma_{\beta\beta}(d_0)\,\alpha' \neq \Omega.\end{aligned}$$

■

Proof. (of Theorems 3 and 5) The estimators $\tilde{\alpha}$, $\tilde{\beta}$, $\tilde{d}$ and $\hat{\Omega}$ satisfy the likelihood equations, so we derive expressions for the derivatives of $L(\alpha,\beta,d,\Omega)$, the concentrated log-likelihood function, with respect to β, α and d.

The expressions for the derivatives of $L(\alpha,\beta,d,\Omega)$ with respect to β and α in the directions b and a are respectively:

$$\begin{aligned}D_\beta\, L(\alpha,\beta,d,\Omega)\;(b) &= tr\left\{\Omega^{-1}\left(\sum_{t=1}^{T}\hat{\varepsilon}_t Z_{1t}'(d)\,b\alpha'\right)\right\} = T tr\left\{\alpha'\Omega^{-1}\left(S_{01}(d)-\alpha\beta' S_{11}(d)\right)b\right\},\\ D_\alpha\, L(\alpha,\beta,d,\Omega)\;(a) &= tr\left\{\Omega^{-1}\left(\sum_{t=1}^{T}\hat{\varepsilon}_t Z_{1t}'(d)\,\beta a'\right)\right\} = T tr\{\Omega^{-1}\left(S_{01}(d)-\alpha\beta' S_{11}(d)\right)\beta a'\},\end{aligned}$$

where $\hat{\varepsilon}_t = Z_{0t}-\alpha\beta' Z_{1t}(d)$
and the expression for the derivative with respect to d is

$$D_d\, L(\alpha,\beta,d,\Omega)\;(d) = tr\left\{\Omega^{-1}\left(\sum_{t=1}^{T}\hat{\varepsilon}_t Z_{1t}'(d)\,\beta\alpha'\right)\right\} = T tr\{\Omega^{-1}\left(S_{01}^{(1)}(d)-\alpha\beta' S_{11}^{(0,1)}(d)\right)\beta\alpha'\}.$$

From these results we can derive the first order conditions that are satisfied at a maximum point. At the point $(\tilde{\alpha},\tilde{\beta},\tilde{d})$ the derivatives are zero in all directions hence the likelihood equations are:

$$\begin{aligned}\tilde{\alpha}'\hat{\Omega}^{-1}\left(S_{01}\left(\tilde{d}\right)-\tilde{\alpha}\tilde{\beta}' S_{11}\left(\tilde{d}\right)\right) &= 0,\\ \left(S_{01}\left(\tilde{d}\right)-\tilde{\alpha}\tilde{\beta}' S_{11}\left(\tilde{d}\right)\right)\tilde{\beta} &= 0,\\ tr\{\hat{\Omega}^{-1}\left(S_{01}^{(1)}\left(\tilde{d}\right)-\tilde{\alpha}\tilde{\beta}' S_{11}^{(0,1)}\left(\tilde{d}\right)\right)\tilde{\beta}\tilde{\alpha}'\} &= 0.\end{aligned} \tag{24}$$

Now substitute $S_{01}^{(1)}\left(\tilde{d}\right) = S_{\varepsilon 1}^{(1)}(\tilde{d}) + \alpha\beta' S_{11}^{(0,1)}(d_0,\tilde{d})$ in the third equation with the obvious definition for $S_{\varepsilon 1}^{(1)}(\tilde{d})$,

$$tr\{\hat{\Omega}^{-1}\left(\begin{array}{c} S_{\varepsilon 1}^{(1)}(\tilde{d})\tilde{\beta}\tilde{\alpha}'-(\tilde{\alpha}-\alpha)\beta' S_{11}^{(0,1)}(d_0,\tilde{d})\tilde{\beta}\tilde{\alpha}'-\tilde{\alpha}(\tilde{\beta}-\beta)' S_{11}^{(0,1)}(d_0,\tilde{d})\tilde{\beta}\tilde{\alpha}'+\\ -\tilde{\alpha}\tilde{\beta}'\left[S_{11}^{(0,1)}\left(\tilde{d}\right)-S_{11}^{(0,1)}(d_0,\tilde{d})\right]\tilde{\beta}\tilde{\alpha}'\end{array}\right)\} = 0$$

and using Taylor expansion, Lemma 9 and consistency of $\tilde{\beta}$ and $\tilde{d}$

$$\begin{aligned}\tilde{\beta}'\left[S_{11}^{(0,1)}\left(\tilde{d}\right)-S_{11}^{(0,1)}(d_0,\tilde{d})\right]\tilde{\beta} &= \tilde{\beta}'S_{11}^{(1,1)}\left(d_0,\tilde{d}\right)\tilde{\beta}(\tilde{d}-d_0)+O_p\left((\tilde{d}-d_0)^2\right) \\ &= \frac{\pi^2}{6}\bar{\Sigma}_{\beta\beta}\left(\tilde{d}-d_0\right)\left(1+o_p\left(1\right)\right),\end{aligned}$$

we get that

$$tr\{\hat{\Omega}^{-1}\left(\begin{array}{c} S_{\varepsilon 1}^{(1)}(\tilde{d})\tilde{\beta}\tilde{\alpha}'-(\tilde{\alpha}-\alpha)\beta' S_{11}^{(0,1)}(d_0,\tilde{d})\tilde{\beta}\tilde{\alpha}'-\tilde{\alpha}(\tilde{\beta}-\beta)'S_{11}^{(0,1)}(d_0,\tilde{d})\tilde{\beta}\tilde{\alpha}'+ \\ -\frac{\pi^2}{6}\tilde{\alpha}\bar{\Sigma}_{\beta\beta}\tilde{\alpha}'\left(\tilde{d}-d_0\right)\left(1+o_p\left(1\right)\right)\end{array}\right)\}=0,$$

so

$$\begin{gathered}\tilde{d}-d_0=\left(\tfrac{\pi^2}{6}\bar{\Sigma}_{\beta\beta}\tilde{\alpha}'\hat{\Omega}^{-1}\tilde{\alpha}\right)^{-1}\times\left(1+o_p\left(1\right)\right) \\ \times\ tr\left\{\hat{\Omega}^{-1}\left[S_{\varepsilon 1}^{(1)}(\tilde{d})\tilde{\beta}\tilde{\alpha}'-(\tilde{\alpha}-\alpha)\beta' S_{11}^{(0,1)}(d_0,\tilde{d})\tilde{\beta}\tilde{\alpha}'-\tilde{\alpha}(\tilde{\beta}-\beta)'S_{11}^{(0,1)}(d_0,\tilde{d})\tilde{\beta}\tilde{\alpha}'\right]\right\}\end{gathered}$$

and therefore

$$\begin{gathered}T^{\frac{1}{2}}\left(\tilde{d}-d_0\right)=\left(\tfrac{\pi^2}{6}\bar{\Sigma}_{\beta\beta}\tilde{\alpha}'\hat{\Omega}^{-1}\tilde{\alpha}\right)^{-1}\times\left(1+o_p\left(1\right)\right) \\ \times\ tr\left\{\hat{\Omega}^{-1}\left[\begin{array}{c} T^{\frac{1}{2}}S_{\varepsilon 1}^{(1)}(\tilde{d})\left(\tilde{\beta}-\beta\right)\tilde{\alpha}'+T^{\frac{1}{2}}S_{\varepsilon 1}^{(1)}(\tilde{d})\beta\tilde{\alpha}' \\ -(\tilde{\alpha}-\alpha)T^{\frac{1}{2}}\beta' S_{11}^{(0,1)}(d_0,\tilde{d})\tilde{\beta}\tilde{\alpha}'-\tilde{\alpha}T^{\frac{1}{2}}(\tilde{\beta}-\beta)'S_{11}^{(0,1)}(d_0,\tilde{d})\tilde{\beta}\tilde{\alpha}'\end{array}\right]\right\}.\end{gathered}\tag{25}$$

Then, using Lemma 9 and consistency of $\tilde{\beta}$, we get

$$T^{1/2}(\tilde{d}-d_0)=O_p\left(1\right)+O_p\left(T^{1/2}\right)\left\|\tilde{\alpha}-\alpha\right\|+O_p\left(T^{d_0-1/2+\eta}\right)\left\|\tilde{\beta}-\beta\right\|.$$

Next consider the second equation in (24) and insert $S_{01}(\tilde{d})=\alpha\beta' S_{11}(d_0,\tilde{d})+S_{\varepsilon 1}(\tilde{d})$,

$$\begin{aligned} 0 &= (S_{\varepsilon 1}\left(\tilde{d}\right)+\alpha\beta' S_{11}\left(d_0,\tilde{d}\right)-\tilde{\alpha}\tilde{\beta}' S_{11}\left(\tilde{d}\right))\tilde{\beta} \\ &= S_{\varepsilon 1}\left(\tilde{d}\right)\tilde{\beta}-(\tilde{\alpha}-\alpha)\tilde{\beta}' S_{11}\left(\tilde{d}\right)\tilde{\beta}+\alpha\beta' S_{11}\left(d_0,\tilde{d}\right)\tilde{\beta}-\alpha\tilde{\beta}' S_{11}\left(\tilde{d}\right)\tilde{\beta}.\end{aligned}$$

Then, standardizing $\tilde{\alpha}-\alpha$ we obtain

$$\begin{aligned} T^{\frac{1}{2}}(\tilde{\alpha}-\alpha) &= \left\{T^{\frac{1}{2}}S_{\varepsilon 1}\left(\tilde{d}\right)\beta+T^{\frac{1}{2}}S_{\varepsilon 1}\left(\tilde{d}\right)(\tilde{\beta}-\beta)+T^{\frac{1}{2}}\alpha\beta' S_{11}\left(d_0,\tilde{d}\right)\tilde{\beta}-T^{\frac{1}{2}}\alpha\tilde{\beta}S_{11}\left(\tilde{d}\right)\tilde{\beta}\right\} \\ &\quad\times\left[\tilde{\beta}' S_{11}\left(\tilde{d}\right)\tilde{\beta}\right]^{-1}\end{aligned}$$

and rearranging terms and using Lemma 9 and consistency of $\tilde{\beta}$ we get

$$\begin{aligned} T^{\frac{1}{2}}(\tilde{\alpha}-\alpha) &= \left\{\begin{array}{c} T^{\frac{1}{2}}S_{\varepsilon 1}\left(\tilde{d}\right)\beta+T^{\frac{1}{2}}S_{\varepsilon 1}\left(\tilde{d}\right)(\tilde{\beta}-\beta)-T^{\frac{1}{2}}\alpha\left(\tilde{\beta}-\beta\right)' S_{11}\left(d_0,\tilde{d}\right)\tilde{\beta} \\ -T^{\frac{1}{2}}\alpha\tilde{\beta}'\left\{S_{11}\left(\tilde{d}\right)-S_{11}\left(d_0,\tilde{d}\right)\right\}\tilde{\beta}\end{array}\right\} \\ &\quad\times\left[a_0\bar{\Sigma}_{\beta\beta}\right]^{-1}\left(1+o_p\left(1\right)\right).\end{aligned}$$

Then using Taylor expansion and Lemma 9,

$$\begin{aligned}\tilde{\beta}'\left\{S_{11}\left(\tilde{d}\right)-S_{11}\left(d_0,\tilde{d}\right)\right\}\tilde{\beta} &= \tilde{\beta}'\left\{S_{11}^{(1,0)}\left(d_0,\tilde{d}\right)\right\}\tilde{\beta}\left(\tilde{d}-d_0\right)+O_p\left(\left(\tilde{d}-d_0\right)^2\right)\\ &= c_0\bar{\Sigma}_{\beta\beta}\left(\tilde{d}-d_0\right)(1+o_p(1))\end{aligned}$$

so that using again Lemma 9, it holds for all $\eta>0$,

$$\begin{aligned}T^{\frac{1}{2}}(\tilde{\alpha}-\alpha) &= O_p(1)+T^{1/2}O_p\left(T^{d_0-1+\eta}\right)o_p\left(T^{1/2-d_0}\right)+O_p\left(T^{\frac{1}{2}}T^{-\kappa}\right)\\ &= O_p(1)+o_p(T^{\eta})+O_p\left(T^{\frac{1}{2}}T^{-\kappa}\right)\end{aligned}$$

so $\tilde{\alpha}-\alpha=O_p\left(T^{-\kappa}+T^{\eta-1/2}\right)$. In fact in the limit

$$T^{\frac{1}{2}}(\tilde{\alpha}-\alpha) = \left\{T^{\frac{1}{2}}S_{\varepsilon 1}\left(\tilde{d}\right)\tilde{\beta}-T^{\frac{1}{2}}\alpha\dot{\Sigma}_{\beta\beta}(d_0)\left(\tilde{d}-d_0\right)+O_p\left(T^{d_0-1/2+\eta}\right)\left(\tilde{\beta}-\beta\right)\right\}\times \tag{26}$$

$$\times\{\Sigma_{\beta\beta}^{-1}(d_0)(1+o_p(1))\} \tag{27}$$

Consider now the first equation (24) and insert $S_{01}\left(\tilde{d}\right)=\alpha\beta' S_{11}(d_0,\tilde{d})+S_{\varepsilon 1}(\tilde{d})$ to get

$$\begin{aligned}0 &= \tilde{\alpha}'\hat{\Omega}^{-1}\left(S_{\varepsilon 1}\left(\tilde{d}\right)+\alpha\beta' S_{11}\left(d_0,\tilde{d}\right)-\tilde{\alpha}\tilde{\beta}' S_{11}\left(\tilde{d}\right)\right)\\ &= \tilde{\alpha}'\hat{\Omega}^{-1}\left(S_{\varepsilon 1}\left(\tilde{d}\right)+\alpha\beta'\{S_{11}\left(d_0,\tilde{d}\right)-S_{11}\left(\tilde{d}\right)\}-\tilde{\alpha}(\tilde{\beta}-\beta)' S_{11}\left(\tilde{d}\right)-(\tilde{\alpha}-\alpha)\beta' S_{11}\left(\tilde{d}\right)\right).\end{aligned}$$

We next multiply by $\bar{\gamma}$ from the right and insert $\tilde{\beta}-\beta=\bar{\gamma}U_T$,

$$0=\tilde{\alpha}'\hat{\Omega}^{-1}\left(S_{\varepsilon 1}\left(\tilde{d}\right)\bar{\gamma}+\alpha\beta'\{S_{11}\left(d_0,\tilde{d}\right)-S_{11}\left(\tilde{d}\right)\}\bar{\gamma}-\tilde{\alpha}U_T'\bar{\gamma}' S_{11}\left(\tilde{d}\right)\bar{\gamma}-(\tilde{\alpha}-\alpha)\beta' S_{11}\left(\tilde{d}\right)\bar{\gamma}\right)$$

so that

$$\begin{aligned}T^{d_0}U_T' = {} & \left(\tilde{\alpha}'\hat{\Omega}^{-1}\tilde{\alpha}\right)^{-1}\{\tilde{\alpha}'\hat{\Omega}^{-1}T^{1-d_0}S_{\varepsilon 1}\left(\tilde{d}\right)\bar{\gamma}\\ & +\tilde{\alpha}'\hat{\Omega}^{-1}T^{1-d_0}\alpha\beta'\{S_{11}\left(d_0,\tilde{d}\right)-S_{11}\left(\tilde{d}\right)\}\bar{\gamma}\\ & -\tilde{\alpha}'\hat{\Omega}^{-1}T^{1-d_0}(\tilde{\alpha}-\alpha)\beta' S_{11}\left(\tilde{d}\right)\bar{\gamma}\}\left[T^{1-2d_0}\bar{\gamma}' S_{11}\left(\tilde{d}\right)\bar{\gamma}\right]^{-1}.\end{aligned} \tag{28}$$

Then, using Taylor expansion again and following Lemma 9, for any $\eta>0$,

$$\begin{aligned}\beta'\{S_{11}\left(d_0,\tilde{d}\right)-S_{11}\left(\tilde{d}\right)\}\bar{\gamma} &= -\beta'\{S_{11}^{(1,0)}\left(d_0,\tilde{d}\right)\}\bar{\gamma}(\tilde{d}-d_0)+O_p\left((\tilde{d}-d_0)^2\right)\\ &= O_p\left(T^{d_0-1+\eta}\right)\left(\tilde{d}-d_0\right)\end{aligned}$$

and by Lemmas 10 and 9, consistency of $\tilde{\alpha}$, $\tilde{\Omega}$ and using the rate of convergence for $\tilde{d}$ and $\tilde{\alpha}$,

$$\begin{aligned}T^{d_0}U_T' &= O_p(1)\left\{O_p(1)+O_p(T^{\eta})\left[\left\|\tilde{d}-d\right\|+\|\tilde{\alpha}-\alpha\|\right]\right\}\\ &= O_p(1)\{O_p(1)+O_p\left(T^{\eta-\kappa}\right)+O_p(T^{\eta-1/2})\}\\ &= O_p(1)\end{aligned} \tag{29}$$

and therefore $\tilde{\beta}-\beta=O_p\left(T^{-d_0}\right)$, $0.5<d_0\leq 1$.

Now substituting (26) into (25) and ignoring the negligible terms in $\tilde{\beta}-\beta$, we find that in the limit

$$\begin{aligned}&\left(\tilde{d}-d_0\right)\left\{\left[\ddot{\Sigma}_{\beta\beta}\left(d_0\right)-\dot{\Sigma}_{\beta\beta}\left(d_0\right)\Sigma_{\beta\beta}^{-1}\left(d_0\right)\dot{\Sigma}_{\beta\beta}\left(d_0\right)\right]\alpha'\Omega^{-1}\alpha\right\}=\\&=-\,tr\left\{\Omega^{-1}\alpha\left(\dot{\Sigma}_{\beta\beta}\left(d_0\right)\Sigma_{\beta\beta}^{-1}\left(d_0\right)\beta'S_{\varepsilon 1}\left(d_0\right)-\beta'S_{\varepsilon 1}^{(1)}(d_0)\right)\right\}\left(1+o_p\left(1\right)\right)\end{aligned}\tag{30}$$

and therefore,

$$\tilde{d}-d_0=-\omega^{-1}tr\left\{\Omega^{-1}\alpha\left[\dot{\Sigma}_{\beta\beta}\left(d_0\right)\Sigma_{\beta\beta}^{-1}\left(d_0\right)\beta'S_{1\varepsilon}\left(d_0\right)-\beta'S_{1\varepsilon}^{(1)}(d_0)\right]\right\}\left(1+o_p\left(1\right)\right)$$

and the distribution of $\tilde{d}$ follows using Lemma 11.

For the distribution of $\tilde{\alpha}$ we can first write

$$\begin{aligned}\tilde{d}-d_0&=-\omega^{-1}tr\left\{\Omega^{-1}\alpha\left[\frac{c_0}{a_0}\beta'S_{1\varepsilon}\left(d_0\right)-\beta'S_{1\varepsilon}^{(1)}(d_0)\right]\right\}\left(1+o_p\left(1\right)\right)\\&=-\omega^{-1}\left(1+o_p\left(1\right)\right)\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t'\Omega^{-1}\alpha\left[\frac{c_0}{a_0}\beta'Z_{1t}\left(d_0\right)-\beta'Z_{1t}^{(1)}(d_0)\right]\end{aligned}$$

so that

$$\begin{aligned}T^{\frac{1}{2}}(\tilde{\alpha}-\alpha)&=\frac{\left(1+o_p\left(1\right)\right)}{T^{1/2}}\sum_{t=1}^{T}\left\{\begin{array}{c}\varepsilon_tZ_{1t}'\left(d_0\right)\beta\\-\omega^{-1}\alpha\dot{\Sigma}_{\beta\beta}\left(d_0\right)\varepsilon_t'\Omega^{-1}\alpha\left[\frac{c_0}{a_0}\beta'Z_{1t}\left(d_0\right)-\beta'Z_{1t}^{(1)}(d_0)\right]\end{array}\right\}\Sigma_{\beta\beta}^{-1}\left(d_0\right)\\&=\frac{\left(1+o_p\left(1\right)\right)}{T^{1/2}}\sum_{t=1}^{T}\left\{\begin{array}{c}\frac{1}{a_0}\varepsilon_tZ_{1t}'\left(d_0\right)\beta\bar{\Sigma}_{\beta\beta}^{-1}\\-\frac{c_0}{\omega a_0}\alpha\varepsilon_t'\Omega^{-1}\alpha\left[\frac{c_0}{a_0}\beta'Z_{1t}\left(d_0\right)-\beta'Z_{1t}^{(1)}(d_0)\right]\end{array}\right\}.\end{aligned}\tag{31}$$

Taking vec's and using that

$$vec\left(AXB\right)=\left(B'\otimes A\right)vec\left(X\right),tr\left(A'BCD'\right)=vec\left(A\right)\left(D\otimes B\right)vec\left(C\right),$$

and ignoring $o_p\left(1\right)$ terms,

$$T^{\frac{1}{2}}vec(\tilde{\alpha}-\alpha)=T^{-\frac{1}{2}}\sum_{t=1}^{T}\left\{\begin{array}{c}\frac{1}{a_0}\left(\bar{\Sigma}_{\beta\beta}^{-1}\beta'Z_{1t}\left(d_0\right)\otimes I\right)vec\left(\varepsilon_t\right)\\-\frac{c_0}{\omega a_0}vec\left(\alpha\right)tr\left\{\varepsilon_t'\Omega^{-1}\alpha\left[\frac{c}{a}\beta'Z_{1t}\left(d_0\right)-\beta'Z_{1t}^{(1)}(d_0)\right]\right\}\end{array}\right\}$$

$$=T^{-\frac{1}{2}}\sum_{t=1}^{T}\left\{\begin{array}{c}\frac{1}{a_0}\left(\bar{\Sigma}_{\beta\beta}^{-1}\beta'Z_{1t}\left(d_0\right)\otimes I\right)\varepsilon_t\\-\frac{c_0}{\omega a_0}vec\left(\alpha\right)vec\left(\varepsilon_t'\Omega^{-1}\right)\left(\left[\frac{c}{a}\beta'Z_{1t}\left(d_0\right)-\beta'Z_{1t}^{(1)}(d_0)\right]'\otimes I\right)vec\left(\alpha\right)\end{array}\right\}$$

$$=\quad T^{-\frac{1}{2}}\sum_{t=1}^{T}\frac{1}{a_0}\left(\bar{\Sigma}_{\beta\beta}^{-1}\beta'Z_{1t}\left(d_0\right)\otimes I\right)\varepsilon_t\tag{32}$$

$$-T^{-\frac{1}{2}}\frac{c_0}{\omega a_0}vec\left(\alpha\right)vec\left(\alpha\right)'\sum_{t=1}^{T}\left(\left[\frac{c}{a}\beta'Z_{1t}\left(d_0\right)-\beta'Z_{1t}^{(1)}(d_0)\right]\otimes I\right)\Omega^{-1}\varepsilon_t.\tag{33}$$

Then the distribution for $T^{\frac{1}{2}}vec(\tilde{\alpha}-\alpha)$ follows by a standard martingale difference CLT, noting that the contributions to its asymptotic variance are

$$\begin{aligned} V\left((33)\right) &= \frac{c_0^2}{\omega a_0^2} vec\left(\alpha\right) vec\left(\alpha\right)' \\ V\left((32)\right) &= \bar{\Sigma}_{\beta\beta}^{-1} \otimes \Omega \end{aligned}$$

$$\begin{aligned} Cov\left((32),(33)\right) &= -\frac{c_0}{\omega a_0^2} vec\left(\alpha\right) vec\left(\alpha\right)' \\ &\times \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} E\left(\left[\frac{c_0}{a_0}\beta' Z_{1t}\left(d_0\right) - \beta' Z_{1t}^{(1)}\left(d_0\right)\right] \otimes I\right) \left(Z_{1t}'\left(d_0\right)\beta\bar{\Sigma}_{\beta\beta}^{-1} \otimes I_n\right) \end{aligned}$$

$$\begin{aligned} Cov\left((32),(33)\right) &= -\frac{c_0}{\omega a_0^2} vec\left(\alpha\right) vec\left(\alpha\right)' \\ &\times \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} E\left(\left[\frac{c_0}{a_0}\beta' Z_{1t}\left(d_0\right) - \beta' Z_{1t}^{(1)}\left(d_0\right)\right] Z_{1t}'\left(d_0\right)\beta\bar{\Sigma}_{\beta\beta}^{-1} \otimes I\right) \\ &= -\frac{c_0}{\omega a_0^2} vec\left(\alpha\right) vec\left(\alpha\right)' \left(\left[\frac{c_0}{a_0} a_0\bar{\Sigma}_{\beta\beta} - c_0\bar{\Sigma}_{\beta\beta}\right] \bar{\Sigma}_{\beta\beta}^{-1} \otimes I\right) = 0. \end{aligned}$$

Finally recall (28).

$$\begin{aligned} T^{d_0}U_T' &= \left(\tilde{\alpha}'\hat{\Omega}^{-1}\tilde{\alpha}\right)^{-1} \{\tilde{\alpha}'\hat{\Omega}^{-1}T^{1-d_0}S_{\varepsilon 1}\left(\tilde{d}\right)\bar{\gamma} \\ &+\tilde{\alpha}'\hat{\Omega}^{-1}T^{1-d_0}\alpha\beta'\{S_{11}\left(d_0,\tilde{d}\right) - S_{11}\left(\tilde{d}\right)\}\bar{\gamma} \\ &-\tilde{\alpha}'\hat{\Omega}^{-1}T^{1-d_0}(\tilde{\alpha}-\alpha)\beta' S_{11}\left(\tilde{d}\right)\bar{\gamma}\} \left[T^{1-2d_0}\bar{\gamma}' S_{11}\left(\tilde{d}\right)\bar{\gamma}\right]^{-1} \end{aligned} \tag{34}$$

and note what happened to the middle term

$$\begin{aligned} \beta'\{S_{11}\left(d_0,\tilde{d}\right) - S_{11}\left(\tilde{d}\right)\}\bar{\gamma} &= -\beta'\{S_{11}^{(1,0)}\left(d_0,\tilde{d}\right)\}\bar{\gamma}(\tilde{d}-d_0) + O_p\left((\tilde{d}-d_0)^2\right) \\ &= O_p\left(T^{d_0-1+\eta}\right)\left(\tilde{d}-d_0\right). \end{aligned}$$

By Lemma 10 and the $T^{\frac{1}{2}}$ consistency of $\tilde{\alpha}$ for $d_0 > \frac{1}{2}$, the last term of (28) converges in probability to zero and the consistency of $\hat{\Omega}$ then implies that

$$T^{d_0}U_T = \left[\bar{\gamma}' T^{1-2d_0}S_{11}\left(d_0\right)\bar{\gamma}\right]^{-1} \bar{\gamma}' T^{1-d_0}S_{\epsilon 1}'\left(d_0\right)\Omega^{-1}\alpha\left(\alpha'\Omega^{-1}\alpha\right)^{-1} + o_P\left(1\right),$$

which converges in d towards the limit given in the theorem. ■

11 Appendix C

In this appendix an alternative proof of main results of this paper is provided. This proof is based on Lemma 1 in Andrews and Sun (2004) and shows that sufficient conditions for the existence of a consistent sequence of solutions of a sequence of stochastic optimization problems are satisfied.

Proof. Assuming $r = 1$ and that Ω is known the first derivatives of the log-likelihood $L(a, b, d) = L(a, b, \Omega, d)$ are

$$\begin{aligned} L_\beta(\theta) &= \frac{\partial}{\partial b} L(a,b,d) = T\left(S_{10}(d) - S_{11}(d)\, ba'\right)\Omega^{-1} a \\ L_\alpha(\theta) &= \frac{\partial}{\partial a} L(a,b,d) = T\Omega^{-1}\left(S_{01}(d) - ab' S_{11}(d)\right) b \\ L_d(\theta) &= \frac{\partial}{\partial d} L(a,b,d) = Ttr\left\{ab'\left(S_{10}^{(1)}(d) - S_{11}^{(1)}(d)\, ba'\right)\Omega^{-1}\right\} \end{aligned}$$

where $\theta = (b', a', d)'$, while the second derivatives are

$$\begin{aligned} L_{\beta\beta}(\theta) &= \frac{\partial}{\partial b'} L_\beta = -Ta'\Omega^{-1} a S_{11}(d) \\ L_{\beta\alpha}(\theta) &= \frac{\partial}{\partial a'} L_\beta = -T\left\{2S_{11}(d)\, ba' - S_{10}(d)\right\}\Omega^{-1} \\ L_{\beta d}(\theta) &= \frac{\partial}{\partial d} L_\beta = -T\left\{2S_{11}^{(1)}(d)\, ba' - S_{10}^{(1)}(d)\right\}\Omega^{-1} a \\ L_{\alpha d}(\theta) &= \frac{\partial}{\partial d} L_\alpha = -T\Omega^{-1}\left(2ab' S_{11}^{(1)}(d) - \dot{S}_{01}(d)\right) b \\ L_{\alpha\alpha}(\theta) &= \frac{\partial}{\partial a'} L_\alpha = -Tb' S_{11}(d)\, b\Omega^{-1} \\ L_{dd}(\theta) &= \frac{\partial}{\partial d} L_d = -Ttr\left\{ab'\left(S_{11}^{(1,1)}(d)\, ba' + S_{11}^{(2,0)}(d)\, ba' - S_{10}^{(2)}(d)\right)\Omega^{-1}\right\}. \end{aligned}$$

We check the conditions of Lemma 1 in Andrews and Sun (2004) for

$$B_T = diag\left(\left(\bar{\gamma} T^{d_0}\ \bar{\beta} T^{1/2}\right), I_{k+1} T^{1/2}\right).$$

Now we have that

$$\begin{aligned} \bar{\gamma}' T^{-d_0} L_\beta(\theta_0) &\to {}_p \bar{\gamma}' T^{1-d_0} S_{1\varepsilon}(d_0)\,\Omega^{-1}\alpha \to_d \bar{\gamma}'\Omega^{1/2}\int_0^1 W_{d_0} dW'\ \Omega^{-1/2}\alpha \\ T^{-1/2} L_\alpha(\theta_0) &\to {}_p T^{1/2}\Omega^{-1} S_{\varepsilon 1}(d_0)\,\beta \to_d N\left(0, a_0\bar{\Sigma}_{\beta\beta}\Omega^{-1}\right) \\ T^{-1/2} L_d(\theta_0) &\to {}_p T^{1/2} tr\left\{\alpha\beta' S_{1\varepsilon}^{(1)}(d_0)\,\Omega^{-1}\right\} \to_p N\left(0, \frac{\pi^2}{6}\bar{\Sigma}_{\beta\beta} tr\left\{\alpha'\Omega^{-1}\alpha\right\}\right) \end{aligned}$$

and

$$\beta' T^{-1/2} L_\beta(\theta_0) \to_p \beta' T^{1/2} S_{1\varepsilon}(d_0)\,\Omega^{-1}\alpha \to_d N\left(0, a_0\bar{\Sigma}_{\beta\beta}\alpha'\Omega_0^{-1/2}\alpha\right)$$

while

$$
\begin{aligned}
T^{-2d_0}\bar{\gamma}' L_{\beta\beta}(\theta_0)\bar{\gamma} &\to_p -\alpha'\Omega^{-1}\alpha\bar{\gamma}'\Omega^{1/2}\int_0^1 W_{d_0}W_{d_0}'\ \Omega^{1/2}\bar{\gamma} \\
T^{-d_0-1/2}\bar{\gamma}' L_{\beta\alpha}(\theta_0) &= -T^{-d_0+1/2}\bar{\gamma}' S_{11}(d)\beta\alpha'\Omega^{-1} \to_p 0 \\
T^{-d_0-1/2}\bar{\gamma}' L_{\beta d}(\theta_0) &= -T^{-d_0+1/2}\bar{\gamma}' S_{11}^{(1)}(d)\beta\alpha'\Omega^{-1}\alpha \to_p 0 \\
T^{-1}L_{\alpha d}(\theta_0) &= -\Omega^{-1}\alpha\beta' S_{11}^{(1)}(d)\beta \to_p -\bar{\Sigma}_{\beta\beta}c_0\Omega^{-1}\alpha \\
T^{-1}L_{\alpha\alpha}(\theta_0) &\to_p -a_0\bar{\Sigma}_{\beta\beta}\Omega^{-1} \\
T^{-1}L_{dd}(\theta_0) &= -tr\left\{\alpha\beta'\left(S_{11}^{(1,1)}(d_0)\beta\alpha' + S_{1\varepsilon}^{(2)}(d_0)\right)\Omega^{-1}\right\} \to_p -\frac{\pi^2}{6}tr\left\{\bar{\Sigma}_{\beta\beta}\alpha'\Omega^{-1}\alpha\right\}.
\end{aligned}
$$

and

$$
\begin{aligned}
T^{-d_0-1/2}\beta' L_{\beta\beta}(\theta_0)\bar{\gamma} &\to_p 0 \\
T^{-1}\beta' L_{\beta\beta}(\theta_0)\beta &\to_p -a_0\alpha'\Omega^{-1}\alpha\bar{\Sigma}_{\beta\beta} \\
T^{-1}\beta' L_{\beta\alpha}(\theta_0) &= -\beta' S_{11}(d)\beta\alpha'\Omega^{-1} \to_p -a_0\bar{\Sigma}_{\beta\beta}\alpha'\Omega^{-1} \\
T^{-1}\beta' L_{\beta d}(\theta_0) &= -\beta' S_{11}^{(1)}(d)\beta\alpha'\Omega^{-1}\alpha \to_p -c_0\bar{\Sigma}_{\beta\beta}\bar{\Sigma}_{\beta\beta}\alpha'\Omega^{-1}\alpha.
\end{aligned}
$$

Therefore $\left(B_T^{-1}\right)' L_{\theta\theta}(\theta_0) B_T^{-1}$ converges to a matrix that is positive definite with probability one.

The fourth point in Andrews and Sun's Lemma can be checked if

$$
\begin{aligned}
T^{1-2d_0}\left\|\bar{\gamma}'\left\{S_{11}(d) - S_{11}(d_0)\right\}\bar{\gamma}\right\| &\to_p 0 \\
\left\|b' S_{11}(d)b - \beta' S_{11}(d_0)\beta\right\| &\to_p 0
\end{aligned}
$$

and if the second statement also holds with S_{11} replaced by $S_{11}^{(1)}$, $S_{11}^{(1,1)}$ or $S^{(2,0)}$, uniformly for θ so that $\left\|\left(B_T^{-1}\right)'(\theta - \theta_0)\right\| \le k_T$ for some $k_T \to \infty$ (e.g. $k_T = \log T$). The first statement is equivalent to show that

$$
\sup_{|d-d_0|\le KT^{-1/2}\log T} T^{1-2d_0}\left\|\bar{\gamma}'\left\{S_{11}(d) - S_{11}(d_0)\right\}\bar{\gamma}\right\| \to_p 0,
$$

which follows by pointwise convergence and tightness of $T^{1-2d_0}\bar{\gamma}' S_{11}(d)\bar{\gamma}$, cf. Theorem 1 of Łasak (2010).

For the second statement, we have that by the triangle inequality

$$
\begin{aligned}
\left\|b' S_{11}(d)b - \beta' S_{11}(d_0)\beta\right\| \le\ & \left\|b' S_{11}(d)b - b' S_{11}(d_0)b\right\| \\
& + \left\|b' S_{11}(d_0)b - \beta' S_{11}(d_0)\beta\right\|.
\end{aligned}
$$

Now, for $\|b-\beta\| \le T^{-d_0}\log T$,

$$
\begin{aligned}
& T^{1-2d_0} b'\left\{S_{11}(d) - S_{11}(d_0)\right\}b \\
= \ & T^{-2d_0}\sum_{t=1}^{T}(b\pm\beta)'\left\{Z_{1t}(d)Z_{1t}'(d) \pm Z_{1t}(d_0)Z_{1t}'(d) - Z_{1t}(d_0)Z_{1t}'(d_0)\right\}(b\pm\beta).
\end{aligned}
$$

A typical term is then

$$\begin{aligned}
& T^{-2d_0} \sum_{t=1}^{T} (b-\beta)' Z_{1t}(d) \left\{ Z_{1t}'(d) - Z_{1t}'(d_0) \right\} \beta \\
= & T^{-2d_0} \sum_{t=1}^{T} (b-\beta)' Z_{1t}(d) \left\{ \Delta^{-d} - \Delta^{-d_0} \right\} \Delta X_t' \beta \\
= & T^{-2d_0} \sum_{t=1}^{T} (b-\beta)' \left\{ \Delta^{-d} - 1 \right\} \Delta X_t \left\{ \Delta^{-d} - \Delta^{-d_0} \right\} \Delta^{d_0} \varepsilon_t' \beta \\
= & T^{-2d_0} \sum_{t=1}^{T} (b-\beta)' \sum_{j=1}^{t} \phi_j \varepsilon_t \left\{ \Delta^{d_0-d} - 1 \right\} \varepsilon_t' \beta \qquad (35)
\end{aligned}$$

where $\phi_j \sim j^{d-1}$ as $j \to \infty$, whereas the weights of the filter $\Delta^{d_0-d} - 1$ can be bounded by $|d_0 - d| j^{-1} \log j$ for $|d_0 - d| \leq T^{-1/2} \log T$. Then

$$\sum_{t=1}^{T} \sum_{j=1}^{t} \phi_j \varepsilon_t \left\{ \Delta^{d_0-d} - 1 \right\} \varepsilon_t'$$

is $O_p(T)$ using the same techniques as in the proof of Lemma 8, and therefore (35) is $o_p(1)$ uniformly in d and b. ■

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] Andrews, D. W. K., Sun, Y. (2004), Adaptive Local Polynomial Whittle Estimation of Long-range Dependence, Econometrica, 72, 569–614.

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

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

[5] Chen, W., Hurvich, C. (2003a), Estimating fractional cointegration in the presence of polynomial trends, Journal of Econometrics, 117, 95-121.

[6] Chen, W., Hurvich, C. (2003b), Semiparametric Estimation of Multivariate Fractional Cointegration, Journal of American Statistical Association, 98, 629-642.

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

[8] Christensen, B. J., Nielsen, M. Ø. (2006), Asymptotic normality of narrow-band least square in the stationary fractional cointegration model and volatility forecasting, Journal of Econometrics, 133, 343-371.

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

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

[11] Dolado, J. J., Marmol, F. (2004), Asymptotic inference results for Multivariate long-memory processes, The Econometrics Journal, 7, 168-190.

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

[13] Doornik, J. A., Ooms, M. (2006), Introduction to Ox, www.doornik.com/ox/OxIntro.pdf.

[14] 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.

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

[16] Fisher, I. (1896), Appreciation and interest, Publications of the American Economic Association, 11, 1–98.

[17] Franchi, M. (2009), A representation theory for polynomial cofractionality in vector autoregressive models, Econometric Theory, in press.

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

[19] 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.

[20] Granger, C. W. J. (1986), Developments in the Study of Cointegrated Economic Variables, Oxford Bulletin of Economics and Statistics, 48, 213-28.

[21] Hassler, U., Marmol, F., Velasco, C. (2008), Fractional cointegration in the presence of linear trends, Journal of Time Series Analysis, 29, 1088-1103.

[22] Hualde, J., Robinson, P. M. (2006), Semiparametric Estimation of Fractional Cointegration, Working Paper 07/06, Universidad de Navarra.

[23] Hualde, J., Robinson, P. M. (2007), Root-N-Consistent Estimation Of Weak Fractional Cointegration, Journal of Econometrics, 127, 450-484.

[24] Iacone, F. (2009), A Semiparametric Analysis of the Term Structure of the US Interest Rates, Oxford Bulletin of Economics and statistics, 71, 475-490.

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

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

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

[28] Johansen, S. (2009), Representation of cointegrated autoregressive processes with application to fractional processes, Econometric Reviews, 28, 121-145.

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

[30] Johansen, S., Nielsen, M. Ø. (2010a), Likelihood inference for a nonstationary fractional autoregressive model, Journal of Econometrics, forthcoming.

[31] Johansen, S., Nielsen, M. Ø. (2010b), Likelihood inference for a vector autoregressive model which allows for fractional and cofractional processes, work in progress.

[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] Łasak, K. (2010), Likelihood based testing for no fractional cointegration, Journal of Econometrics, forthcoming.

[35] Łasak, K., Velasco, C. (2010), Fractional Cointegration Rank Estimation, work in progress.

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

[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 (6), 1809-1844.

[40] Nielsen, M. Ø. (2007), Local Whittle analysis of stationary fractional cointegration and implied-realized volatility relation, Journal of Business and Economic Statistics, 25, 427-446.

[41] Nielsen, M. Ø., Frederiksen, P. (2007), Fully modified narrow-band least squares estimation of stationary fractional cointegration, Preprint.

[42] Nielsen, M. Ø. (2010), Nonparametric cointegration analysis of fractional systems with unknown integration orders, Journal of Econometrics, in press.

[43] Phillips, P. C. B. (1991), Optimal Inference in Cointegrated Systems, Econometrica, 59 (2), 283-306.

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

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

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

[47] Robinson, P. M., Iacone, F. (2005), Cointegration in fractional systems with deterministic trends, Journal of Econometrics, 129, 263-298.

[48] Robinson, P. M., Marinucci, D. (2001), Narrow-Band Analysis of Nonstationary Process, The Annals of Statistics, 29, 947-986.

[49] Robinson, P. M., Marinucci, D. (2003), Frequency Domain Analysis of Fractional Cointegration, in P. M. Robinson, Time Series with Long Memory, Oxford University Press.

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

[51] Rossi, E., Santucci de Magistris, P. (2009), A No Arbitrage Fractional Cointegration Analysis of the Range Based Volatility, CREATES Research Paper, 2009-31.

[52] Sibbertsen, P., Kruse, R. (2009), Testing for a break in persistence under long-range dependencies, Journal of Time Series Analysis, 30, 263-285.

[53] Sims, C. A., Stock, J. H., Watson, M. W. (1990), Inference in linear time series with some unit root, Econometrica, 58, 113-144.

[54] Soderlind, P., Svensson, L. (1997), New techniques to extract market expectations from financial instruments, Journal of Monetary Economics, 40, 383-429.

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

Research Papers 2008

2008-37: Dennis Kristensen: Uniform Convergence Rates of Kernel Estimators with Heterogenous, Dependent Data

2008-38: Christian M. Dahl and Emma M. Iglesias: The limiting properties of the QMLE in a general class of asymmetric volatility models

2008-39: Roxana Chiriac and Valeri Voev: Modelling and Forecasting Multivariate Realized Volatility

2008-40: Stig Vinther Møller: Consumption growth and time-varying expected stock returns

2008-41: Lars Stentoft: American Option Pricing using GARCH models and the Normal Inverse Gaussian distribution

2008-42: Ole E. Barndorff-Nielsen, Silja Kinnebrock and Neil Shephard: Measuring downside risk — realised semivariance

2008-43: Martin Møller Andreasen: Explaining Macroeconomic and Term Structure Dynamics Jointly in a Non-linear DSGE Model

2008-44: Christian M. Dahl, Henrik Hansen and John Smidt: The cyclical component factor model

2008-45: Christian M. Dahl and Yu Qin: The limiting behavior of the estimated parameters in a misspecified random field regression model

2008-46: Bent Jesper Christensen, Christian M. Dahl and Emma M. Iglesias: Semiparametric Inference in a GARCH-in-Mean Model

2008-47: Charlotte Christiansen: Mean Reversion in US and International Short Rates

2008-48: Tim Bollerslev, George Tauchen and Hao Zhou: Expected Stock Returns and Variance Risk Premia

2008-49: Tim Bollerslev: Glossary to ARCH (GARCH)

2008-50: Giuseppe Cavaliere, Anders Rahbek and A.M.Robert Taylor: Testing for Co-integration in Vector Autoregressions with Non-Stationary Volatility

2008-51: Bent Jesper Christensen and Michael Sørensen: Optimal inference in dynamic models with conditional moment restrictions

2008-52: Katarzyna Lasak: Likelihood based testing for no fractional cointegration

2008-53: Katarzyna Lasak: Maximum likelihood estimation of fractionally cointegrated systems