AARHUS UNIVERSITY
BUSINESS AND SOCIAL SCIENCES
DEPARTMENT OF ECONOMICS AND BUSINESS

CREATES
Center for Research in Econometric Analysis of Time Series

On the identification of fractionally cointegrated VAR models with the F(d) condition

Federico Carlini and Paolo Santucci de Magistris

CREATES Research Paper 2013-44

Department of Economics and Business
Aarhus University
Fuglesangs Allé 4
DK-8210 Aarhus V
Denmark

Email: oekonomi@au.dk
Tel: +45 8716 5515

On the identification of fractionally cointegrated VAR models with the $\mathcal{F}(d)$ condition

Federico Carlini* Paolo Santucci de Magistris*

December 11, 2013

Abstract

This paper discusses identification problems in the fractionally cointegrated system of Johansen (2008) and Johansen and Nielsen (2012). The identification problem arises when the lag structure is over-specified, such that there exist several equivalent re-parametrization of the model associated with different fractional integration and cointegration parameters. The properties of these multiple non-identified sub-models are studied and a necessary and sufficient condition for the identification of the fractional parameters of the system is provided. The condition is named $\mathcal{F}(d)$. The assessment of the $\mathcal{F}(d)$ condition in the empirical analysis is relevant for the determination of the fractional parameters as well as the lag structure.

Keywords: Fractional Cointegration; Cofractional Models; Identification; Lag Selection.

JEL Classification: C19, C32

*The authors acknowledge support from CREATES - Center for Research in Econometric Analysis of Time Series (DNRF78), funded by the Danish National Research Foundation. The authors are grateful to Søren Johansen, Morten Nielsen, Niels Haldrup and Rocco Mosconi for insightful comments on this work. The authors would like to thank the participants to the Third Long Memory Symposium (Aarhus 2013).

1 Introduction

The last decade has witnessed an increasing interest in the statistical definition and evaluation of the concept of *fractional cointegration*, as a generalization of the idea of cointegration to processes with fractional degrees of integration. In the context of long-memory processes, fractional cointegration allows linear combinations of $I(d)$ processes to be $I(d-b)$, with $d \in \mathbb{R}_+$ and $0 < b \leq d$. More specifically, the concept of fractional cointegration implies the existence of one, or more, common stochastic trends, integrated of order d, with short-period departures from the long-run equilibrium integrated of order $d-b$. The coefficient b is the degree of fractional reduction obtained by the linear combination of $I(d)$ variables, namely the *cointegration gap*.

Interestingly, the seminal paper by Engle and Granger (1987) already introduced the idea of common trends between $I(d)$ processes, but the subsequent theoretical works, see among many others Johansen (1988), have mostly been dedicated to cases with integer orders of integration. Notable methodological works in the field of fractional cointegration are Robinson and Marinucci (2003) and Christensen and Nielsen (2006), which develop regression-based semi-parametric methods to evaluate whether two fractional stochastic processes share common trends. More recently, Nielsen and Shimotsu (2007) provide a testing procedure to evaluate the cointegration rank of the multivariate coherence matrix of two, or more, fractionally differenced series. Despite the effort spent in defining testing procedures for the presence of fractional cointegration, the literature in this area lacked a coherent multivariate model explicitly characterizing the joint behaviour of fractionally cointegrated processes. Only recently, Johansen (2008) and Johansen and Nielsen (2012) have proposed the FCVAR$_{d,b}$ model, an extension of the well-known VECM to fractional processes, which represents a tool for a direct modeling and testing of fractional cointegration. Johansen (2008) and Johansen and Nielsen (2012) study the properties of the model and provide a method to obtain consistent estimates when the lag structure of the model is correctly specified.

The present paper shows that the FCVAR$_{d,b}$ model is not globally identified, i.e. for a given number of lags, k, there may exist several sub-models with the same conditional densities but different values of the parameters, and hence cannot be identified. The multiplicity of not-identified sub-models can be characterized for any $\text{FCVAR}_{d,b}$ model with k lags. An analogous identification problem, for the FIVAR_b model, induced by the generalized lag operator is discussed in Tschernig et al. (2013a,b).

A solution for this identification problem is provided in this paper. It is proved that the $I(1)$ condition in the VECM of Johansen (1988) can be generalized to the fractional context. This condition is named $\mathcal{F}(d)$, and it is a necessary and sufficient condition for the identification of the system. This condition can be used to correctly identify the lag structure of the model and to consistently estimate the parameter vector.

The consequence of the lack of identification of the $\text{FCVAR}_{d,b}$ is investigated from a statistical point of view. Indeed, as a consequence of the identification problem, the expected likelihood function is maximized in correspondence of several parameter vectors when the lag order is not correctly specified. Hence, the fractional and co-fractional parameters cannot be uniquely estimated if the true lag structure is not correctly determined. Therefore, a lag selection procedure, integrating the likelihood ratio test with an evaluation of the $\mathcal{F}(d)$ condition, is proposed and tested. A simulation study shows that the proposed method provides the correct lag specification in most cases.

Finally, a further identification issue is discussed. It is proved that there is a potentially large number of parameters sets associated with different choices of lag length and cointegration rank for which the conditional density of the $\text{FCVAR}_{d,b}$ model is the same. This problem has practical consequences when testing for the nullity of the cointegration rank and the true lag length is unknown. For example, under certain restrictions on the sets of parameters, the $\text{FCVAR}_{d,b}$ with full rank and k lags is equivalent to the $\text{FCVAR}_{d,b}$ with rank 0 and $k+1$ lags. It is shown that the evaluation of the $\mathcal{F}(d)$ condition provides a solution to this identification problem and works in most cases.

This paper is organized as follows. Section 2 discusses the identification problem from a theoretical point of view. Section 3 discusses the consequences of the lack of identification on the inference on the parameters of the $\text{FCVAR}_{d,b}$ model. Section 4 presents the method to optimally select the number of lags and provides evidence, based on simulation, on the performance of the method in finite sample. Section 5 discusses the problems when the cointegration rank and the lag length are both unknown. Section

6 concludes the paper.

2 The Identification Problem

This Section provides a discussion of the identification problem related to the $\text{FCVAR}_{d,b}$ model

$$\mathcal{H}_k: \quad \Delta^d X_t = \alpha\beta'\Delta^{d-b}L_b X_t + \sum_{i=1}^{k}\Gamma_i \Delta^d L_b^i X_t + \varepsilon_t \quad \varepsilon_t \sim iidN(0,\Omega) \tag{1}$$

where X_t is a p-dimensional vector, α and β are $p \times r$ matrices, where r defines the cointegration rank.[1] Ω is the positive definite covariance matrix of the errors, and Γ_j, for $j = 1, \ldots, k$, are $p \times p$ matrices loading the short-run dynamics. The operator $L_b := 1 - \Delta^b$ is the so called *fractional lag operator*, which, as noted by Johansen (2008), is necessary for characterizing the solutions of the system. $\mathcal{H}_k$ defines the model with k lags and $\theta = vec(d, b, \alpha, \beta, \Gamma_1, ..., \Gamma_k, \Omega)$ is the parameter vector. Similarly to Johansen (2010), the concept of identification and equivalence between two models is formally introduced by the following definition.

Definition 2.1 *Let $\{P_\theta, \theta \in \Theta\}$ be a family of probability measures, that is, a statistical model. We say that a parameter function $g(\theta)$ is identified if $g(\theta_1) \neq g(\theta_2)$ implies that $P_{\theta_1} \neq P_{\theta_2}$. On the other hand, if $P_{\theta_1} = P_{\theta_2}$ and $g(\theta_1) \neq g(\theta_2)$, the two models are equivalent or not identified.*

As noted by Johansen and Nielsen (2012), the parameters of the $\text{FCVAR}_{d,b}$ model in (1) are not identified, i.e. there exist several *equivalent* sub-models associated with different values of the parameter vector, θ.

An illustration of the identification problem is provided by the following example. Consider the $\text{FCVAR}_{d,d}$ model with one lag[2]

$$\mathcal{H}_1: \Delta^d X_t = \alpha\beta' L_d X_t + \Gamma_1 \Delta^d L_d X_t + \varepsilon_t$$

[1] The results of this Section are obtained under the maintained assumption that the cointegration rank is known and such that $0 < r < p$. An extension to case of unknown rank and number of lags is presented in Section 5.

[2] To simplify the exposition, we consider the case $\text{FCVAR}_{d,b}$ with $d = b$.

where $d > 0$. Consider the following two restrictions, leading to the sub-models:

$$\mathcal{H}_{1,0}: \qquad \mathcal{H}_1 \text{ under the constraint } \Gamma_1 = 0 \tag{2}$$

$$\mathcal{H}_{1,1}: \qquad \mathcal{H}_1 \text{ under the constraint } \Gamma_1 = I_p + \alpha\beta' \tag{3}$$

Interestingly, these two sets of restrictions lead to equivalent sub-models with different parameter vectors. The sub-model $\mathcal{H}_{1,0}$ can be formulated as:

$$\Delta^{d_1} X_t = \alpha\beta' L_{d_1} X_t + \varepsilon_t \tag{4}$$

where d_1 is the fractional parameter under $\mathcal{H}_{1,0}$, and the restriction $\Gamma_1 = 0$ corresponds to a FCVAR$_{d,d}$ model with no lags, $\mathcal{H}_0$. After a simple manipulation, the sub-model $\mathcal{H}_{1,1}$ can be written as

$$\Delta^{2d_2} X_t = \alpha\beta' L_{2d_2} X_t + \varepsilon_t \tag{5}$$

where d_2 is the fractional parameter under $\mathcal{H}_{1,1}$. From (5) it emerges that also $\mathcal{H}_{1,1}$ is equivalent to $\mathcal{H}_0$. This means that the two sub-models, $\mathcal{H}_{1,0}$ and $\mathcal{H}_{1,1}$, are equivalent, with $d_1 = 2d_2$. The fractional order of the system is the same in both cases, i.e. $\mathcal{F}(d_1) = \mathcal{F}(2d_2)$. Hence, under $\mathcal{H}_{1,0}$ the process X_t has the same fractional order as under $\mathcal{H}_{1,1}$, but the latter is represented by an integer multiple of the parameter d_2. In the example above, the identification condition is clearly violated, as the conditional densities of $\mathcal{H}_{1,0}$ and $\mathcal{H}_{1,1}$ are

$$p(X_1, ..., X_T, \theta_1 | X_0, X_{-1}, \ldots) = p(X_1, ..., X_T, \theta_2 | X_0, X_{-1}, \ldots) \tag{6}$$

where $\theta_1 = vec(d_1, \alpha, \beta, \Gamma_1^{(1)}, \Omega)$ and $\theta_2 = vec(\frac{1}{2}d_1, \alpha, \beta, \Gamma_1^{(2)}, \Omega)$ with $\Gamma_1^{(1)} = 0$ and $\Gamma_1^{(2)} = I_p + \alpha\beta'$.

The identification problem outlined above has practical consequences when the model with k_0 lags and a FCVAR$_{d,b}$ model with k lags are considered. Suppose that the $\mathcal{H}_{k_0}$ model is

$$\mathcal{H}_{k_0}: \quad \Delta^{d_0} X_t = \alpha_0 \beta_0' \Delta^{d_0 - b_0} L_{b_0} X_t + \sum_{i=0}^{k_0} \Gamma_i^0 \Delta^{d_0} L_{b_0}^i X_t + \varepsilon_t \quad \varepsilon_t \sim N(0, \Omega_0) \tag{7}$$

with k_0 lags, $|\alpha_{0,\perp}' \Gamma^0 \beta_{0,\perp}| \neq 0$ where $\Gamma^0 = I_p - \sum_{i=1}^{k_0} \Gamma_i^0$. When a model $\mathcal{H}_k$ with $k > k_0$ is considered for X_t, then $\mathcal{H}_{k_0}$ corresponds to the set of restrictions $\Gamma_{k_0+1} = \Gamma_{k_0+2} = ... = \Gamma_k = 0$ imposed on $\mathcal{H}_k$. As shown in the example above, there are several equivalent sub-models to that under the restriction $\Gamma_{k_0+1} = \Gamma_{k_0+2} = ... = \Gamma_k = 0$.

Therefore, the aim of this Section is to study the number and the nature of these equivalent sub-models, in order to provide a necessary and sufficient condition to identify the fractional parameters d_0 and b_0 as the parameters d and b of the model $\mathcal{H}_k$.

The following Proposition states the necessary and sufficient condition, called the $\mathcal{F}(d)$ condition, for identification of the parameters of the model $\mathcal{H}_{k_0}$.

Proposition 2.2 *For any $k_0 \geq 0$ and $k \geq k_0$, the $F(d)$ condition, $|\alpha_\perp' \Gamma \beta_\perp| \neq 0$, where $\Gamma = I_p - \sum_{i=1}^{k} \Gamma_i$, is a necessary and sufficient condition for the identification of the set of parameters of $\mathcal{H}_{k_0}$ in equation (7).*

Corollary 2.3

i) Given k_0 and k, with $k \geq k_0$, the number of equivalent sub-models that can be obtained from $\mathcal{H}_k$ is $m = \lfloor \frac{k+1}{k_0+1} \rfloor$, where $\lfloor x \rfloor$ denotes the greatest integer less or equal to x.

ii) For any $k \geq k_0$, all the equivalent sub-models are found for parameter values $d_j = d_0 - \frac{j}{j+1} b_0$ and $b_j = b_0/(j+1)$ for $j = 0, 1, ..., m-1$.

The Proposition 2.2 has important consequences. First, the condition $|\alpha_\perp' \Gamma \beta_\perp| \neq 0$ holds only for the sub-model for which $d = d_0$ and $b = b_0$, i.e. for the sub-model corresponding to the restrictions $\Gamma_{k_0+1} = \Gamma_{k_0+2} = ... = \Gamma_k = 0$. In the example above, the $\mathcal{F}(d)$ condition is verified only for $\mathcal{H}_{1,0}$, while $|\alpha_\perp' \Gamma \beta_\perp| = 0$ for $\mathcal{H}_{1,1}$, since $\Gamma = I_p - (I_p + \alpha\beta') = -\alpha\beta'$. Second, for $k \gg k_0$, the $(m-1)$-th sub-model is such that $d_{m-1} \approx d_0 - b_0$ and $b_{m-1} \approx 0$, i.e. located close to the boundary of the parameter space. As a consequence of corollary i), there are cases for which $k > k_0$ doesn't imply lack of identification. For example, when $k = 2$ and $k_0 = 1$ there are no sub-models of $\mathcal{H}_2$ that are equivalent to the one in correspondence of $d = d_0$, $b = b_0$, $\Gamma_1 = \Gamma_1^0$ and $\Gamma_2 = 0$.

Table 1 summarizes the number of equivalent sub-models for different values of k_0 and k. When k_0 is small there are several equivalent sub-models for small choices of k. When k_0 increases, multiple equivalent sub-models are obtained only for large k. For example, when $k_0 = 5$, then two equivalent sub-models are obtained only from suitable restrictions of the $\mathcal{H}_{11}$ model.

$k_0 \downarrow k \rightarrow$	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
0	1	2	3	4	5	6	7	8	9	10	11	12	13
1	–	1	1	2	2	3	3	4	4	5	5	6	6
2	–	–	1	1	1	2	2	2	3	3	3	4	4
3	–	–	–	1	1	1	1	2	2	2	2	3	3
4	–	–	–	–	1	1	1	1	1	2	2	2	2
5	–	–	–	–	–	1	1	1	1	1	1	2	2

Table 1: Table reports the number of equivalent models (m) for different combinations of k and k_0. When $k_0 > k$ the $\mathcal{H}_k$ is under-specified.

The next Section discusses the consequences of the lack of identification on the estimation of the FCVAR$_{d,b}$ parameters when the true number of lags is unknown.

3 Identification and Inference

This Section illustrates, by means of numerical examples, the problems in the estimation of the parameters of the FCVAR$_{d,b}$ that are induced by the lack of identification outlined in Section 2. In particular, the $\mathcal{F}(d)$ condition can be used to correctly identify the fractional parameters d and b when model $\mathcal{H}_k$ is fitted on the data.

As shown in Johansen and Nielsen (2012), the parameters of the FCVAR$_{d,b}$ can be estimated in two steps. First, the parameters $\hat{d}$ and $\hat{b}$ are obtained by maximizing the profile log-likelihood

$$\ell_T(\psi) = -\log\det(S_{00}(\psi)) - \sum_{i=1}^{r} \log(1 - \lambda_i(\psi)) \tag{8}$$

where $\psi = (d, b)'$. $\lambda(\psi)$ and $S_{00}(\psi)$ are obtained from the residuals, $R_{it}(\psi)$ for $i = 1, 2,$

of the reduced rank regression of $\Delta^d Y_t$ on $\sum_{j=1}^k \Delta^d L_b^j Y_t$ and $\Delta^{d-b} L_b Y_t$ on $\sum_{j=1}^k \Delta^d L_b^j Y_t$, respectively. The moment matrices $S_{ij}(\psi)$ for $i,j=1,2$ are $S_{ij}(\psi) = T^{-1}\sum_{t=1}^T \sum_{t=1}^T R_{it}(\psi)R'_{jt}(\psi)$ and $\lambda_i(\psi)$ for $i=1,\dots,p$ are the solutions, sorted in decreasing order, of the generalized eigenvalue problem

$$\det\left[\lambda S_{11}(\psi) - S_{10}(\psi)S_{00}^{-1}(\psi)S_{01}(\psi)\right]. \tag{9}$$

Second, given $\hat{d}$ and $\hat{b}$, the estimates $\hat{\alpha}$, $\hat{\beta}$, $\hat{\Gamma}_j$ for $j=1,\dots,k$ and $\hat{\Omega}$ are found by reduced rank regression.

The values of ψ that maximize $\ell_T(\psi)$ must be found numerically. Therefore, we explore, by means of Monte Carlo simulations, the effect of the lack of identification of the FCVAR$_{d,b}$ model on the expected profile likelihood when $k > k_0$. Since the asymptotic value of $\ell_T(\psi)$ is not a closed-form function of the model parameters, we approximate the asymptotic behaviour of $\ell_T(\psi)$ by averaging over S simulations, setting a large number T of observations. This provides an estimate of the expected profile likelihood, $E[\ell_T(\psi)]$. Therefore, we generate $S=100$ times from model (7) with $T=50,000$ observations and different choices of k_0 and $p=2$. The fractional parameters of the system are $d_0=0.8$ and b_0 is set equal to d_0 in order to simplify the readability of the results without loss of generality. The cointegration vectors are $\alpha=[0.5,-0.5]$ and $\beta=[1,-1]$, and the matrices Γ_i^0 for $i=1,...,k_0$ are chosen such that the roots of the characteristic polynomial are outside the fractional circle, see Johansen (2008). [3] The average profile log-likelihood, $\bar{\ell}_T(\psi)$, and the average $\mathcal{F}(d)$ condition, $\bar{\mathcal{F}}(d)$, are then computed with respect to a grid of alternative values for $d=[d_{\min},\dots,d_{\max}]$.[4]

Figure 1 reports the values of $\bar{\ell}_T(d)$ and $\bar{\mathcal{F}}(d)$ when $k=1$ lags are chosen but $k_0=0$. It clearly emerges that two equally likely sub-models are found corresponding to $d=0.4$ and $d=0.8$. However, $\bar{\mathcal{F}}(d)$ is equal to zero when $d=0.4$. Consistently with the theoretical results presented in Section 2, the other value of the parameter d that maximizes $\bar{\ell}_T(d)$ is found around $d=0.8$, where $\bar{\mathcal{F}}(d)$ is far from zero. Similarly, as reported in Figure 2, when $k=2$ and $k_0=0$, the likelihood function presents three humps around $d=0.8$,

[3]The purpose of these simulations is purely illustrative, so that we do not explore the behaviour of $E[\ell_T(\psi)]$ for other parameter values. All the source codes are available upon request from the authors.

[4]The values of $d_{\min}$ and $d_{\max}$ presented in the graphs change in order to improve the clearness of the plots.

$d = 0.4$ and $d = 0.2667 = d_0/3$. As in the previous case, the estimates corresponding to $d = 0.4$ and $d = 0.2667$ should be discarded due to the nullity of the $\mathcal{F}(d)$ condition.

A slightly more complex evidence arises when $k_0 > 0$. Figures 3 and 4 report $\bar{\ell}_T(d)$ and $\bar{\mathcal{F}}(d)$ when $k_0 = 1$ while $k = 2$ and $k = 3$ are chosen. When $k = 2$ the $\bar{\ell}_T(d)$ function has a single large hump in the region of $d = 0.8$, thus supporting the theoretical results outlined above, i.e. when $k = 2$ and $k_0 = 1$ there is no lack of identification. However, another interesting evidence emerges. The $\bar{l}_T(d)$ function is flat and high in the region around $d = 0.5$. This may produce identification problems in finite samples. This issue will be further discussed in Section 3.1. When $k = 3$ we expect $m = \frac{4}{2} = 2$ equivalent sub-models in correspondence of $d = d_0 = 0.8$ and $d = d_0/2 = 0.4$. Indeed, in Figure 4, the line $\bar{\ell}_T(d)$ has two humps around the values of $d = 0.4$ and $d = 0.8$. As expected, in the region around $d = 0.4$ the average $\mathcal{F}(d)$ condition is near 0.

3.1 Identification in Finite Sample

In Section 2, the mathematical identification issues of the $\text{FCVAR}_{d,b}$ have been discussed in theory. The purpose of this Section is to shed some light on the consequences of the lack of mathematical identification in finite samples. From the analysis above, we know that the expected profile likelihood displays multiple equivalent maxima in correspondence of fractions of d_0 for some $k > k_0$.

In finite samples, however, the profile likelihood function displays multiple humps, but just one global maximum when k is larger than k_0. Figure 5 reports the finite sample profile likelihood function, $\ell_T(d)$, of model $\mathcal{H}_1$ obtained from two simulated paths of (7) with $k_0 = 0$ with $T = 1000$. The plots highlight the behaviour of $\ell_T(d)$ and the consequences of the lack of identification, since the global maximum of $\ell_T(d)$ is around $d = 0.4$ in Panel a), while it is around 0.8 in Panel b).

Moreover, the lag structure of the $\text{FCVAR}_{d,b}$ model induces poor finite sample identification, namely *weak* identification, also for those cases in which mathematical identification is expected. For example, as shown in Figure 3, when $k_0 = 1$ and $k = 2$ the average profile likelihood is high in a neighbourhood of $d = 0.5$, even though in theory there is no sub-model equivalent to the one corresponding to $d = 0.8$. The problem worsens if we look at the profile likelihoods, $\ell_T(d)$, for a given $T = 1000$. As in Figure 5, Figure 6 reports the shape of the finite sample profile likelihood function, $\ell_T(d)$, relative to two simulated paths of (7) when $k_0 = 1$ and $\mathcal{H}_2$ is estimated. When the global maximum is in a neighbourhood of $d = 0.4$, Panel a), the $\mathcal{F}(d)$ is close to zero, thus suggesting that the estimated matrix $\hat{\Gamma}_1$ and $\hat{\Gamma}_2$ are such that $|\alpha'_\perp \Gamma \beta_\perp| = 0$. This evidence suggests that, in empirical applications, it is crucial to evaluate the $\mathcal{F}(d)$ condition when selecting the optimal lag length.

4 Lag selection and the $\mathcal{F}(d)$ condition

In practical applications the true number of lags is unknown. Commonly, the lag selection in the VECM framework is carried out following a *general-to-specific* approach. Starting from a large value of k, the optimal lag length is chosen by a sequence of likelihood-ratio tests for the hypothesis $\Gamma_k = 0$, until the nullity of the matrix Γ_k is rejected. At each step of this iteration, the profile likelihood function $\ell_T(d)$ must be computed. If k is larger than k_0, then there is a non-zero probability that the maximum of $\ell(d)_T$ will be found in a neighborhood of the values of d, that are fractions of d_0 and for which $|\alpha_\perp \Gamma \beta_\perp| = 0$. For example, similarly to the evidence shown in Panel a) of Figure 5, it may happen that when $k = 1$, $\max \ell(d)_T$ is found in a region near $d = 0.4$, when $d_0 = 0.8$ and $k_0 = 0$. If the likelihood ratio test

$$LR = 2 \cdot \left[\ell(\hat{d}_T^{(k=1)}) - \ell(\hat{d}_T^{(k=0)}) \right] \tag{10}$$

rejects the null hypothesis, then the set of parameters which maximizes the likelihood in this case will correspond to $\theta = (d_0/2, b_0/2, \alpha, \beta, \Gamma_1 = I_p + \alpha\beta')$.

In order to avoid this inconvenience, we suggest to integrate the top-down approach for the selection of the lags with an evaluation of the $\mathcal{F}(d)$ condition. Since the value of $|\hat{\alpha}'_\perp \hat{\Gamma} \hat{\beta}_\perp|$ is a point estimate, it is required to compute confidence bands around its value in order to evaluate if it is statistically different from zero. Therefore, we rely on a bootstrap approach to evaluate the nullity of $|\hat{\alpha}'_\perp \hat{\Gamma} \hat{\beta}_\perp|$. The suggested algorithm for the lag selection in the FCVAR model is

1. Evaluate $L_k = \max \ell_T(d)$ for the FCVAR for a given large k;

2. Evaluate $L_{k-1} = \max \ell_T(d)$ for the FCVAR with $k-1$;

3. Compute the value of the LR test (10) for k and $k-1$, which is distributed as $\chi^2(p^2)$ where p^2 are the degrees of freedom, see Johansen and Nielsen (2012).

4. Iterate points 2. and 3. until the null hypothesis is rejected, in $\tilde{k}$.

5. Evaluate the $\mathcal{F}(d)$ condition in $\hat{d}_{\tilde{k}}, \hat{b}_{\tilde{k}}$, i.e. $|\hat{\alpha}'_{\perp,\tilde{k}} \hat{\Gamma}_{\tilde{k}} \hat{\beta}_{\perp,\tilde{k}}|$, namely $\mathcal{F}(\hat{d}_{\tilde{k}})$.

6. Generate S pseudo trajectories from the re-sampled residuals of the $\mathcal{H}_{\tilde{k}}$ model.

7. For fixed $\hat{d}_{\tilde{k}}, \hat{b}_{\tilde{k}}$, estimate the matrices $\hat{\alpha}^s_{\perp,\tilde{k}}$, $\hat{\beta}^s_{\perp,\tilde{k}}$ and $\hat{\Gamma}^s_1, .., \hat{\Gamma}^s_{\tilde{k}}$ with reduced rank regression, for $s = 1, ..., S$.

8. Compute the $\mathcal{F}^s(d)$ condition, for $s = 1, ..., S$.

9. Compute the quantiles, q_α and $q_{1-\alpha}$, of the empirical distribution of $\mathcal{F}^s(d)$.

10. If both $\mathcal{F}(\hat{d}_{\tilde{k}})$ and 0 belong to the bootstrapped confidence interval, then iterate 1.-9. for $\tilde{k} - i$, for $i = 1, ..., \tilde{k}$ until the LR test rejects the null and $\mathcal{F}(\hat{d}_{\tilde{k}-i})$ is statistically different from zero.

Table 2 reports the results on the performance of the lag selection procedure that exploits the information on the $\mathcal{F}(d)$ condition to infer the correct number of lags. The lag selection method follows the procedure outlined above, starting from $k = 10$ lags. It clearly emerges that in more than 95% of the cases the true number of lags is selected, thus avoiding the identification problems discussed in Section 2. A different evidence emerges from Table 3. The selection procedure based only on likelihood ratio tests is not robust to the identification problem and it has a much lower coverage probability. Indeed, only in 50% of the cases the correct lag length is selected with 500 observations. As expected, the performance slightly improves when $T = 1000$ and the percentage of correctly specified models increases to 65% of the cases.

$k_0 \downarrow k \rightarrow$	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
	T=500										
0	97	0	1	0	0	0	0	0	1	0	1
1	1	96	0	0	0	1	0	0	1	0	1
2	0	2	91	2	1	2	0	0	1	1	0
	T=1000										
0	93	3	0	1	1	0	0	0	0	2	0
1	0	95	2	1	1	0	0	0	1	0	0
2	0	0	96	2	0	0	1	0	0	0	1

Table 2: Table reports the percentage coverage probabilities in which a specific lag length k is selected using the $\mathcal{F}(d)$ condition together with the LR test. The reported results are based on 100 generated paths from the $\mathcal{H}_{k_0}$ model with $k_0 = 0, 1, 2$ and $T = 500$ and $T = 1000$ observations. The bootstrapped confidence intervals for the $\mathcal{F}(d)$ condition are based on $S = 200$ draws.

$k_0 \downarrow k \rightarrow$	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
	T=500										
0	56	0	0	0	0	0	3	6	9	13	13
1	0	57	0	0	0	2	2	7	10	11	11
2	0	0	46	1	2	4	4	9	11	10	13
	T=1000										
0	64	3	1	1	1	4	4	1	7	7	7
1	0	67	2	2	1	3	2	3	6	8	6
2	0	0	69	2	3	2	3	1	3	10	7

Table 3: Table reports the percentage coverage probabilities in which a specific lag length k is selected with a general-to-specific approach using a sequence of LR tests. The reported results are based on 100 generated paths from the $\mathcal{H}_{k_0}$ model with $k_0 = 0, 1, 2$ and $T = 500$ and $T = 1000$ observations.

5 Unknown cointegration rank

This Section extends the previous results to the case of unknown rank, r, which is of relevance in empirical applications. The FCVAR$_{d,b}$ model with cointegration rank $0 \leq r \leq p$ is defined as:

$$\mathcal{H}_{r,k}: \quad \Delta^d X_t = \Pi \Delta^{d-b} L_b X_t + \sum_{i=1}^{k} \Gamma_i \Delta^d L_b^i X_t + \varepsilon_t$$

where r is the rank of the $p \times p$ matrix Π.

Compared to the case discussed in previous sections, model $\mathcal{H}_{r,k}$ exhibits further identification issues. For example, the model with $k = 1$ lag and rank $0 \leq r \leq p$, is given by

$$\mathcal{H}_{r,1}: \quad \Delta^d X_t = \Pi \Delta^{d-b} L_b X_t + \Gamma_1 \Delta^d L_b X_t + \varepsilon_t \tag{11}$$

where the parameters $\theta = (d, b, \Pi, \Gamma_1)$. Consider the following two sub-models

$$\mathcal{H}_{p,0}: \quad \Delta^{d_1} X_t = \Pi^1 \Delta^{d_1 - b_1} L_{b_1} X_t + \varepsilon_t \tag{12}$$

and

$$\mathcal{H}_{0,1}: \quad \Delta^{d_2} X_t = \Gamma_1^2 \Delta^{d_2} L_{b_2} X_t + \varepsilon_t \tag{13}$$

The sub-model $\mathcal{H}_{p,0}$ is a reparameterization of $\mathcal{H}_{0,1}$ because (12) can be written as

$$\left[\Delta^{d_1 - b_1}(-\Pi^1) + \Delta^{d_1}(I_p + \Pi^1)\right] X_t = \varepsilon_t \tag{14}$$

and (13) is given by

$$\left[\Delta^{d_2}(I - \Gamma_1^2) + \Delta^{d_2 + b_2}(\Gamma_1^2)\right] X_t = \varepsilon_t \tag{15}$$

If $I - \Gamma_1^2 = \Pi^1$, $d_1 = d_2 + b_2$, $b_1 = b_2$ the two sub-models represent the same process and $d_1 \geq b_1 > 0$ implies $d_2 + b_2 > b_2$. Hence, the probability densities

$$p(X_1, \ldots, X_T; \theta_1 | X_{-1}, \ldots) = p(X_1, \ldots, X_T; \theta_2 | X_{-1}, \ldots)$$

when

$$\theta_1 = (d_1, b_1, \Pi^1, 0) \quad \theta_2 = (d_2 + b_2, b_2, 0, I - \Pi^1)$$

However, the sub-model $\mathcal{H}_{0,1}$ is not always a reparameterization of $\mathcal{H}_{p,0}$. In fact, given the expansions in (14) and (15), it follows that

$$p(X_1, \ldots, X_T; \theta_3 | X_{-1}, X_{-2}, \ldots) = p(X_1, \ldots, X_T; \theta_4 | X_{-1}, X_{-2}, \ldots) \tag{16}$$

where

$$\theta_3 = (d_2, b_2, 0, \Gamma_1^2) \quad \theta_4 = (d_2 - b_2, b_2, I - \Gamma_1^2, 0)$$

The equality (16) holds if and only if θ_4 is such that $d_2 - b_2 \geq b_2 > 0$. This implies that $\mathcal{H}_{0,1} = \mathcal{H}_{p,0} \cap \{d \geq 2b\}$. Hence, the nesting structure $\mathcal{H}_{0,1} \subset \mathcal{H}_{p,0}$ follows.

Next proposition extends this example for a general number of lags k and rank r.[5]

Proposition 5.1 *Consider the FCVAR$_{d,b}$, $\mathcal{H}_{r,k}$, with $k > 0$ and $0 \leq r \leq p$. The following propositions hold:*

- *For any $k > 0$, model $\mathcal{H}_{0,k}$ is equivalent to $\mathcal{H}_{p,k-1}$, if the restriction $d > 2b$ holds.*
- *For any $k > 0$, model $\mathcal{H}_{0,k}$ is equivalent to $\mathcal{H}_{r,k-1}$ with $0 < r < p$, if and only if $|\alpha'_{\perp}\Gamma\beta_{\perp}| = 0$ in the latter.*
- *The nesting structure of the FCVAR$_{d,b}$ model is represented by the following scheme:*

$$
\begin{array}{ccccccccc}
\mathcal{H}_{0,0} & \subset & \mathcal{H}_{0,1} & \subset & \mathcal{H}_{0,2} & \subset & \cdots & \subset & \mathcal{H}_{0,k} \\
\cap & & \cap & & \cap & & & & \cap \\
\mathcal{H}_{1,0} & \subset & \mathcal{H}_{1,1} & \subset & \mathcal{H}_{1,2} & \subset & \cdots & \subset & \mathcal{H}_{1,k} \\
\cap & & \cap & & \cap & & & & \cap \\
\vdots & & \vdots & & \vdots & & \ddots & & \vdots \\
\cap & & \cap & & \cap & & & & \cap \\
\mathcal{H}_{p,0} & \subset & \mathcal{H}_{p,1} & \subset & \mathcal{H}_{p,2} & \subset & \cdots & \subset & \mathcal{H}_{p,k}
\end{array}
\quad \textit{with} \quad
\begin{array}{c}
\mathcal{H}_{0,1} \subset \mathcal{H}_{p,0} \\
\mathcal{H}_{0,2} \subset \mathcal{H}_{p,1} \\
\vdots \\
\vdots \\
\mathcal{H}_{0,k} \subset \mathcal{H}_{p,k-1}
\end{array}
$$

Clearly, the nesting structure of the FCVAR impacts on the joint selection of the number of lags and the cointegration rank. Indeed, the likelihood ratio statistic for

[5] A similar identification problem arises in the FAR(k) in Johansen and Nielsen (2010)

$$\Delta^d X_t = \pi \Delta^{d-b} L_b X_t + \sum_{i=1}^{k} \gamma_i \Delta^d L_b^i X_t + \varepsilon_t$$

Similarly to the FCVAR$_{d,b}$ model, the FAR(k) has the following nesting structure:

$$
\begin{array}{ccccccccc}
\mathcal{H}_{0,0} & \subset & \mathcal{H}_{0,1} & \subset & \mathcal{H}_{0,2} & \subset & \cdots & \subset & \mathcal{H}_{0,k} \\
\cap & & \cap & & \cap & & & & \cap \\
\mathcal{H}_{1,0} & \subset & \mathcal{H}_{1,1} & \subset & \mathcal{H}_{1,2} & \subset & \cdots & \subset & \mathcal{H}_{1,k}
\end{array}
\quad \text{with} \quad
\begin{array}{c}
\mathcal{H}_{0,1} \subset \mathcal{H}_{1,0} \\
\mathcal{H}_{0,2} \subset \mathcal{H}_{1,1} \\
\vdots \\
\vdots \\
\mathcal{H}_{0,k} \subset \mathcal{H}_{1,k-1}
\end{array}
$$

cointegration rank r, $LR_{r,k}$, is given by

$$-2\log LR_{r,k}(\mathcal{H}_{r,k}|\mathcal{H}_{p,k}) = T\cdot(\ell_{r,k}(\hat{d}_r,\hat{b}_r) - \ell_{p,k}(\hat{d}_p,\hat{b}_p))$$

where $\ell_{r,k}$ is the profile likelihood of the $\text{FCVAR}_{d,b}$ model with rank r and k lags. Analogously, $\hat{d}_{r,k}$ and $\hat{b}_{r,k}$ are the arguments that maximize $\ell_{r,k}$. The asymptotic properties of the LR_r statistics for given k are provided in Johansen and Nielsen (2012).

Under the null hypothesis $\mathcal{H}_{0,k}$, it follows from Proposition 5.1 that the LR tests $LR_{0,k} = -2\log LR(\mathcal{H}_{0,k}|\mathcal{H}_{p,k})$ is equal to $LR_{p,k-1} = -2\log LR(\mathcal{H}_{p,k-1}|\mathcal{H}_{p,k})$.[6] Hence, the equality of the test statistics $LR_{0,k}$ and $LR_{p,k-1}$ influences the top-down sequence of tests for the joint identification of the cointegration rank and the lag length. Indeed, assuming that the top-down procedure for the optimal lag selection terminates in $\mathcal{H}_{p,k-1}$, then it would impossible to test whether the optimal model is $\mathcal{H}_{p,k-1}$ or $\mathcal{H}_{0,k}$. Therefore a problem of joint selection of k and $r > 0$ arises in the $\text{FCVAR}_{d,b}$ when rank is potentially equal to 0 or p. A trivial solution to this issue is to exclude the models with rank equal to zero when selecting rank and lag length.

6 Conclusion

This paper discussed in detail the identification problem in the $\text{CFVAR}_{d,b}$ model of Johansen (2008) such that the fractional order of the system cannot be uniquely determined when the lag structure is over-specified. In particular, the multiplicity of equivalent sub-models is provided in closed form given k and k_0. It is also shown that a necessary and sufficient condition for the identification is that the $\mathcal{F}(d)$ condition, i.e. $|\alpha'_\perp\Gamma\beta_\perp| \neq 0$, is fulfilled. A simulation study highlights the practical problem of multiple humps in the expected profile log-likelihood function as a consequence of the identification problem and the over-specification of the lag structure. The simulations also show that the true parameters can be detected by evaluating the $\mathcal{F}(d)$ condition. The simulation study also reveals a problem of weak identification, characterized by the presence of local and global maxima of the profile likelihood function in finite samples. It is also shown that the

[6]Both tests, $LR_{p,k-1}$ and $LR_{0,k}$ are asymptotically $\chi^2(p^2)$ distributed.

$\mathcal{F}(d)$ condition is necessary and sufficient for identification also when the cointegration rank is unknown and such that $0 < r < p$. It is proved that model $\mathcal{H}_{0,k}$ is equivalent to model $\mathcal{H}_{p,k-1}$ under certain conditions on d and b, but the $\mathcal{F}(d)$ does not provide any information for the identification in this case. A solution to this issue, which does not exclude rank equal to zero, is left for future research.

References

Christensen, B. J. and Nielsen, M. O. (2006). Semiparametric analysis of stationary fractional cointegration and the implied-realized volatility relation. *Journal of Econometrics*.

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

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

Johansen, S. (2008). A representation theory for a class of vector autoregressive models for fractional processes. *Econometric Theory*, Vol 24, 3:651–676.

Johansen, S. and Nielsen, M. Ø. (2012). Likelihood inference for a fractionally cointegrated vector autoregressive model. *Econometrica*, 80(6):2667–2732.

Johansen, S. and Nielsen, M. r. (2010). Likelihood inference for a nonstationary fractional autoregressive model. *Journal of Econometrics*, 158(1):51–66.

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

Robinson, P. M. and Marinucci, D. (2003). Semiparametric frequency domain analysis of fractional cointegration. In Robinson, P. M., editor, *Time Series with Long Memory*, pages 334–373. Oxford University Press.

Tschernig, R., Weber, E., and Weigand, R. (2013a). Fractionally integrated var models with a fractional lag operator and deterministic trends: Finite sample identification and two-step estimation. Technical report.

Tschernig, R., Weber, E., and Weigand, R. (2013b). Long-run identification in a fractionally integrated system. *Journal of Business & Economic Statistics*, 31(4):438–450.

Proof of Proposition 2.2 when $k_0 = 0$ and $k = 1$

Let us define the FCVAR$_{d,b}$ model with one lag, $\mathcal{H}_1$, as

$$\Delta^d X_t = \alpha\beta' \Delta^{d-b} L_b X_t + \Gamma_1 \Delta^d L_b X_t + \varepsilon_t \tag{17}$$

which can be written as

$$\left\{\Delta^d \left[I + \alpha\beta' - \Gamma_1\right] + \Delta^{d-b}\left[-\alpha\beta'\right] + \Delta^{d+b}\Gamma_1\right\} X_t = \varepsilon_t \tag{18}$$

Similarly, the model $\mathcal{H}_{k_0}$ with $k_0 = 0$ lags in (7) can be rewritten as

$$\left\{\Delta^{d_0}[I + \alpha_0\beta_0'] + \Delta^{d_0-b_0}[-\alpha_0\beta_0']\right\} X_t = \varepsilon_t \tag{19}$$

Imposing $I + \alpha\beta' - \Gamma_1 = 0$, it follows that

$$\Delta^{d+b}\Gamma_1 = (I + \alpha_0\beta_0^{'})\Delta^{d_0} \tag{20}$$

and the condition

$$-\alpha\beta'\Delta^{d-b} = -\alpha_0\beta_0^{'}\Delta^{d_0-b_0} \tag{21}$$

it is satisfied when $d = d_0 - b_0/2$ and $b = b_0/2$. The other equivalent sub-model corresponding to $\mathcal{H}_{k_0}$ in (7) with $k_0 = 0$ with $\alpha_0\beta_0^{'} = \alpha\beta'$, $\Gamma_1 = 0$, $d = d_0$ and $b = b_0$.

When the $\mathcal{H}_{k_0}$ has $k_0 = 0$ lags and model $\mathcal{H}_k$ with $k > 0$ is considered, then the $k-$th model can be rewritten as

$$\sum_{i=-1}^{k} \Psi_i \Delta^{d+ib} X_t = \varepsilon_t \tag{22}$$

where $\sum_{i=-1}^{k} \Psi_i = I_p$, $\Psi_{-1} = -\alpha\beta'$ and $\Psi_0 = \alpha\beta' + \Gamma$.

Similarly, the $\mathcal{H}_{k_0}$ model with k_0 lags is given by

$$\sum_{i=-1}^{0} \Psi_{i,0}\Delta^{d_0+ib_0} = \varepsilon_t, \quad \text{with} \quad \Psi_{-1,0} + \Psi_{0,0} = I_p. \tag{23}$$

It is possible to show, that $k + 1$ sub-models equivalent to $\mathcal{H}_{k_0}$ can be obtained imposing suitable restrictions on the matrices Ψ_i $i = -1, ..., k$ of the model $\mathcal{H}_k$. The equivalent sub-models, $\mathcal{H}_{k,j}$, $j = 0, 1, \ldots, k$, are found in correspondence of

$$\begin{aligned} \Psi_{-1} &= \Psi_{-1,0} \quad \text{corresponding to} \quad d - b = d_0 - b_0 \\ \Psi_j &= \Psi_{0,0} \quad \text{corresponding to} \quad d + jb = d_0 \\ \Psi_s &= 0, \quad s \neq j \end{aligned} \tag{24}$$

This system entails that all sub-models $\mathcal{H}_{k,j}$, $j = 1, \ldots, k$ are such that $\Psi_{-1} = -\alpha\beta' = -\alpha_0\beta_0' = \Psi_{-1,0}$ and $\Psi_0 = 0$. This implies that $\alpha\beta' + \Gamma = \Psi_0 = 0$. Hence, the sub-models for $j = 1, ..., k$ are such that $|\alpha_\perp' \Gamma \beta_\perp| = 0$. Only for $j = 0$, the condition $|\alpha_\perp' \Gamma \beta_\perp| \neq 0$ is satisfied.

Hence, verifying $|\alpha_\perp' \Gamma \beta_\perp| \neq 0$ is sufficient for the identification of the parameters of $\mathcal{H}_{k_0}$.■

The following Table summarizes the set of restrictions that have to be imposed on a $\mathcal{H}_k$ model in order to find $k + 1$ sub-models equivalent to the model $\mathcal{H}_{k_0}$ with $k_0 = 0$:

Matrices in $\mathcal{H}_k$ →	Ψ_{-1}	Ψ_0	Ψ_1	Ψ_2	Ψ_3	$\cdots$	Ψ_k
$\mathcal{H}_{k,0}$	$\Psi_{-1,0}$	$\Psi_{0,0}$	0	0	0	$\cdots$	0
$\mathcal{H}_{k,1}$	$\Psi_{-1,0}$	0	$\Psi_{0,0}$	0	0	$\cdots$	0
$\mathcal{H}_{k,2}$	$\Psi_{-1,0}$	0	0	$\Psi_{0,0}$	0	$\cdots$	0
$\mathcal{H}_{k,3}$	$\Psi_{-1,0}$	0	0	0	$\Psi_{0,0}$	$\cdots$	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$\mathcal{H}_{k,k}$	$\Psi_{-1,0}$	0	0	0	0	$\cdots$	$\Psi_{0,0}$

Table 4: Restrictions imposed on the $\mathcal{H}_k$ model when the model $\mathcal{H}_{k_0}$ is a FCVAR$_{d,b}$ with $k_0 = 0$ lags.

Proof of Proposition 2.2

Let us define the model $\mathcal{H}_{k_0}$ under $k_0 \geq 0$ as

$$\sum_{i=-1}^{k_0} \Psi_{i,0}\Delta^{d_0+ib_0}X_t = \varepsilon_t \tag{25}$$

and the model $\mathcal{H}_k$ with $k > k_0$ as

$$\sum_{i=-1}^{k} \Psi_i\Delta^{d+ib}X_t = \varepsilon_t \tag{26}$$

It is possible to show, that, for a given k_0, m sub-models equivalent to the DGP (25) can be obtained imposing suitable restrictions on the matrices Ψ_i $i = -1, ..., k$ of the model $\mathcal{H}_k$. The equivalent sub-models, $\mathcal{H}_{k,j}$, $j = 0, 1, \ldots, m-1$, are found in correspondence of

$$\begin{aligned}
\Psi_{-1} &= \Psi_{-1,0} \quad \text{corresponding to} \quad d - b = d_0 - b_0 \qquad (27)\\
\Psi_{(\ell+1)(j+1)-1} &= \Psi_{\ell,0} \quad \text{corresponding to} \quad d + [(\ell+1)(j+1)-1]b = d_0 + \ell b_0,\\
& \quad \text{for} \quad \ell = 0, \ldots, k_0 \quad j = 0, 1, \ldots, m-1\\
\Psi_s &= 0 \quad \text{for} \quad s \neq (\ell+1)(j+1)-1,\\
& \quad \text{and} \quad \ell = 0, \ldots, k_0 \quad j = 0, 1, \ldots, m-1.
\end{aligned}$$

The restriction $\Psi_0 = \alpha\beta' + \Gamma = 0$, implying $|\alpha_\perp'\Gamma\beta_\perp| = 0$ with $\Gamma = I - \sum_{i=1}^{k}\Gamma_i$, is always imposed for the sub-models $\mathcal{H}_{k,j}$ when $j \geq 1$.

As in the case $k_0 = 0$, $\Psi_{-1,0} = -\alpha_0\beta_0'$ and $\Psi_{-1} = -\alpha\beta$ load the terms $\Delta^{d_0-b_0}X_t$ and $\Delta^{d-b}X_t$ respectively. This implies that $d_0 - b_0 = d - b$. For a given $j > 0$, a system of $k_0 + 2$ equations (27) in d and b is derived from the restrictions on the matrices Ψ_i. The solution of this system is found for $b = b_0/(j+1)$ and $d = d_0 - \frac{j}{j+1}b_0$. All sub-models $\mathcal{H}_{k,j}$, $j = 1, \ldots, k$ are such that $\Psi_{-1} = -\alpha\beta' = -\alpha_0\beta_0' = \Psi_{-1,0}$ and $\Psi_0 = 0$. This implies that $\alpha\beta' + \Gamma = \Psi_0 = 0$. Hence, the sub-models for $j = 1, ..., k$ are such that $|\alpha_\perp'\Gamma\beta_\perp| = 0$. Only for $j = 0$, the condition $\alpha_\perp'\Gamma\beta_\perp| \neq 0$ is satisfied.

For a given $k > k_0$, the number of possible restriction to be imposed on Ψ_i that satisfies the system in (27) is $\lfloor \frac{k+1}{k_0+1} \rfloor$. Hence, the number of equivalent sub-models is $m = \lfloor \frac{k+1}{k_0+1} \rfloor$.

Finally, the following Table reports the set of restrictions to be imposed on the $\mathcal{H}_6$ model to have $m = \lfloor \frac{7}{2} \rfloor = 3$ sub-models equivalent to the model $\mathcal{H}_{k_0}$ with $k_0 = 1$ lags.

Matrices in $\mathcal{H}_6 \rightarrow$	Ψ_{-1}	Ψ_0	Ψ_1	Ψ_2	Ψ_3	Ψ_4	Ψ_5	Ψ_6
$\mathcal{H}_{6,0}$	$\Psi_{-1,0}$	$\Psi_{0,0}$	$\Psi_{1,0}$	0	0	0	0	0
$\mathcal{H}_{6,1}$	$\Psi_{-1,0}$	0	$\Psi_{0,0}$	0	$\Psi_{1,0}$	0	0	0
$\mathcal{H}_{6,2}$	$\Psi_{-1,0}$	0	0	$\Psi_{0,0}$	0	0	$\Psi_{1,0}$	0

Table 5: Restrictions imposed on the $\mathcal{H}_6$ model when the model $\mathcal{H}_{k_0}$ is a FCVAR$_{d,b}$ with $k_0 = 1$ lag.

■

Proof of Proposition 5.1

Consider the model

$$\mathcal{H}_{r,k}: \quad \Delta^d X_t = \Pi\Delta^{d-b} L_b X_t + \sum_{j=1}^{k} \Gamma_j \Delta^{d-b} L_b X_t + \varepsilon_t$$

It can be written as

$$\sum_{i=-1}^{k} \Psi_j \Delta^{d+ib} X_t = \varepsilon_t$$

where $\Psi_{-1} = -\Pi$, $\Psi_0 = I + \Pi - \sum_{i=1}^{k} \Gamma_i$ and $\Psi_k = -(1)^{k+1}\Gamma_k$.

Consider the two sub-models of $\mathcal{H}_{r,k}$ with the following two restrictions:

$$\mathcal{H}_{p,k-1}: \qquad \Pi \text{ is a } p \times p \text{ matrix and } \Gamma_k = 0$$

$$\mathcal{H}_{0,k}: \qquad \Pi{=}0$$

The model $\mathcal{H}_{p,k-1}$ can be written as:

$$\sum_{i=-1}^{k-1} \tilde{\Psi}_i \Delta^{\tilde{d}+i\tilde{b}} X_t = \varepsilon_t$$

where $\tilde{\Psi}_{-1} = \Pi$, $\tilde{\Psi}_0 = I + \Pi - \sum_{i=1}^{k-1} \Gamma_i$ and $\tilde{\Psi}_{k-1} = (-1)^k \Gamma_{k-1}$.

The model $\mathcal{H}_{0,k}$ can be written as:

$$\sum_{i=0}^{k} \bar{\Psi}_i \Delta^{\bar{d}+i\bar{b}} X_t = \varepsilon_t$$

because $\bar{\Psi}_{-1} = 0$, $\bar{\Psi}_0 = I + 0 - \sum_{i=1}^{k} \Gamma_i$ and $\bar{\Psi}_k = (-1)^{k+1}\Gamma_k$.

The two sub-models are equal if

$$\begin{cases} \tilde{\Psi}_{-1} = \bar{\Psi}_0 \\ \tilde{\Psi}_0 = \bar{\Psi}_1 \\ \vdots \\ \tilde{\Psi}_{k-1} = \bar{\Psi}_k \end{cases} \quad \text{and} \quad \begin{cases} \tilde{d} - \tilde{b} = \bar{d} \\ \tilde{d} = \bar{d} + \bar{b} \\ \vdots \\ \tilde{d} + (k-1)\tilde{b} = \bar{d} + k\bar{b} \end{cases} \tag{28}$$

Given that the FCVARd, b model assumes that $d \geq b > 0$, it implies that $\tilde{d} \geq \tilde{b}$ and $\bar{d} \geq \bar{b}$. The inequality $\tilde{d} \geq \tilde{b}$ is always verified but $\bar{d} \geq \bar{b}$ is verified if and only if $\tilde{d} \geq 2\tilde{b}$. Therefore, $\mathcal{H}_{0,k} \subset \mathcal{H}_{p,k-1}$.

Consider the case in which Π is a reduced rank matrix with $0 < r < p$. $\mathcal{H}_{r,k-1}$ and $\mathcal{H}_{0,k}$ are equivalent if the systems of equations (28) hold. In this case, $\bar{\Psi}_0$ is equal to $I - \sum_{i=1}^{k} \bar{\Gamma}_i = -\tilde{\alpha}\tilde{\beta}'$. Hence, the models $\mathcal{H}_{r,k-1}$ are equivalent to $\mathcal{H}_{0,k}$ if and only if the $\mathcal{F}(d)$ condition is equal to 0. ■

A Figures

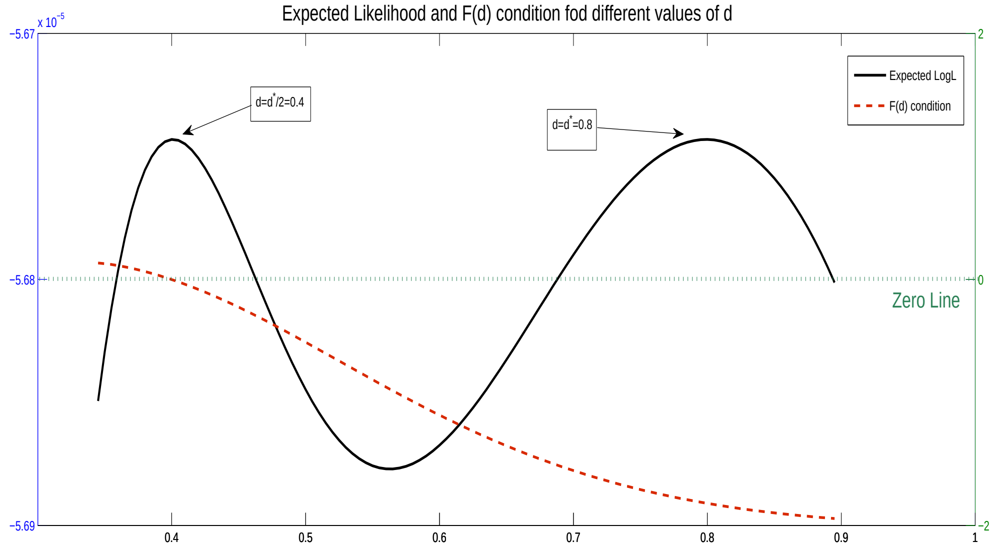


Figure 1: Figure reports simulated values of $\bar{l}(d)$ and $\bar{\mathcal{F}}(d)$ for different values of $d \in [0.2, 1.2]$. The DGP is generated with $k_0 = 0$ lags and a model $\mathcal{H}_k$ with $k = 1$ lags is fitted.

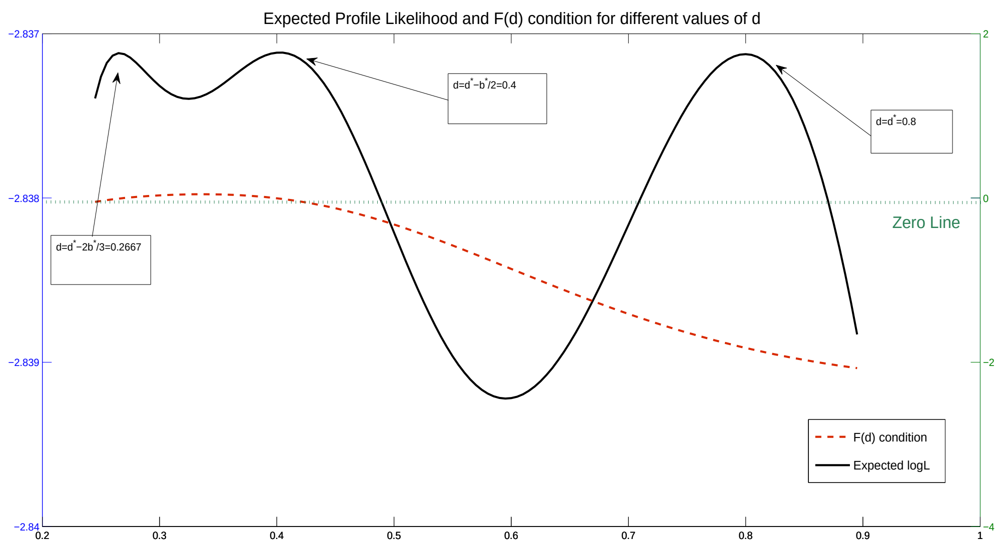


Figure 2: Figure reports simulated values of $\bar{l}(d)$ and $\bar{\mathcal{F}}(d)$ for different values of $d \in [0.2, 1.2]$. The DGP is generated with $k_0 = 0$ lags and a model $\mathcal{H}_k$ with $k = 2$ lags is fitted.

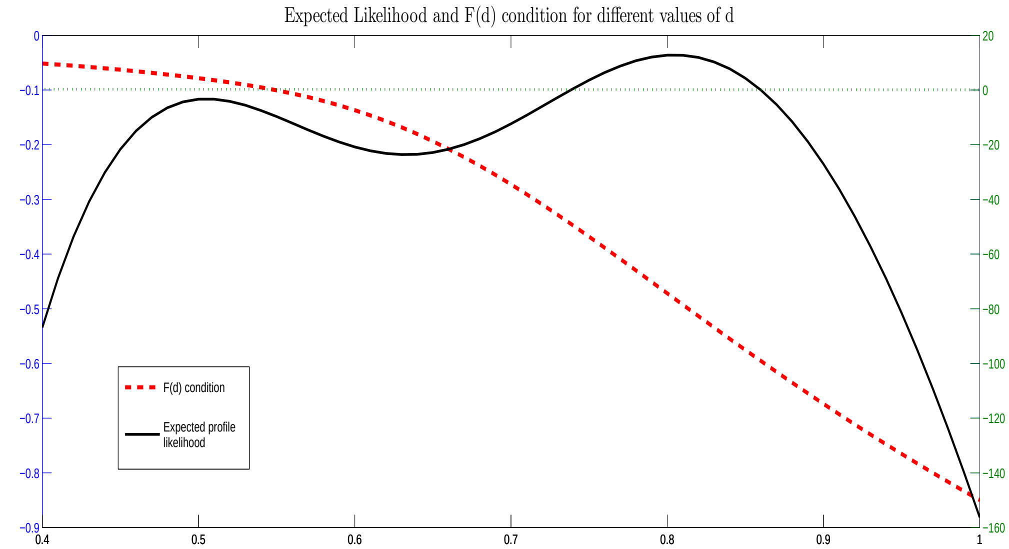


Figure 3: Figure reports simulated values of $\bar{l}(d)$ and $\bar{\mathcal{F}}(d)$ for different values of $d \in [0.4, 1]$. The DGP is generated with $k_0 = 1$ lags and a model $\mathcal{H}_k$ with $k = 2$ lags is fitted.

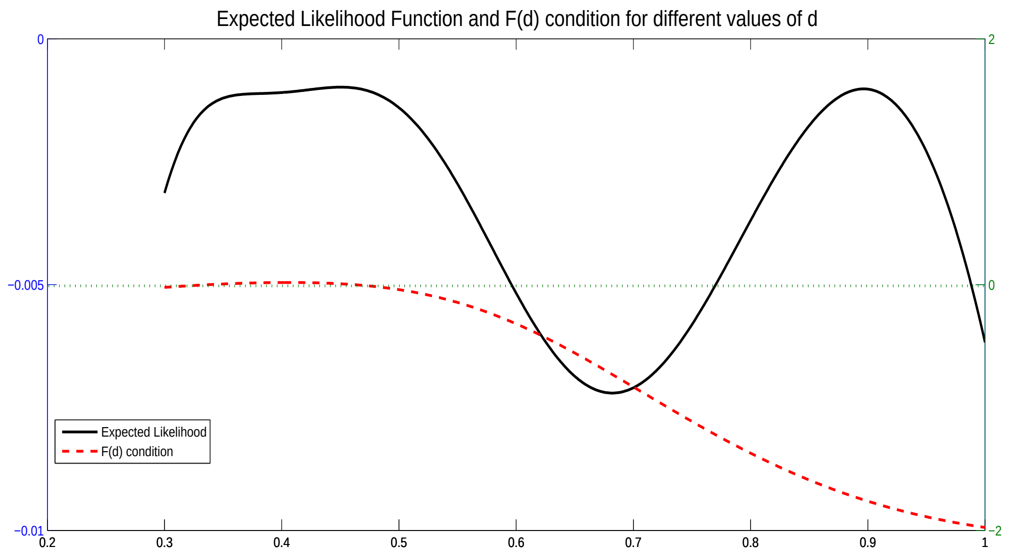
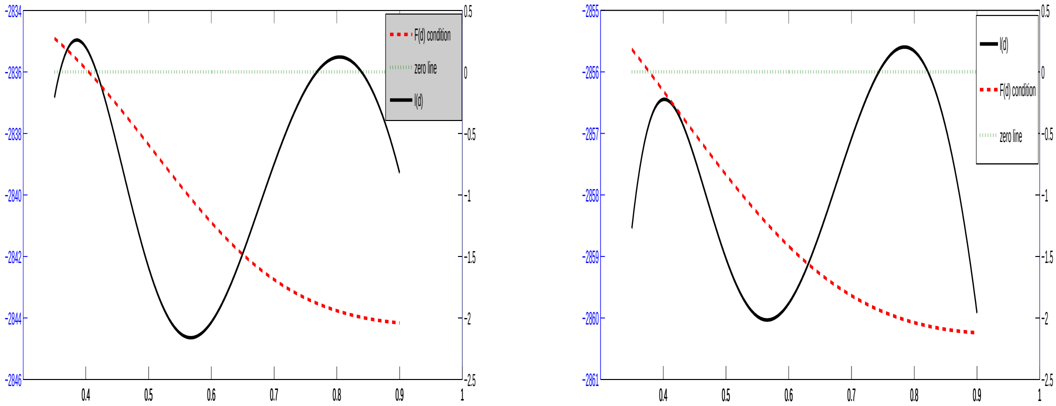
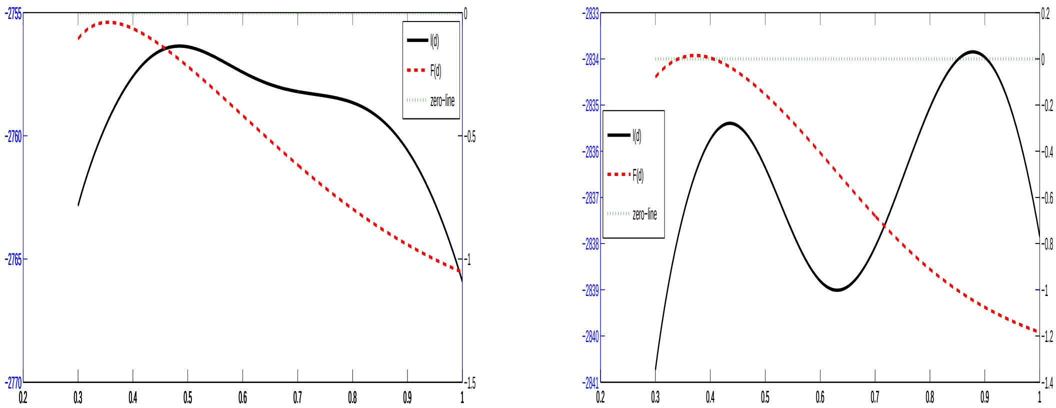


Figure 4: Figure reports simulated values of $\bar{l}(d)$ and $\bar{\mathcal{F}}(d)$ for different values of $d \in [0.3, 0.8]$. The DGP is generated with $k_0 = 1$ lags and a model $\mathcal{H}_k$ with $k = 3$ lags is fitted.



(a) Maximum around $d = 0.4$ (b) Maximum around $d = 0.8$

Figure 5: Figure reports the values of the profile likelihood $l(d)$ and $\mathcal{F}(d)$ for different values of $d \in [0.35, 0.9]$ for two different simulated path with $T = 1000$ of the FCVAR$_{d,d}$ when $k_0 = 0$ and model $\mathcal{H}_1$ is estimated in the data.



(a) Maximum around $d = 0.4$ (b) Maximum around $d = 0.8$

Figure 6: Figure reports the values of the profile likelihood $l(d)$ and $\mathcal{F}(d)$ for different values of $d \in [0.35, 0.9]$ for two different simulated path with $T = 1000$ of the FCVAR$_{d,d}$ when $k_0 = 1$ and model $\mathcal{H}_2$ is estimated in the data.

Research Papers 2013

2013-26: Nima Nonejad: Long Memory and Structural Breaks in Realized Volatility: An Irreversible Markov Switching Approach

2013-27: Nima Nonejad: Particle Markov Chain Monte Carlo Techniques of Unobserved Compdonent Time Series Models Using Ox

2013-28: Ulrich Hounyo, Sílvia Goncalves and Nour Meddahi: Bootstrapping pre-averaged realized volatility under market microstructure noise

2013-29: Jiti Gao, Shin Kanaya, Degui Li and Dag Tjøstheim: Uniform Consistency for Nonparametric Estimators in Null Recurrent Time Series

2013-30: Ulrich Hounyo: Bootstrapping realized volatility and realized beta under a local Gaussianity assumption

2013-31: Nektarios Aslanidis, Charlotte Christiansen and Christos S. Savva: Risk-Return Trade-Off for European Stock Markets

2013-32: Emilio Zanetti Chini: Generalizing smooth transition autoregressions

2013-33: Mark Podolskij and Nakahiro Yoshida: Edgeworth expansion for functionals of continuous diffusion processes

2013-34: Tommaso Proietti and Alessandra Luati: The Exponential Model for the Spectrum of a Time Series: Extensions and Applications

2013-35: Bent Jesper Christensen, Robinson Kruse and Philipp Sibbertsen: A unified framework for testing in the linear regression model under unknown order of fractional integration

2013-36: Niels S. Hansen and Asger Lunde: Analyzing Oil Futures with a Dynamic Nelson-Siegel Model

2013-37: Charlotte Christiansen: Classifying Returns as Extreme: European Stock and Bond Markets

2013-38: Christian Bender, Mikko S. Pakkanen and Hasanjan Sayit: Sticky continuous processes have consistent price systems

2013-39: Juan Carlos Parra-Alvarez: A comparison of numerical methods for the solution of continuous-time DSGE models

2013-40: Daniel Ventosa-Santaulària and Carlos Vladimir Rodríguez-Caballero: Polynomial Regressions and Nonsense Inference

2013-41: Diego Amaya, Peter Christoffersen, Kris Jacobs and Aurelio Vasquez: Does Realized Skewness Predict the Cross-Section of Equity Returns?

2013-42: Torben G. Andersen and Oleg Bondarenko: Reflecting on the VPN Dispute

2013-43: Torben G. Andersen and Oleg Bondarenko: Assessing Measures of Order Flow Toxicity via Perfect Trade Classification

2013-44: Federico Carlini and Paolo Santucci de Magistris: On the identification of fractionally cointegrated VAR models with the F(d) condition