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

Paolo Santucci de Magistris and Federico Carlini

CREATES Research Paper 2014-43

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 †

May 15, 2016

Abstract

This paper discusses identification problems in the fractionally cointegrated system of Johansen (2008) and Johansen and Nielsen (2012). It is shown that several equivalent reparametrizations of the model associated with different fractional integration and cointegration parameters may exist for any choice of the lag-length when the true cointegration rank is known. The properties of these multiple non-identified 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 $\mathcal{F}(d)$. This is a generalization of the well-known $I(1)$ condition to the fractional case. Imposing a proper restriction on the fractional integration parameter, d, is sufficient to guarantee identification of all model parameters and the validity of the $\mathcal{F}(d)$ condition. The paper also illustrates the indeterminacy between the cointegration rank and the lag-length. It is also proved that the model with rank zero and k lags may be an equivalent reparametrization of the model with full rank and $k-1$ lags. This precludes the possibility to test for the cointegration rank unless a proper restriction on the fractional integration parameter is imposed.

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

JEL Classification: C18, C32, C52

*CREATES, Department of Economics and Business Economics, Aarhus University.

†**Corresponding author**: CREATES, Department of Economics and Business Economics, Aarhus University, Fuglesangs Alle 4, 8210 Aarhus V, Denmark. Tel.: +45 8716 5319. E-mail address: psantucci@econ.au.dk.
The authors acknowledge support from CREATES - Center for Research in Econometric Analysis of Time Series (DNRF78), funded by the Danish National Research Foundation.

1 Introduction

The past 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 $d, b \in \mathbb{R}_+$ with $0 < b \leq d$. More specifically, the concept of fractional cointegration implies the existence of 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*.

Notable methodological works in the field of fractional cointegration are Robinson and Marinucci (2003) and Christensen and Nielsen (2006) that develop regression-based semi-parametric methods to evaluate whether two fractional stochastic processes share common trends. Analogously, Hualde and Velasco (2008) propose to check for the absence of cointegration by comparing the estimates of the cointegration vector obtained with OLS and those obtained with a GLS type of estimator. Breitung and Hassler (2002) propose a multivariate score test statistic to determine the cointegration rank that is obtained by solving a generalized eigenvalue problem of the type proposed by Johansen (1988). Alternatively, Robinson and Yajima (2002) and Nielsen and Shimotsu (2007) suggest a testing procedure to evaluate the cointegration rank of the multivariate coherence matrix of two, or more, fractionally differenced series. Chen and Hurvich (2003, 2006) estimate cointegrated spaces and subspaces by the eigenvectors corresponding to the r smallest eigenvalues of an averaged periodogram matrix of tapered and differenced observations.

Despite the effort spent in defining testing procedures for the presence of fractional cointegration, for a long time the literature in this area lacked a fully parametric multivariate model explicitly characterizing the joint behaviour of fractionally cointegrated processes. Interestingly, Granger (1986, p.222) 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. 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 is a tool for a direct modeling and testing of fractional cointegration. Johansen (2008) studies the properties of the model while Łasak (2010) suggests a profile likelihood approach to estimate the parameters and to test the hypothesis of absence of cointegration relations in the Granger (1986) model under the assumption that $d = 1$. Recently, Johansen and Nielsen (2012) have extended the estimation method of Łasak (2010) to the FCVAR$_{d,b}$ model, deriving the asymptotic properties of the profile maximum likelihood estimator when $0 \leq d-b < 1/2$ and $b \neq 1/2$. Other contributions in the parametric framework for fractional cointegration are in Avarucci and Velasco (2009), Franchi (2010) and Łasak and Velasco (2015).

This paper shows that the FCVAR$_{d,b}$ model is not globally identified when the number of lags, k, is unknown. For a given number of lags, several sub-models with the same conditional densities but different values of the parameters may exist. Hence the parameters of the FCVAR$_{d,b}$ model

cannot be uniquely identified. The multiplicity of not-identified sub-models can be determined for any FCVAR$_{d,b}$ model with k lags. An analogous identification problem, for the FIVAR$_b$ model is discussed in Tschernig et al. (2013a,b). This paper provides a detailed illustration of the identification problem in the FCVAR$_{d,b}$ framework. It is proved that the $I(1)$ condition in the VECM of Johansen (1988) can be generalized to the fractional context. In analogy with the $I(1)$ condition for integer orders of integration, this condition is named $\mathcal{F}(d)$, and it is a necessary and sufficient condition for the identification of the parameters of the system. If the $\mathcal{F}(d)$ condition is not satisfied, the FCVAR$_{d,b}$ parameters, including fractional and co-fractional parameters, d and b, cannot be uniquely determined.

This paper studies the problems of identification in the FCVAR$_{d,b}$ model along the following lines. First, Proposition 2.2 extends the results in Theorem 3 of Johansen and Nielsen (2012), highlighting the close relationship between the lag structure and the lack of identification, and deriving a necessary and sufficient condition for identification associated to any lag-length. Proposition 2.2 also highlights the consequence of the indeterminacy of the lag-length on the fractional parameters d and b, showing that the lack of identification is specific to a subset of all the possible choices of the number of lags. Second, the paper shows the consequence of the lack of identification on the likelihood function, both asymptotically and in finite samples. Differently from the standard case, where the integration orders are fixed to integer values, the estimation of the FCVAR$_{d,b}$ involves the maximization of the profile log-likelihood with respect to d and b, but the latter is affected by the indeterminacy generated by the over-specification of the lag-length. As expected, the lack of mathematical identification generates multiple absolute maxima in the profile log-likelihood function associated to different values of d and b when the number of lags is over-specified, thus confirming the statement in Proposition 2.2. Moreover, an interesting clue emerges from the finite sample analysis. Indeed, in finite samples, the profile log-likelihood function displays multiple maxima also when the identification is theoretically guaranteed. Moreover, a further identification issue, that emerges when the cointegration rank is unknown, is discussed. It is proved that there is a potentially large number of parameter sets associated with different choices of lag-length and cointegration rank for which the conditional density of the 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, it can be shown that, under certain restrictions, the FCVAR$_{d,b}$ with full rank and k lags is equivalent to the FCVAR$_{d,b}$ with rank 0 and $k+1$ lags. This last finding precludes the possibility to test for the absence of cointegration when the true number of lags is unknown based on the unrestricted FCVAR$_{d,b}$ model. Finally, we prove that the FCVAR$_{d,b}$ is identified for any lag $k > 1$, both in the known and unknown rank cases, if the fractional parameter d is restricted to be equal to the true fractional order, such that the $\mathcal{F}(d)$ condition is satisfied by construction. Building on this result, we show that to solve the identification problem it is sufficient to restrict the parameter set of d to belong to the sub-interval of $\mathbb{R}^+$ that includes the true fractional order, d_0, but excludes other values of $d < d_0$ associated to equivalent models. The information about the true fractional order can be obtained by the exact local Whittle estimator of Shimotsu and Phillips (2005).

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 FCVAR$_{d,b}$ model both asymptotically and in finite samples. Section 4 discusses the problems when the cointegration rank and the lag-length are both unknown. Section 5 concludes the paper.

2 The Identification Problem

This section provides a discussion of the identification problem related to the 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, and r defines the cointegration rank. Ω is the positive definite covariance matrix of the errors, and Γ_j, $j=1,\ldots,k$, 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 and obtaining the Granger representation for fractionally cointegrated processes. Following Definition 1 in Johansen and Nielsen (2012, p.2672), if X_t follows (1), then X_t is a fractional process of order d, denoted as $\mathcal{F}(d)$, and co-fractional of order $d-b$. The symbol $\mathcal{H}_k$ defines the model with k lags and $\theta = vec(d,b,\alpha,\beta,\Gamma_1,...,\Gamma_k,\Omega)$ is the parameter vector. The parameter space of model $\mathcal{H}_k$ is

$$\Theta_{\mathcal{H}_k} = \{\alpha\in\mathbb{R}^{p\times r_0}, \beta\in\mathbb{R}^{p\times r_0}, \Gamma_j\in\mathbb{R}^{p\times p}, j=1,\ldots,k, d\in\mathbb{R}^+, b\in\mathbb{R}^+, d\geq b>0, \Omega>0\},$$

where r_0 is the true cointegration rank and it is assumed known.[1]

Similarly to Johansen (2010), the concept of identification and equivalence between two models is formally introduced by the following definition.

Definition 2.1 *Let $\{\mathcal{P}=P_\theta,\theta\in\Theta\}$ 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 parameter function $g(\theta)$ is not identified. In this case, the statistical models P_{θ_1} and P_{θ_2} are equivalent.*

It can be shown that the parameters of the FCVAR$_{d,b}$ model in (1) are not identified, i.e. several *equivalent* sub-models associated with different values θ, can be found.

Example 1: An illustration of the identification problem is provided by the following example. Consider the FCVAR$_{d,b}$ model with one lag,

$$\mathcal{H}_1: \Delta^d X_t = \alpha\beta'\Delta^{d-b}L_b X_t + \Gamma_1\Delta^d L_b X_t + \varepsilon_t, \tag{2}$$

[1]The results of this Section are obtained under the maintained assumption that the true cointegration rank is known and such that $0<r_0<p$. An extension to the case of unknown rank and number of lags is presented in Section 4.

which can be written as

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

First, examine the restriction, $\mathcal{H}_1^{(0)}:\Gamma_1^0=0$. Under $\mathcal{H}_1^{(0)}$, the model in equation (2) can be rewritten as

$$\left\{\Delta^{d_0}[I_p+\alpha\beta']+\Delta^{d_0-b_0}[-\alpha\beta']\right\}X_t=\varepsilon_t.$$

Second, consider instead the restriction $\mathcal{H}_1^{(1)}:I_p+\alpha\beta'-\Gamma_1^1=0$. It follows that

$$\left\{\Delta^{d_1-b_1}\left[-\alpha\beta'\right]+\Delta^{d_1+b_1}[I_p+\alpha\beta']\right\}X_t=\varepsilon_t.$$

Given that the condition $\alpha\beta'\Delta^{d_0-b_0}=\alpha\beta'\Delta^{d_1-b_1}$ must hold in both sub-models,[2] hence model (2) under $\mathcal{H}_1^{(0)}$ is equivalent to the model (2) under $\mathcal{H}_1^{(1)}$ if and only if

$$[I_p+\alpha\beta^{'}]\Delta^{d_0}=[I_p+\alpha\beta^{'}]\Delta^{d_1+b_1}.$$

This leads to the system of two equations in d_0, b_0, d_1 and b_1

$$\begin{cases}d_0-b_0=d_1-b_1\\ d_0=d_1+b_1\end{cases}\tag{3}$$

which has a unique solution when $d_1=d_0-b_0/2$ and $b_1=b_0/2$. Since the restrictions $\mathcal{H}_1^{(0)}$ and $\mathcal{H}_1^{(1)}$ lead to equivalent descriptions of the data, it follows that the fractional order of X_t implied by both models must be the same. However, in $\mathcal{H}_1^{(0)}$ the fractional order is represented by the parameter d_0, i.e. $X_t\sim\mathcal{F}(d_0)$ since $\Delta^{d_0}X_t\sim\mathcal{F}(0)$, while in $\mathcal{H}_1^{(1)}$ the fractional order is given by the sum d_1+b_1, i.e. $X_t\sim\mathcal{F}(d_1+b_1)$. The identification condition defined in 2.1 is clearly violated, as the conditional densities of $\mathcal{H}_1^{(0)}$ and $\mathcal{H}_1^{(1)}$ are such that

$$p_{\mathcal{H}_1^{(0)}}(X_1,...,X_T,\theta_0|X_0,X_{-1},\ldots)=p_{\mathcal{H}_1^{(1)}}(X_1,...,X_T,\theta_1|X_0,X_{-1},\ldots),\tag{4}$$

where $\theta_0=vec(d_0,b_0,\alpha,\beta,\Omega)$ and $\theta_1=vec(d_1,b_1,\alpha,\beta,\Gamma_1^1,\Omega)$ with $\Gamma_1^1=I_p+\alpha\beta'$.

Example 1 can be extended to a generic lag-length $k_0\geq 0$. Consider the model $\mathcal{H}_{k_0}$

$$\mathcal{H}_{k_0}:\quad \Delta^{d_0}X_t=\alpha_0\beta_0'\Delta^{d_0-b_0}L_{b_0}X_t+\sum_{i=1}^{k_0}\Gamma_i^0\Delta^{d_0}L_{b_0}^iX_t+\varepsilon_t\quad \varepsilon_t\sim N(0,\Omega_0),\tag{5}$$

with $k_0\geq 0$ lags, and $|\alpha_{0,\perp}'\Gamma^0\beta_{0,\perp}|\neq 0$ with $\Gamma^0=I_p-\sum_{i=1}^{k_0}\Gamma_i^0$. When a model $\mathcal{H}_k$ with $k>k_0$ is considered, then $\mathcal{H}_{k_0}$ is associated with the set of restrictions $\mathcal{H}_k^{(0)}:\Gamma_{k_0+1}=\Gamma_{k_0+2}=...=\Gamma_k=0$ imposed on $\mathcal{H}_k$. However, there may be several alternative restrictions on $\Gamma_{k_0+1},\Gamma_{k_0+2},...,\Gamma_k$

[2]Note that this paper does not discuss the identification of the matrices α and β. As noted in Johansen (1995a, p.177), the product $\alpha\beta'$ is identified but not the matrices α and β because if there was an $r\times r$ matrix ξ, the product $\alpha\beta'$ would be equal to $\alpha_\xi\beta_\xi'$ where $\alpha_\xi=\alpha\xi$ and $\beta_\xi=\beta(\xi')^{-1}$.

leading to an equivalent sub-model as the one obtained under $\mathcal{H}_k^{(0)}$.

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

Proposition 2.2 *Consider a FCVAR$_{d,b}$ model with k lags,*

i) Given $k > k_0 \geq 0$, the $\mathcal{F}(d)$ condition, defined as $|\alpha'_\perp \Gamma \beta_\perp| \neq 0$ with $\Gamma = I_p - \sum_{i=1}^k \Gamma_i$, is a necessary and sufficient condition for the identification of the set of parameters of $\mathcal{H}_k$ in equation (5).

ii) 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.

iii) 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 $j = 0, 1, ..., m-1$.

Proof in Appendix A.1.

Proposition 2.2 has several important consequences that are worth being discussed in detail. First of all, the $\mathcal{F}(d)$ condition only holds for the sub-model of $\mathcal{H}_k$ for which $d = d_0$ and $b = b_0$, i.e. for the sub-model of $\mathcal{H}_k$ corresponding to the restriction $\mathcal{H}_k^{(0)} : \Gamma_{k_0+1} = \Gamma_{k_0+2} = ... = \Gamma_k = 0$. In the Example 1, the $\mathcal{F}(d)$ condition is only verified for $\mathcal{H}_1^{(0)}$, while for $\mathcal{H}_1^{(1)}$ we have that $|\alpha'_\perp \Gamma^1 \beta_\perp| = 0$, since $\Gamma^1 = I_p - (I_p + \alpha\beta') = -\alpha\beta'$. Note that the assumption $|\alpha'_{0,\perp} \Gamma^0 \beta_{0,\perp}| \neq 0$ imposed on model (5) guarantees that it is not possible to find restrictions on $\mathcal{H}_{k_0}$ for which two or more sub-models are equivalent. In this sense Proposition 2.2 generalizes Theorem 3 in Johansen and Nielsen (2012). Indeed, while in Johansen and Nielsen (2012) the $\mathcal{F}(d)$ condition is only imposed on the $\mathcal{H}_{k_0}$ model with $k = k_0$ by assumption, Proposition 2.2.*i*) shows that a necessary and sufficient condition for the identification of the parameters of any $\mathcal{H}_k$ model, with $k > k_0$, is the validity of the $\mathcal{F}(d)$ condition. This has important consequences in practical applications when the true number of lags is unknown and it is potentially over-specified.[3]

When $d = b = 1$, then the FCVAR$_{d,b}$ model reduces to the usual $VECM$ model and the $\mathcal{F}(d)$ condition reduces to the $I(1)$ condition that excludes solutions of the $VECM$ that are integrated of order 2 or higher, see for example the discussion in Johansen (2009). Indeed, the $\mathcal{F}(d)$ condition has analogies in the classical $I(1)$ and $I(2)$ context and it can be better understood by looking at the $I(2)$ cointegration model as discussed in Johansen (1995b). The model is

$$\Delta^2 X_t = \Gamma \Delta X_{t-1} + \Pi X_{t-2} + \sum_{i=1}^{k-2} \Psi_i \Delta^2 X_{t-i} + \epsilon_t. \tag{6}$$

which can be found by imposing proper restrictions on the Π_i matrices of the the unrestricted $VAR(k)$ on X_t, $X_t = \sum_{i=1}^k \Pi_i X_{t-i} + \epsilon_t$. Depending on the restrictions imposed on the matrices Π, Γ and $\Psi_1, ..., \Psi_{k-2}$, model (6) allows for three types of statistical models: $I(0)$, $I(1)$ and $I(2)$.

[3]When the number of lags is under-specified there is no identification problem, but the model is misspecified and the results in Johansen and Nielsen (2012) do not hold.

If Π has full rank, then $X_t \sim I(0)$, see Theorem 1 in Johansen (1995b). If $\Pi = \alpha'\beta$ and the matrix $\alpha'_\perp \Gamma \beta_\perp$ has full rank, it follows from Theorem 2 in Johansen (1995b) that $X_t \sim I(1)$. If instead the matrix $\alpha'_\perp \Gamma \beta_\perp$ is of reduced rank, then X_t contains both $I(2)$ and $I(1)$ common trends, whose number depends on the rank of Π and $\alpha'_\perp \Gamma \beta_\perp$. This means that the condition on the rank of $\alpha'_\perp \Gamma \beta_\perp$ determines two distinct models, which in turn may imply alternative explanations of the relationships between economic series. Similarly, a model for multiple (or polynomial) fractional cointegration can be obtained by proper restrictions of the unrestricted $VAR_{d,b}$ model, see Johansen (2008, p.667), as

$$\Delta^d X_t = \Delta^{d-2b}(\alpha\beta' L_b X_t - \Gamma \Delta^b L_b X_t) + \sum_{i=1}^{k} \Psi_i \Delta^d L_b^i X_t + \epsilon_t. \tag{7}$$

Depending on the rank of $\alpha'_\perp \Gamma \beta_\perp$ it is possible to find cointegration relations of order $I(d-b)$ and $I(d-2b)$. Setting $d = 2$ and $b = 1$ we obtain model (6) with $I(2)$ and $I(1)$ trends. It is important to stress that the condition $|\alpha'_{0,\perp}\Gamma^0 \beta_{0,\perp}| \neq 0$ imposed on model (5) excludes the possibility that the FCVAR$_{d,b}$ model with k_0 lags can be re-written as model (7), thus ruling out polynomial fractional cointegration.[4] Consider model $\mathcal{H}_1^{(1)}$ in Example 1 again, where $|\alpha'_\perp \Gamma \beta_\perp| = 0$. After simple algebraical manipulations, model $\mathcal{H}_1^{(1)}$ can be formulated as

$$\Delta^{d_2} X_t = \Delta^{d_2 - 2b_1}(\alpha\beta' L_{b_1} X_t - \Gamma^1 \Delta^{b_1} L_{b_1} X_t) + \epsilon_t \tag{8}$$

where $d_2 = d_1 + b_1$ and $\Gamma^1 = -\alpha\beta'$. This example illustrates the close link between the possibility of polynomial fractional cointegration and the indeterminacy of lag-length and FCVAR$_{d,b}$ parameters as illustrated in Proposition 2.2. In particular, imposing the $\mathcal{F}(d)$ condition on the FCVAR$_{d,b}$ model does not only guarantee that the parameters d, b and $\Gamma_1, ..., \Gamma_k$ are correctly identified, but also rules out cases of polynomial fractional cointegration.

In addition, Proposition 2.2.*ii*) characterizes the number of equivalent sub-models of $\mathcal{H}_k$ for a given k_0, showing that their multiplicity depends on k and k_0. Analogously to the example above, this means that models with polynomial fractional cointegration up to order $m = \lfloor \frac{k+1}{k_0+1} \rfloor$ can be obtained from the FCVAR$_{d,b}$ model for some combinations of k and k_0. Table 1 summarizes the number of equivalent sub-models for different values of k_0 and k. Interestingly, as a consequence of Proposition 2.2.*ii*), there are cases in which $k > k_0$ does not necessarily imply a lack of identification. For example, when $k = 2$ and $k_0 = 1$ there are no sets of restrictions on $\mathcal{H}_2$ leading to a sub-model equivalent to the one obtained under the restriction $d = d_0$, $b = b_0$, $\Gamma_1 = \Gamma_1^0$ and $\Gamma_2 = 0$. Hence, in this case, the multiplicity, m, of equivalent sub-models is 1. When k_0 is small there are several equivalent sub-models for small choices of k. As k_0 increases, multiple equivalent sub-models are only found for large values of k. For example, when $k_0 = 5$, then two equivalent sub-models can only be found for suitable restrictions of the $\mathcal{H}_{11}$ model. Moreover, Proposition 2.2.*iii*) shows that each sub-model of $\mathcal{H}_k$ equivalent to $\mathcal{H}_{k_0}$ with $|\alpha'_\perp \Gamma \beta_\perp| = 0$ has values of d and b that are fractions of d_0 and b_0. Interestingly, when k is very large compared to k_0, the $(m-1)$-th

[4] The model of Franchi (2010) extends the FCVAR$_{d,b}$ model to a flexible form of polynomial fractional cointegration. An investigation of the identification conditions in Franchi (2010)'s model is left to future research.

$k_0 \downarrow k \rightarrow$	**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.

sub-model is associated with $d_{m-1} \approx d_0 - b_0$ and $b_{m-1} \approx 0$, i.e. located closely to the boundary of the parameter space. Compared to the classic VECM, the parameters d and b must be estimated in the FCVAR$_{d,b}$ model. However, the lack of identification precludes the possibility of uniquely determining the fractional parameters if k is over-specified. Therefore, 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, information on the fractional order of X_t, $\mathcal{F}(d)$, can be used to correctly identify the fractional parameters d and b when model $\mathcal{H}_k$ is estimated on the data.

As shown in Johansen and Nielsen (2012), the parameters of the FCVAR$_{d,b}$ can be estimated following a profile likelihood approach. Indeed, the estimates of the fractional parameters, $\hat{d}$ and $\hat{b}$, are obtained by maximizing the profile log-likelihood

$$\hat{\psi} = \arg\max_{\psi} \ell_T(\psi), \tag{9}$$

where $\psi = (d, b)'$ and

$$\ell_T(\psi) = -\log|S_{00}(\psi)| - \sum_{i=1}^{r} \log(1 - \lambda_i(\psi)). \tag{10}$$

The quantities $\lambda(\psi)$ and $S_{00}(\psi)$ are obtained from the residuals, $R_{it}(\psi)$ for $i = 0, 1$, of the reduced rank regression of $\Delta^d X_t$ on $\Delta^d L_b^j X_t$ and $\Delta^{d-b} L_b X_t$ on $\Delta^d L_b^j X_t$ for $j = 1, .., k$, respectively. The product moment matrices $S_{ij}(\psi)$ for $i, j = 0, 1$ are $S_{ij}(\psi) = T^{-1} \sum_{t=1}^{T} R_{it}(\psi) R'_{jt}(\psi)$ and $\lambda_i(\psi)$ for $i = 1, \ldots, p$ are the solutions, sorted in decreasing order, of the generalized eigenvalue problem

$$|\lambda(\psi) S_{11}(\psi) - S_{10}(\psi) S_{00}^{-1}(\psi) S_{01}(\psi)| = 0. \tag{11}$$

Given $\hat{d}$ and $\hat{b}$, the estimates $\hat{\alpha}$, $\hat{\beta}$, $\hat{\Gamma}_j$, $j = 1, \ldots, k$, and $\hat{\Omega}$ are found by reduced rank regression as in Johansen (1988). Although the the statistical model (5) is defined for all $0 < b_0 \leq d_0$, the

asymptotic properties of the ML estimator are derived in Johansen and Nielsen (2012) when the true values satisfy $0 \leq d_0 - b_0 < 1/2$ and $b_0 \neq 1/2$, for which $\beta_0' X_t$ is (asymptotically) a stationary process. Therefore, the following analysis is carried out for combinations of d_0 and b_0, which satisfy such constraint.

The values of ψ that maximize $\ell_T(\psi)$ must be found numerically. The consequences of the lack of identification of the FCVAR$_{d,b}$ model on the expected profile log-likelihood when $k > k_0$ are therefore explored by means of Monte Carlo simulations. Since the asymptotic value of $\ell_T(\psi)$ is not available in closed-form as a function of the model parameters, the asymptotic behavior of $\ell_T(\psi)$ is approximated averaging, over M simulations, the value of $\ell_T(\psi)$ computed for different values of ψ and a large T. This provides a precise numerical approximation of the expected profile log-likelihood, $E[\ell_T(\psi)]$. Therefore, $M = 100$ simulated paths are generated from model (5) with $T = 50,000$ observations and $p = 2$. The fractional parameters of the system are $d_0 = 0.8$ and $b_0 = d_0$. The assumption $b_0 = d_0$ simplifies the readability of the results without loss of generality, since the plots display $E[\ell_T(d)]$ as a function of d in a two dimensional Cartesian system. The cointegration vector is $\beta_0 = [1, -1]'$, the vector of adjustment coefficients is $\alpha_0 = [0.5, -0.5]'$, and the matrices Γ_i^0, $i = 1, ..., k_0$, for different values of k_0 are chosen such that the roots of the characteristic polynomial are outside the fractional circle, see Johansen (2008). The average profile log-likelihood, $\bar{\ell}_T(d)$, and the average of the function $f(d) = |\hat{\alpha}_\perp'(d)\hat{\Gamma}(d)\hat{\beta}_\perp(d)|$ are computed with respect to a grid of alternative values for $d = [d_{\min}, \ldots, d_{\max}]$. The average of $f(d)$ over the M simulations is a an estimate of the value of the $\mathcal{F}(d)$ condition for different values of d. Hence $\bar{\mathcal{F}}(d) = \frac{1}{M}\sum_{i=1}^{M} f_i(d)$ for $d = [d_{\min}, \ldots, d_{\max}]$ is plotted together with $\bar{\ell}_T(d)$.[5]

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 the two global maxima of $\bar{\ell}_T(d)$ are associated to the pair of values $d = 0.4$ and $d = 0.8$, but when $d = 0.4$ the $\bar{\mathcal{F}}(d)$ line is equal to zero. Similarly, as reported in Figure B.1 in Appendix B, the expected log-likelihood function has three humps around $d = 0.8$, $d = 0.4$ and $d = 0.2667 = d_0/3$ when $k = 2$ and $k_0 = 0$. As in the previous case, when $d = 0.4$ and $d = 0.2667$, the line with $\bar{\mathcal{F}}(d)$ is approximately equal to zero. Consistently with the theoretical results presented in Section 2, the $\bar{\mathcal{F}}(d)$ line is far from zero in $d = 0.8$ also in this case.

Figure 2 reports the contour plot of the expected profile log-likelihood function in the 2-dimensional space of $(d, b) \in \mathbb{R}^2$, with $d \geq b$. The plot clearly highlights the presence of two equivalent peaks located inside the isolines with level -14.1928 that, as expected, are associated with the vectors $\psi_0 = [0.8, 0.8]'$ and $\psi_1 = [0.4, 0.4]'$. Notably, the function $\bar{l}(\psi)$ quickly decreases at the extremes of the parameter space, i.e. when $d > d_0$ and $b > b_0$ or when $d < d_0 - b_0/2$ and $b < b_0/2$. Instead, the function remains rather high and flat in the interval $b_0/2 < b \leq d < d_0$. This may induce further identification problems in finite samples as discussed in Section 3.1.

A slightly more complex evidence arises when $k_0 > 0$. Figures 3 and B.2 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 is globally

[5]Due to space constraints, the results of the Monte Carlo simulations cannot be shown for many combinations of parameter values. The results for different combinations of the parameters confirm the evidence reported here and they are available upon request from the authors. The values of $d_{\min}$ and $d_{\max}$ on the x-axis of the graphs change to improve the clarity of the plots.

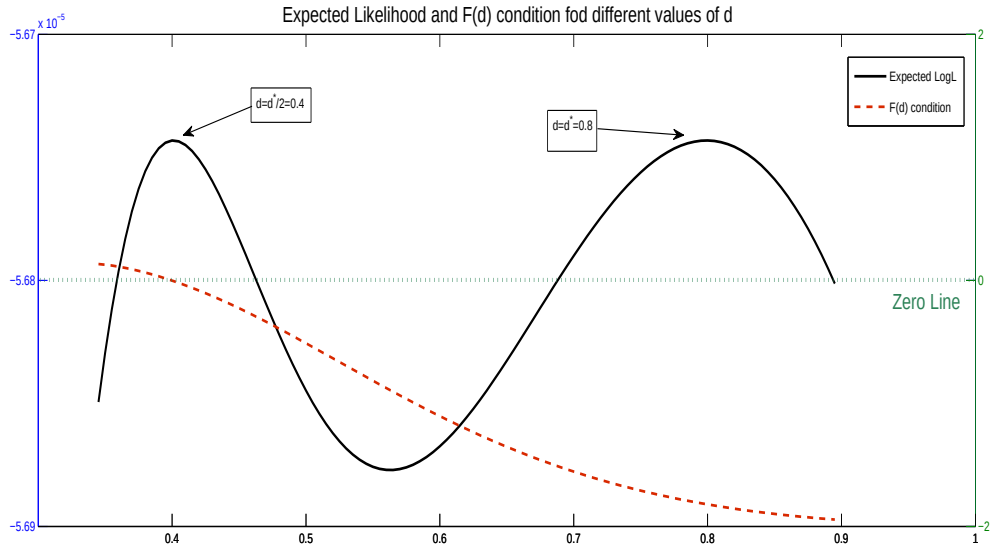


Figure 1: Figure reports simulated values of $\bar{l}(d)$ and $\bar{\mathcal{F}}(d)$ for different values of $d \in [0.2, 1.2]$ on the x-axis. The observations from the DGP are generated with $k_0 = 0$ lags and model $\mathcal{H}_k$ with $k = 1$ lags is estimated. The parameters of the DGP are $d_0 = b_0 = 0.8$, $\beta_0 = [1, -1]'$, $\alpha_0 = [-0.5, 0.5]'$.

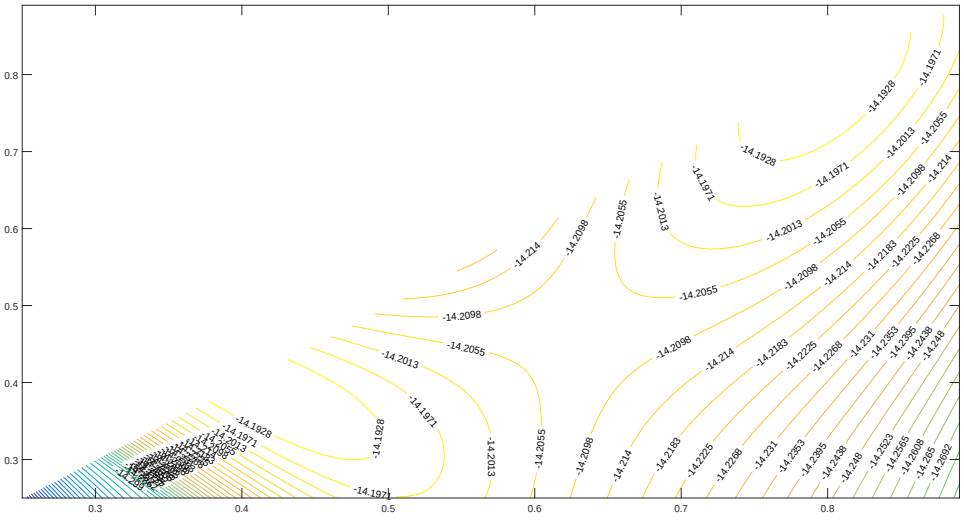


Figure 2: Figure reports the contour plot of the values (rescaled by a 10000) of the function $\bar{l}(\psi)$ for different combinations of $d \in [0.2, 1.2]$ (x-axis) and $b \in [0.2, 1.2]$ (y-axis). The observations from the DGP are generated with $k_0 = 0$ lags and model $\mathcal{H}_k$ with $k = 1$ lag is estimated. The parameters of the DGP are $d_0 = b_0 = 0.8$, $\beta_0 = [1, -1]'$, $\alpha_0 = [-0.5, 0.5]'$. The empty area is associated to values of $b > d$ for which the log-likelihood is not defined.

maximized in the region around $d = 0.8$, 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 $d = 0.5$, possibly inducing 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 associated with $d = d_0 = 0.8$ and $d = d_0/2 = 0.4$. Indeed, by looking at Figure B.2 in Appendix B it emerges that the line $\bar{\ell}_T(d)$ has two global maxima around the values of $d = 0.4$ and $d = 0.8$. As expected, in the region around $d = 0.4$ the $\bar{\mathcal{F}}(d)$ line is close to zero. .

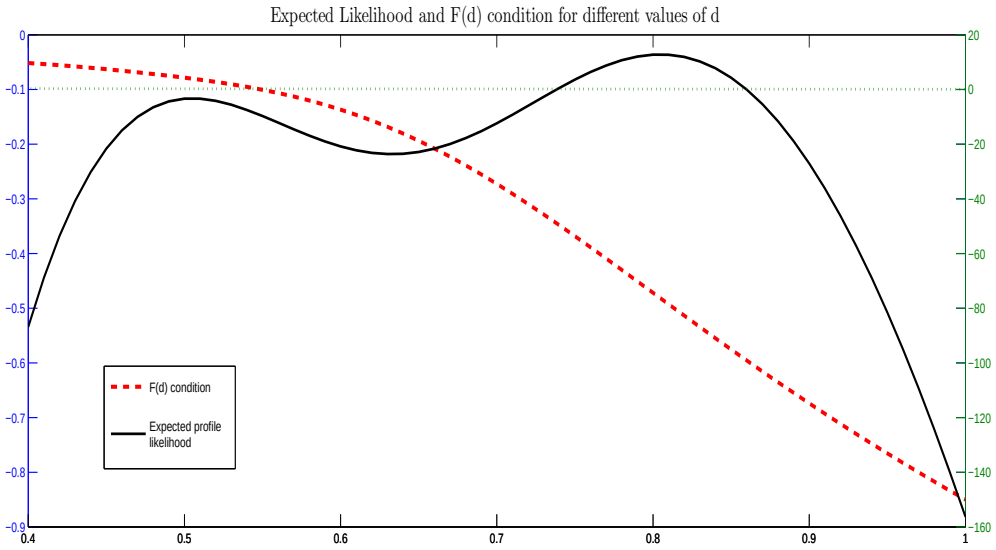
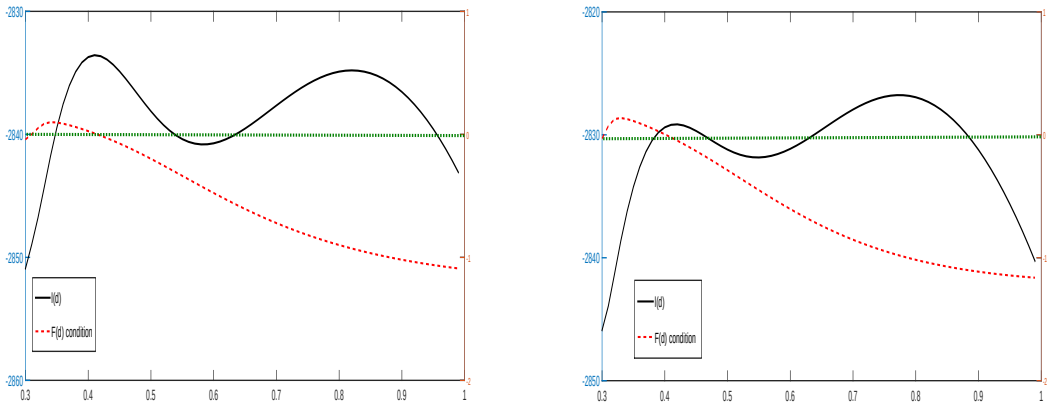


Figure 3: Figure reports simulated values of $\bar{l}(d)$ and $\bar{\mathcal{F}}(d)$ for different values of $d \in [0.4, 1]$ on the x-axis. The observations from the DGP are generated with $k_0 = 1$ lags and model $\mathcal{H}_k$ with $k = 2$ lags is estimated. The parameters of the DGP are $d_0 = b_0 = 0.8$, $\beta_0 = [1, -1]'$, $\alpha_0 = [-0.5, 0.5]'$, and $\Gamma_1 = \left[\begin{smallmatrix} 0.3 & -0.2 \\ 0.4 & -0.5 \end{smallmatrix}\right]$.

3.1 Identification in Finite Samples

In Section 2, the mathematical identification of the FCVAR$_{d,b}$ has been discussed theoretically. The purpose of this Section is to shed light on the consequences of the lack of mathematical identification in finite samples. From the analysis above, we know that for some $k > k_0$, the expected profile log-likelihood displays multiple equivalent maxima associated with fractions of d_0. This section focuses on the consequences of the lack of identification when the sample size, T, is finite.

Figure 4 reports the finite sample profile log-likelihood function, $\ell_T(d)$, against a fine grid of values of d. Each plot reports the function $\ell_T(d)$ obtained by fitting model $\mathcal{H}_1$ on a distinct simulated path of length $T = 1000$, generated under model $\mathcal{H}_0$. The plot clearly highlights the consequences of the lack of identification in finite samples. In Panel a), the global maximum of $\ell_T(d)$ is found around $d = 0.4$, while in Panel b) it is around 0.8. As expected in Panel a), the $f(d)$ line is near 0 when $d = 0.4$, while it is far from zero in Panel b) when $d = 0.8$. As it emerges from the plots in Figure 3, the generalized lag structure of the FCVAR$_{d,b}$ model also induces poor finite



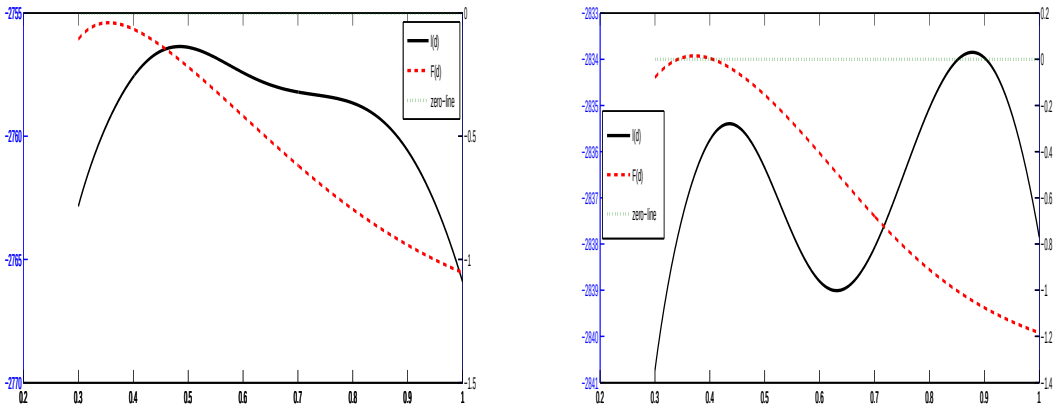
(a) Maximum around $d = 0.4$ (b) Maximum around $d = 0.8$

Figure 4: Figure reports the values of the profile log-likelihood $l(d)$ and $\mathcal{F}(d)$ for different values of $d \in [0.35, 0.9]$ (x-axis) for two different simulated paths with $T = 1000$ of the FCVAR$_{d,d}$ when $k_0 = 0$ and model $\mathcal{H}_1$ is estimated. The parameters of the DGP are $d_0 = b_0 = 0.8$, $\beta_0 = [1, -1]'$, $\alpha_0 = [-0.5, 0.5]'$.

sample identification, namely *weak* identification, for any $k > k_0$. As in Figure 4, Figure 5 reports the finite sample profile log-likelihood function relative to the estimation of the $\mathcal{H}_2$ model on two simulated paths of $\mathcal{H}_1$ with $T = 1000$. In Panel a), the global maximum is in a neighborhood of $d = 0.4$, and the function $f(d)$ is close to zero in $d = 0.4$. Hence, the estimated matrices $\hat{\Gamma}_1$ and $\hat{\Gamma}_2$ are such that $|\alpha'_\perp \Gamma \beta_\perp| = 0$. On the other hand, with another simulated path, the global maximum is found around $d = 0.8$, where the function $f(d)$ is far from zero, Panel b). As it emerges from this example, for any choice of $k > k_0$ there is the risk of obtaining estimates of the fractional parameters, d and b, that are far from the true ones. Tschernig et al. (2013a) discuss an analogous identification problem in the FIVAR$_b$ model. The FIVAR$_b$ extends the FIVAR model allowing the autoregressive structure to depend on the fractional lag operator, hence inducing more flexibility in the short-run term. The FIVAR$_b$ model is defined as

$$\Delta(L, d)Y_t = \sum_{i=1}^{l} \Phi_i L_b^i \Delta(L, d)Y_t + \epsilon_t \tag{12}$$

where Y_t is p-dimensional vector of detrended processes and $\Delta(L, d) = diag(\Delta^{d_1}, \Delta^{d_2}, ..., \Delta^{d_p})$ allows for different integration orders between the elements of Y_t. Similarly to the FCVAR$_{d,b}$ model, when $b = 0$ the matrices Γ_i are not identified, so that b must be larger than 0 also in the FIVAR$_b$ model. Tschernig et al. (2013a) shows that another identification problem arises when the eigenvalues of the characteristic polynomial in the L_b operator, $\Phi(L_b) = I_p - \sum_{i=1}^{k} \Phi_i L_b^i$, are either close to 0 or to 1. Similarly to the FCVAR$_{d,b}$, the lack of identification leads to an high and flat log-likelihood function for a wide range of combinations of d and b. However, in the FCVAR$_{d,b}$ model, the $\mathcal{F}(d)$ condition provides a necessary and sufficient condition for the identification. It is therefore crucial to develop a robust estimation procedure that guarantees that the estimated



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

Figure 5: Figure reports the values of the profile log-likelihood $l(d)$ and $\mathcal{F}(d)$ for different values of $d \in [0.35, 0.9]$ (x-axis) for two different simulated paths with $T = 1000$ of the FCVAR$_{d,d}$ when $k_0 = 1$ and model $\mathcal{H}_2$ is estimated in the data. The parameters of the DGP are $d_0 = b_0 = 0.8$, $\beta_0 = [1, -1]'$, $\alpha_0 = [-0.5, 0.5]'$, and $\Gamma_1 = \left[\begin{smallmatrix} 0.3 & -0.2 \\ 0.4 & -0.5 \end{smallmatrix}\right]$.

FCVAR$_{d,b}$ parameters are correctly identified and satisfy the $\mathcal{F}(d)$ condition also when the lag-length is potentially overspecified.

3.2 Constrained Likelihood

In the previous sections, we have proved that the FCVAR$_{d,b}$ model suffers from identification problems when k is over-specified. In particular, a number of equivalent parametrization associated to fractions of the true d_0 and b_0 can be found for several choices of $k > k_0$. On the other hand, the fractional parameter d is equivalent to the true fractional order of the process X_t only in $d = d_0$. As illustrated above, this identification problem has clear consequences from a statistical point of view since an unique ML estimator of d and b cannot be determined, since the profile log-likelihood function does not have an unique maximum around d_0 and b_0. We therefore propose a new approach that is based on the idea of transforming the unrestricted maximum likelihood problem, whose properties have been studied in Johansen and Nielsen (2012) only for the case $k = k_0$, into a constrained maximum likelihood problem by imposing a very mild restriction on the parameter space of d. In particular, we suggest that $\hat{d}$ and $\hat{b}$ must be the solutions of the following constrained maximum likelihood problem

$$\begin{aligned} \hat{\psi} &= \arg\max_{\psi} \ell_T(\psi), \\ &\text{s.t. } d \geq \delta_{\min} \end{aligned} \tag{13}$$

where $\ell_T(\psi)$ is defined in (10) and $\delta_{\min}$ determines the lower bound on the parameter d. Restricting the parameter space of d is supported by the following lemma, which is a direct derivation of Proposition 2.2.

Lemma 3.1 *Let* $\tilde{\Theta}_{\mathcal{H}_k} = \{d = d_0, b \in [0, d_0], \alpha \in \mathbb{R}^{p\times r}, \beta \in \mathbb{R}^{p\times r}, \Gamma_j \in \mathbb{R}^{p\times p}, j = 1, \ldots, k; \Omega > 0\}$ *be the restricted parameter space of model* $\Theta_{\mathcal{H}_k}$ *with* $d = d_0 \in \mathbb{R}^+$*, then the statistical model* $\mathcal{P} = \{P_\theta : \theta \in \tilde{\Theta}_{\mathcal{H}_k}\}$ *is identified, i.e.* $P_{\theta_1} = P_{\theta_2}$ *implies* $\theta_1 = \theta_2$ *for all* $\theta_1, \theta_2 \in \tilde{\Theta}_{\mathcal{H}_k}$*, and* $|\alpha'_\perp \Gamma \beta_\perp| \neq 0 \ \forall \theta \in \tilde{\Theta}_{\mathcal{H}_k}$.

Proof in Appendix A.2.

It follows from Lemma 3.1 that once the parameter d is fixed to d_0, then all the FCVAR$_{d,b}$ parameters are uniquely identified for any lag $k > k_0$. Under the constraint $d = d_0$, the profile log-likelihood function $\ell_T(\psi)$ only varies with respect to b and it has an unique maximum around b_0. Interestingly, Lemma 3.1 provides theoretical support to the procedure, adopted in Bollerslev et al. (2013) and Caporin et al. (2013), of estimating the FCVAR$_{d,b}$ model by restricting the fractional parameter d to a constant value and by maximizing the profile log-likelihood function with respect to b only. Figure 6 reports the value of the *sliced* profile log-likelihood with respect to different values of b, when the parameter d is kept fixed to the true value $d_0 = 1$. It clearly emerges that,

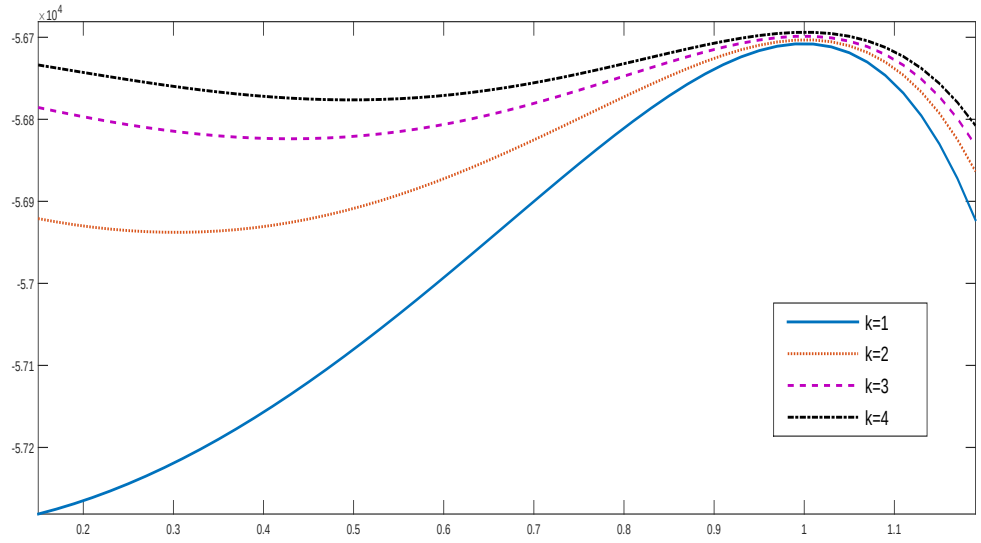


Figure 6: Figure reports the values of the expected profile log-likelihood, $\bar{l}(\psi)$, for different values of $b \in [0.1, 1.2]$ (x-axis) when $d = d_0 = 1$. The sample size is $T = 20000$ and $k_0 = 0$, while $\mathcal{H}_k$ with $k = 1, 2, 3, 4$ is estimated. The parameters of the DGP are $d_0 = b_0 = 1$, $\beta_0 = [1, -1]'$, $\alpha_0 = [-0.5, 0.5]'$.

irrespectively of the choice of $k > k_0$, the profile log-likelihood function is uniquely maximized around b_0. This is a direct consequence of Lemma 3.1. Figure B.3 in the Appendix confirms this result also when $b_0 < d_0$. As expected the value of the *sliced* profile log-likelihood at the optimum is the highest for the model with $k = 4$ lags in both figures, since the model $\mathcal{H}_4$ nests all the other models with $k < 4$. However, the profile log-likelihood function becomes very flat when k increases. This is due to the efficiency loss caused by the inclusion in the model $\mathcal{H}_k$ of matrices of parameters, Γ_j, $j > k_0$, that should be theoretically excluded. This may generate a problem of

weak identification analogous to the one discussed in Section 3.1.

Since there exists an unique maximum of the profile log-likelihood function when d is restricted to d_0, then the asymptotic properties found in Johansen and Nielsen (2012) would still hold. However, since d_0 is unknown in practice, we rely on a constrained optimization method which sets to zero the probability of selecting a maximum outside a given interval for the parameter d. This means that the lower bound $\delta_{\min}$ must be determined such that the optimization of the profile log-likelihood is performed in an area which contains only one maximum. In the following, we illustrate a simple and direct way to select $\delta_{\min}$ in a data-driven fashion. In principle, any semi-parametric estimator of the fractional order of the series, e.g. the exact local Whittle estimator of Shimotsu and Phillips (2005), could be adopted to determine the fractional order of the system and a value for $\delta_{\min}$ could be easily determined by setting a lower bound based on the point estimate. Unfortunately, a multivariate version of the exact local Whittle in which all the processes share the same degree of fractional integration is not yet available in the literature. Indeed, under the assumption of fractional cointegration the multivariate log-likelihood of the model cannot be determined due to the singularity of the coherence matrix at the origin, see the discussion in Nielsen and Shimotsu (2007) among others. Similarly to Nielsen and Shimotsu (2007), we therefore recommend to obtain a semi-parametric estimate of d as

$$\tilde{d} = \frac{1}{p}\sum_{i=1}^{p} \tilde{d}_i \tag{14}$$

where $\tilde{d}_i$ is the univariate exact local Whittle estimate of the parameter d on the i-th series. The exact local Whittle is defined as

$$\tilde{d}_i = \arg\min_{d\in\mathbb{D}} R(d_i, X_{t,i}) \quad i = 1, ..., p \tag{15}$$

with

$$R(d_i, X_{t,i}) = \frac{1}{m}\sum_{j=1}^{m} \log\left(\lambda_j^{-2d_i}\right) + \log\left(\frac{1}{m}\sum_{j=1}^{m} I_{\Delta^{d_i} X_{t,i}}(\lambda_j)\right), \tag{16}$$

where $I_{\Delta^{d_i} X_{t,i}}(\lambda_j)$ is the periodogram of the fractional difference of the series $X_{t,i}$ evaluated at the Fourier frequency λ_j, where the number of frequencies used in the estimation is m and $\mathbb{D}$ is the admissible set of values of d.[6] Under Assumptions 1-5 of Shimotsu and Phillips (2005), $\tilde{d}_i$ is a consistent estimator of d and asymptotically Gaussian with

$$\sqrt{m}(\tilde{d}_i - d_0) \xrightarrow{d} N\left(0, \frac{1}{4}\right). \tag{17}$$

[6] Shimotsu and Phillips (2005) define $\mathbb{D}$ in terms of an upper and lower bound for the parameter d_i, where the length of the interval is smaller or equal to $\frac{9}{2}$. This defines a very large interval of possible values of d, such that we can assume that the true d_0 always belongs to $\mathbb{D}$ without loss of generality.

where the asymptotic variance does not depend on any nuisance parameter and the rate of convergence depends on m. Therefore, once $\tilde{d}$ is estimated, then $\delta_{\min}$ can be determined as

$$\delta_{\min} = \tilde{d} - \gamma \cdot S.E.(\tilde{d}) \tag{18}$$

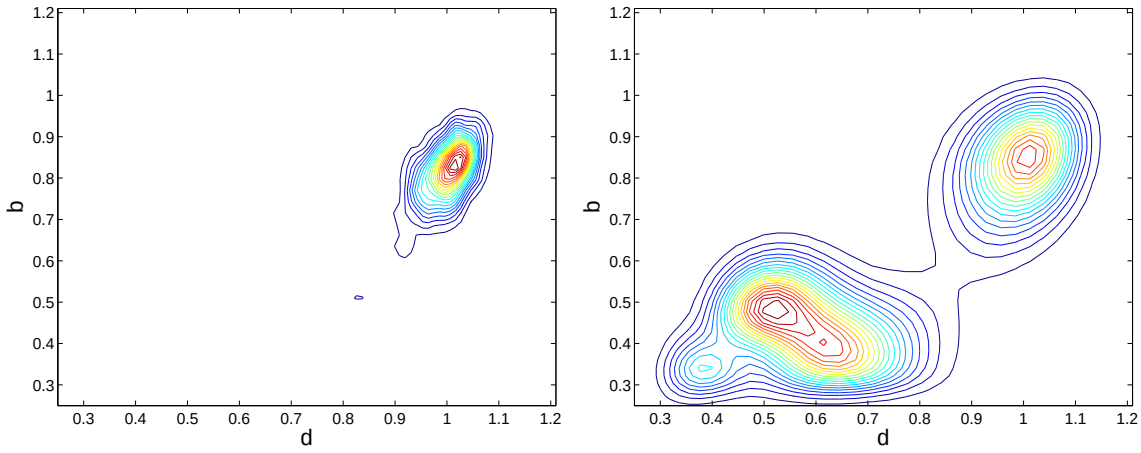
where $S.E.(\tilde{d})$ is the standard-error of the estimator $\tilde{d}$, and γ a positive constant. Following the results of Shimotsu and Phillips (2005), setting $\gamma = 2$ would roughly correspond to a choice of $\delta_{\min}$ associated to the lower bound of a 97.5% confidence interval around the preliminary estimate $\tilde{d}$. Alternatively, you could simply restrict the parameter d in the FCVAR$_{d,b}$ model to the point estimate $\tilde{d}$, obtained with the exact local Whittle estimator. However, next paragraph shows that imposing the lower bound constraint in (13) is sufficient to solve the identification problem with a very mild restriction on the parameter space.

3.2.1 Monte Carlo simulations

In this paragraph, we discuss the results of a number of Monte Carlo simulations to support the need for the approach based on the constrained log-likelihood outlined in (13) as opposed to the unconstrained one when the number of lags is unknown. Figure 7 reports the contour plot of the Monte Carlo estimates of the parameters d and b when a sample of $T = 2500$ observations is generated by the following bivariate FCVAR$_{d,b}$ model

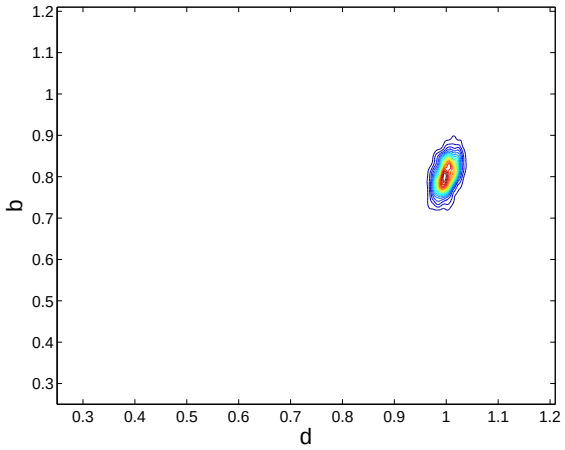
$$\Delta^{d_0} X_t = \alpha_0 \beta_0' \Delta^{d_0 - b_0} L_{b_0} X_t + \varepsilon_t \quad t = 1, \ldots, T \tag{19}$$

where $d_0 = 1$ and $b_0 = 0.8$. For each generated sample, the model $\mathcal{H}_2$ is estimated on the data. According to Proposition 2.2, three equivalent models can be found associated to different combinations of d and b, i.e. $\psi_0 = [1, 0.8]$, $\psi_1 = [0.6, 0.4]$ and $\psi_2 = [0.47, 0.27]$. From Panel a) of Figure 7 it clearly emerges that maximizing the constrained log-likelihood function (13) solves the identification problem discussed above. Indeed, almost the entire probability mass of $\hat{\psi}$, based on $M = 1000$ Monte Carlo estimates, is concentrated around ψ_0. Only in a very limited number of cases the estimates are located around [0.8,0.5], and this could be attributed to the variability of the estimates in finite samples. Instead, when the optimal parameters d and b are found by maximizing the unrestricted likelihood function, see Panel b), a large portion of the probability mass is located away from $\psi_0 = [1, 0.8]$. In particular, when the profile log-likelihood function is not constrained, the bivariate distribution of $\hat{\psi}$ is clearly multi-modal, as a consequence of the lack of identification as outlined in Proposition 2.2. For comparison, Figure 8 reports the distribution of $\hat{\psi}$ when the number of lags is correctly specified, i.e. $k = 0$. Not surprisingly, the distribution of $\hat{\psi}$ is well centered around ψ_0, and the estimates are more efficient than those obtained with $k > 0$ since fewer FCVAR$_{d,b}$ parameters must be estimated under correct lag specification. However, k_0 is unknown in practice and typically needs to be determined by a general-to-specific sequence of LR tests. In Section 4.2 we discuss the nesting structure of the FCVAR$_{d,b}$ model under unknown cointegration rank and lag-length and the optimal sequence of LR tests when the parameter space of d is properly restricted.



(a) Constrained (b) Unconstrained

Figure 7: Figure reports the contour plot of $M = 1000$ Monte Carlo estimates of the parameters d (x-axis) and b (y-axis) when a sample of $T = 2500$ observations is generated by a FCVAR$_{d,b}$ model with $k_0 = 0$, $d_0 = 1$, $b_0 = 0.8$ and the cointegration vectors given by $\beta_0 = [1, -1]'$ and $\alpha_0 = [-0.5, 0.5]'$. Model $\mathcal{H}_2$ is estimated on the data. Panel a) is relative to the estimates based on the constrained log-likelihood (13) where $\gamma = 2$ and $m = T^{0.6}$. Panel b) reports the contour plot for the unrestricted estimates.



(a) Correct Number of Lags

Figure 8: Figure reports the contour plot of $M = 1000$ Monte Carlo estimates of the parameters d (x-axis) and b (y-axis) when a sample of $T = 2500$ observations is generated by a FCVAR$_{d,b}$ model with $k_0 = 0$, $d_0 = 1$, $b_0 = 0.8$ and the cointegration vectors given by $\beta_0 = [1, -1]'$ and $\alpha_0 = [-0.5, 0.5]'$. Model $\mathcal{H}_0$ is estimated on the data.

Figures B.4-B.8 in Appendix B highlight the robustness of the constrained likelihood approach for different sample sizes and different combinations of k_0 and k. When T increases, the estimates based on the unconstrained likelihood still display the identification problem, while the constrained estimates are all centered around d_0 and b_0, see Figure B.4. When $T = 1000$, most of the probability mass is again concentrated around d_0 and b_0 although it is more dispersed, see Figures B.5 and B.6. Finally, the results do not qualitatively change when data are generated under $\mathcal{H}_1$ with

$\Gamma_1 = \left[\begin{smallmatrix} 0.3 & -0.2 \\ -0.4 & 0.5 \end{smallmatrix}\right]$ and model $\mathcal{H}_3$ is estimated, see Figure B.7. As expected, the estimates based on the unconstrained likelihood are clearly bimodal, since two equivalent sub-models of $\mathcal{H}_3$ can be found associated to $k_0 = 1$, see Table 1. Instead, the estimates based on the constrained likelihood are again unimodal and centered around the true values of d and b. Finally, the quality of the constrained estimates slightly deteriorates when $d_0 - b_0 \approx 0.5$, see Figure B.8. In this case, the slow convergence rate makes the profile log-likelihood function extremely flat, although the sample size is moderately large, thus generating more dispersed estimates of ψ. However, compared to the unrestricted estimates which are found everywhere in the interval $0.3 < b < d < 1$, the constrained estimates are much more often concentrated in the region around d_0 and b_0.

4 Unknown cointegration rank

In this section, we extend the previous results to the case in which the cointegration rank and the lag-length are both unknown. This is the relevant case in empirical applications, when testing for the presence of a cointegration relationship between two (or more) fractional processes but there is no preliminary information on the optimal choice of k. The unrestricted FCVAR$_{d,b}$ model is formulated as:

$$\mathcal{H}_{r,k}: \quad \Delta^d X_t = \Pi \Delta^{d-b} L_b X_t + \sum_{i=1}^{k} \Gamma_i \Delta^d L_b^i X_t + \varepsilon_t, \tag{20}$$

where $0 \leq r \leq p$ is the rank of the $p \times p$ matrix Π. The parameter space of model $\mathcal{H}_{r,k}$ is

$$\Theta_{\mathcal{H}_{r,k}} = \{\alpha \in \mathbb{R}^{p\times r}, \beta \in \mathbb{R}^{p\times r}, \Gamma_j \in \mathbb{R}^{p\times p}, j = 1, \ldots, k, d \in \mathbb{R}^+, b \in \mathbb{R}^+, d \geq b > 0, \Omega > 0\}.$$

Compared to the parameter space of $\mathcal{H}_k$ in Section 2, the set $\Theta_{\mathcal{H}_{r,k}}$ also contains the cointegration rank, r, among the unknown parameters. For this reason, model $\mathcal{H}_{r,k}$ exhibits further identification issues than those illustrated in Section 2.

Example 2: Consider the model with $k = 1$ lags and rank $0 \leq r \leq p$, given by

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

where the set of parameters is $\theta = vec(d, b, \Pi, \Gamma_1)$.

Examine now the following two sub-models of $\mathcal{H}_{r,1}$. First, model $\mathcal{H}_{p,0}$ is

$$\mathcal{H}_{p,0}: \quad \Delta^{\tilde{d}} X_t = \tilde{\Pi} \Delta^{\tilde{d}-\tilde{b}} L_{\tilde{b}} X_t + \varepsilon_t,$$

with $\tilde{\theta} = vec(\tilde{d}, \tilde{b}, \tilde{\Pi})$ is the set of parameters. Second, model $\mathcal{H}_{0,1}$ is

$$\mathcal{H}_{0,1}: \quad \Delta^{d^*} X_t = \Gamma_1^* \Delta^{d^*} L_{b^*} X_t + \varepsilon_t.$$

where $\theta^* = vec(d^*, b^*, \Gamma_1^*)$ is the set of parameters.[7] Both $\mathcal{H}_{p,0}$ and $\mathcal{H}_{0,1}$ can be written as

$$\left[\Delta^{\tilde{d}-\tilde{b}}(-\tilde{\Pi}) + \Delta^{\tilde{d}}(I_p + \tilde{\Pi})\right] X_t = \varepsilon_t, \tag{21}$$

and

$$\left[\Delta^{d^*}(I - \Gamma_1^*) + \Delta^{d^*+b^*}(\Gamma_1^*)\right] X_t = \varepsilon_t. \tag{22}$$

Imposing the restrictions $\tilde{d} = d^* + b^*$, $\tilde{b} = b^*$ and $-\tilde{\Pi} = I_p - \Gamma_1^*$ on model $\mathcal{H}_{p,0}$ in (21), it results that $\mathcal{H}_{p,0}$ and $\mathcal{H}_{0,1}$ are equivalent. Indeed, the probability densities are

$$p_{\mathcal{H}_{p,0}}(X_1, \ldots, X_T; \tilde{\theta} | X_0, X_{-1} \ldots) = p_{\mathcal{H}_{0,1}}(X_1, \ldots, X_T; \theta^* | X_0, X_{-1}, \ldots), \tag{23}$$

when $\tilde{\theta} = vec(d^* + b^*, b^*, \Gamma_1^* - I_p, 0)$ and $\theta^* = vec(d^*, b^*, 0, \Gamma_1^*)$.

However, the sub-model $\mathcal{H}_{0,1}$ is not always a re-parametrization of $\mathcal{H}_{p,0}$. Indeed, applying the restrictions $d^* = \tilde{d} - \tilde{b}$, $b^* = \tilde{b}$ and $\Gamma_1^* = I_p + \tilde{\Pi}$ on model $\mathcal{H}_{0,1}$ in (22), it follows that

$$p_{\mathcal{H}_{p,0}}(X_1, \ldots, X_T; \tilde{\theta} | X_0, X_{-1}, \ldots) = p_{\mathcal{H}_{0,1}}(X_1, \ldots, X_T; \theta^* | X_0, X_{-1}, \ldots), \tag{24}$$

where $\tilde{\theta} = vec(\tilde{d}, \tilde{b}, \tilde{\Pi}, 0)$ and $\theta^* = vec(\tilde{d} - \tilde{b}, \tilde{b}, 0, I_p + \tilde{\Pi})$. However, the equality (24) holds if and only if $\tilde{d} - \tilde{b} \geq \tilde{b} > 0$, i.e. $\tilde{d} \geq 2\tilde{b}$. This implies that $\mathcal{H}_{0,1} = \mathcal{H}_{p,0} \cap \left\{\tilde{d} \geq 2\tilde{b}\right\}$. Hence, $\mathcal{H}_{0,1} \subseteq \mathcal{H}_{p,0}$. The next proposition extends this example for any combination of k and r.

Proposition 4.1 *Consider an unrestricted FCVAR$_{d,b}$ model*

$$\mathcal{H}_{r,k}: \quad \Delta^d X_t = \Pi \Delta^{d-b} L_b X_t + \sum_{j=1}^{k} \Gamma_j \Delta^{d-b} L_b X_t + \varepsilon_t \tag{25}$$

where $0 \leq r \leq p$ is the rank of the matrix Π and k is the number of lags. Consider the following sub-models of $\mathcal{H}_{r,k}$: $\mathcal{H}_{p,k-1}$ with parameter set $\tilde{\theta} = vec(\tilde{d}, \tilde{b}, \tilde{\Pi}, \tilde{\Gamma}_1, ..., \tilde{\Gamma}_{k-1}, \tilde{\Omega})$, and $\mathcal{H}_{0,k}$ with parameter set $\theta^* = vec(d^*, b^*, \Gamma_1^*, ..., \Gamma_k^*, \Omega^*)$.

i) For any $k > 0$, model $\mathcal{H}_{0,k}$ is equivalent to $\mathcal{H}_{p,k-1}$ if the condition $\tilde{d} \geq 2\tilde{b}$ imposed on model $\mathcal{H}_{p,k-1}$ is satisfied. Hence $\mathcal{H}_{0,k} = \mathcal{H}_{p,k-1} \cap \left\{\tilde{d} \geq 2\tilde{b}\right\}$.

[7] Note that to maintain the notation as light as possible and avoid the double subscript for the parameters, we use $\tilde{\theta}$ and θ^*, instead of $\theta_{p,0}$ and $\theta_{1,0}$, to indicate the parameter sets of $\mathcal{H}_{p,0}$ and $\mathcal{H}_{0,1}$ respectively.

ii) The nesting structure of the FCVAR$_{d,b}$ model is represented by the following scheme:

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

Proof in Appendix A.3.

It follows from Proposition 4.1i) that model $\mathcal{H}_{0,k}$ can always be re-parametrized as model $\mathcal{H}_{p,k-1}$. On the other hand, model $\mathcal{H}_{p,k-1}$ can be formulated as $\mathcal{H}_{0,k}$ only when the condition $\tilde{d} \geq 2\tilde{b}$ on model $\mathcal{H}_{p,k-1}$ holds. This leads to the peculiar nesting structure displayed in Proposition 4.1.*ii*). Notably the interpretation of the two models $\mathcal{H}_{p,k-1}$ and $\mathcal{H}_{0,k}$ is slightly different, although they are equivalent descriptions of the data. In model $\mathcal{H}_{p,k-1}$, the the process X_t has p non-common stochastic trends fractional order $\tilde{d} - \tilde{b}$. Instead, in model $\mathcal{H}_{0,k}$, then the process X_t has p non-common stochastic trends fractional order d^*.

The following Corollary shows that indeterminacy between cointegration rank and lag-length is not limited to $\mathcal{H}_{p,k-1}$ and $\mathcal{H}_{0,k}$, but it can be extended to any cointegration rank $0 < s < p$.

Corollary 4.2 *For any $k > 0$, model $\mathcal{H}_{s,k-1}$ with $0 < s < p$ and $\tilde{d} \geq 2\tilde{b}$ is equivalent to $\mathcal{H}_{0,k}$, if and only if the matrix $\Gamma^* = I_p - \sum_{j=1}^{k} \Gamma_j^*$ in model $\mathcal{H}_{0,k}$ has rank equal to s.*

Proof in Appendix A.4.

In other words, if the matrix $\Gamma^* = I_p - \sum_{j=1}^{k} \Gamma_j^*$ in $\mathcal{H}_{0,k}$ has reduced rank of order $0 < s < p$, the models $\mathcal{H}_{s,k-1}$ and $\mathcal{H}_{0,k}$ are equivalent under $\tilde{d} \geq 2\tilde{b}$ in $\mathcal{H}_{s,k-1}$. This means that $\mathcal{H}_{0,k} \subseteq \mathcal{H}_{s,k-1}$ for any $0 < s \leq p$, if $rank(\Gamma) = s$.

4.1 Univariate model

A similar identification problem, due to indeterminacy between d, b and k, arises also in the univariate FAR(k) model studied in Johansen and Nielsen (2010)

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

where Y_t is an univariate process and π is a scalar. Following the same procedure of the proof of Proposition 4.1, it follows that $\mathcal{H}_{0,k} = \mathcal{H}_{1,k-1} \cap \left\{ \tilde{d} \geq 2\tilde{b} \right\}$, where $\mathcal{H}_{0,k}$ defines here the FAR model with $\pi = 0$ and k lags, while $\mathcal{H}_{1,k-1}$ defines the FAR model with $\pi \neq 0$ and $k-1$ lags. Therefore,

the FAR(k) model has the following circular nesting structure:

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

In Johansen and Nielsen (2010), the theoretical results are obtained under the maintained assumption that the true number of lags k_0 is known.

4.2 Model selection under unknown rank and lag-length

The peculiar nesting structure of the FCVAR$_{d,b}$ obviously impacts on the joint selection of the number of lags and the cointegration rank. Indeed, the likelihood ratio statistic for cointegration rank r, denoted as $LR_{r,k} := -2\log LR(\mathcal{H}_{r,k}|\mathcal{H}_{p,k})$, see Johansen and Nielsen (2012, p.2698), is given by

$$-2\log LR(\mathcal{H}_{r,k}|\mathcal{H}_{p,k}) = T(\ell_T^{(r,k)}(\hat{d}_{r,k},\hat{b}_{r,k}) - \ell_T^{(p,k)}(\hat{d}_{p,k},\hat{b}_{p,k})), \tag{26}$$

where $\ell_T^{(r,k)}$ is the profile log-likelihood of the 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 $\ell_T^{(r,k)}$. The asymptotic properties of the $LR_{r,k}$ test, under the maintained assumption of correct specification of the lag-length, i.e. $k = k_0$, are provided in Johansen and Nielsen (2012). Unfortunately, the values of the profile log-likelihoods $\ell_T^{(0,k)}(\hat{d}_{0,k},\hat{b}_{0,k})$ and $\ell_T^{(p,k-1)}(\hat{d}_{p,k-1},\hat{b}_{p,k-1})$ are equal when $\tilde{d} \geq 2\tilde{b}$ in model $\mathcal{H}_{p,k-1}$, and the number of the parameters of the model $\mathcal{H}_{p,k-1}$ is the same as in $\mathcal{H}_{0,k}$. Hence, the equality of $\ell_T^{(0,k)}(\hat{d}_{0,k},\hat{b}_{0,k})$ and $\ell_T^{(p,k-1)}(\hat{d}_{p,k-1},\hat{b}_{p,k-1})$ influences the general-to-specific sequence of tests for the joint selection of the cointegration rank and the lag-length. Indeed, assuming that the general-to-specific procedure for the optimal lag selection terminates in $\mathcal{H}_{p,k-1}$, then it would be impossible to know whether the optimal model is $\mathcal{H}_{p,k-1}$ or $\mathcal{H}_{0,k}$ if the estimates $\hat{d}_{p,k-1}$ and $\hat{b}_{p,k-1}$ are such that $\hat{d}_{p,k-1} \geq 2\hat{b}_{p,k-1}$.

Therefore, a problem of joint selection of k and $r > 0$ arises in the FCVAR$_{d,b}$ when the cointegration rank is unknown and potentially equal to 0 or p. Moreover, under $\mathcal{H}_{0,k}$ with $k > 0$, the parameter b is defined but it does not have the usual interpretation as *cointegration gap*. A test for the null hypothesis that $r = 0$ has been proposed by Łasak (2010) and extended in Łasak and Velasco (2015) to allow for multiple degrees of fractional cointegration. Łasak (2010) derives the asymptotic distribution of the maximum eigenvalue and trace tests for the null hypothesis of absence of cointegration relation in the Granger (1986) system

$$\mathcal{H}_k: \quad \Delta^d X_t = \alpha\beta'\Delta^{d-b}L_b X_t + \sum_{i=1}^{k}\Gamma_i\Delta^d X_{t-i} + \varepsilon_t \quad \varepsilon_t \sim iidN(0,\Omega), \tag{27}$$

under the assumption that $d = 1$. It should be noted that in the FVECM model of Granger

(1986), the problem of identification discussed above does not arise since the operator L_b does not enter in the short-run terms. Indeed, under $r = 0$, the parameter b is not defined, implying that $\mathcal{H}_{0,k}$ and $\mathcal{H}_{p,k-1}$ are distinct models in the FVECM framework. In other words, the problem of joint indeterminacy between cointegration rank and number of lags does not affect model (27). However, as noted by Johansen (2008), it is not possible to obtain a Granger representation theorem for fractionally cointegrated processes under the FVECM representation. Łasak and Velasco (2015) guarantee a Granger representation theorem also under short-run dynamics by assuming that the *pre-whitened* series $X_t^* = A(L)X_t$ follows a FVECM with $k = 0$.[8] Alternatively, a solution to the indeterminacy in the FCVAR$_{d,b}$ framework is to rely on a preliminary estimate of the cointegration rank based on a frequency domain procedure, following for example the testing procedure of Nielsen and Shimotsu (2007). Instead, in the section below, we show that it is sufficient to impose a constraint the fractional parameter d to solve in the problem of indeterminacy of cointegration rank and lag-length.

4.2.1 Model selection with an identification restriction

Unfortunately, a solution to the joint indeterminacy of cointegration rank and lag-length is not available within the unrestricted FCVAR$_{d,b}$ framework. However, a simple solution to the identification problem caused by the indeterminacy of cointegration rank and lag-length can be achieved by a suitable restriction of the parameter space of d. Consider the model with unknown rank and unknown lag structure. The model can be expressed by the parameter set $\Theta_{r,k} = \{d_0 \in \mathbb{R}^+, b \in (0, d_0], \Gamma_j \in \mathbb{R}^{p\times p}, j = 1, \dots, k, \alpha \in \mathbb{R}^{p\times r}, \beta \in \mathbb{R}^{p\times r}, \Omega > 0\}$ where $0 \leq r \leq p$ and $k \geq 0$ are unknown. The following lemma holds

Lemma 4.3 *Let* $\tilde{\Theta}_{\mathcal{H}_{r,k}} = \{d = d_0, b \in [0, d_0], \alpha \in \mathbb{R}^{p\times r}, \beta \in \mathbb{R}^{p\times r}, \Gamma_j \in \mathbb{R}^{p\times p}, j = 1, \dots, k; \Omega > 0\}$ *be the restricted parameter space of model* $\Theta_{\mathcal{H}_{r,k}}$ *with* $d = d_0 \in \mathbb{R}^+$ *for* $0 \leq r \leq p$ *and* $k \geq 0$*, then the nesting structure for the statistical models* $\mathcal{P} = \{P_\theta : \theta \in \Theta_{r,k}\}_{k=0,1,\dots}^{r=0,\dots,p}$ *can be written as*

$$
\begin{array}{ccccccc}
\mathcal{H}_{0,0} & \subset & \mathcal{H}_{0,1} & \subset & \cdots & \subset & \mathcal{H}_{0,k} \\
\cap & & \cap & & & & \cap \\
\mathcal{H}_{1,0} & \subset & \mathcal{H}_{1,1} & \subset & \cdots & \subset & \mathcal{H}_{1,k} \\
\vdots & & \vdots & & & & \vdots \\
\cap & & \cap & & & & \cap \\
\mathcal{H}_{p,0} & \subset & \mathcal{H}_{p,1} & \subset & \cdots & \subset & \mathcal{H}_{p,k}
\end{array}
$$

Proof in Appendix A.5.

When $d = d_0$ is fixed, Lemma 4.3 proves that the FCVAR$_{d,b}$ has a nesting structure that does not exhibit the problem outlined above, since $\mathcal{H}_{p,k-1}$ and $\mathcal{H}_{0,k}$ are two distinct models. Analogously to the discussion in Section 3.2, we suggest that the estimates of $d_{r,k}$ and $b_{r,k}$, for any $0 \leq r \leq p$ and $k \geq 0$, must be the solutions of the following constrained maximum likelihood problem

[8] Only when $k = 0$, the FVECM and the FCVAR$_{d,b}$ model are equivalent, meaning that in this case also the FVECM model allows for a Granger representation.

$$\hat{\psi}_{r,k} = \arg\max_{\psi} \ell_T^{(r,k)}(\psi_{r,k}), \quad \text{s.t. } d_{r,k} \geq \delta_{\min} \tag{28}$$

where the lower bound on the parameter $d_{r,k}$, $\delta_{\min}$, can be determined by a preliminary estimate of the fractional order of the process. Therefore, under the constraint $d_{r,k} \geq \delta_{\min}$, we can test $\mathcal{H}_{p,k}$ against $\mathcal{H}_{p,k-1}$, without the risk of having an equivalent parametrization in $\mathcal{H}_{0,k}$ under the null hypothesis. In particular, the general-to-specific sequence of LR tests consists of iterating the tests $LR_{p,k-1} := -2\log LR(\mathcal{H}_{p,k-1}|\mathcal{H}_{p,k})$ over k with fixed p (full rank) until the null hypothesis is rejected in k^*. Subsequently, the cointegration rank, i.e. the rank of the matrix Π in model (20), can be determined by a sequence of LR tests, LR_{r,k^*}, as in (26), over $r \in [0, p]$ with k fixed to k^*. It is important to stress that $\delta_{\min}$ does not depend on r and k so that it can be determined before the general-to-specific sequence of LR tests for the determination of lag-length and cointegration rank.

5 Conclusion

This paper discussed in detail some identification problems that affect the FCVAR$_{d,b}$ model of Johansen (2008). The main finding is that the fractional parameters 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. Furthermore, the simulations reveal a problem of weak identification, characterized by the presence of local and global maxima of the profile likelihood function in finite samples. We also prove that it is sufficient to restrict d to d_0 to solve the identification problem. However, since d_0 is unknown, we impose a lower-bound restriction on d, where the lower bound is determined on the basis of a preliminary semiparametric estimate of d_0. This imposes the mildest restriction on the parameter space of the FCVAR$_{d,b}$ model. The Monte Carlo simulations show that the estimates of the model parameters are unimodal and centered around the true values in most cases. It is also proved that model $\mathcal{H}_{0,k}$ is equivalent to model $\mathcal{H}_{p,k-1}$ under certain conditions on d and b. Unfortunately, the $\mathcal{F}(d)$ condition does not provide any information for the identification in this case, but it is again sufficient to impose a suitable lower bound restriction on the parameter space of d to solve this identification problem and retrieve a nesting structure of FCVAR$_{d,b}$ model that allows testing for the unknown lag-length and cointegration rank in the standard general-to-specific fashion.

Acknowledgements. The authors are grateful to Niels Haldrup, Søren Johansen, Katarzyna

Łasak and Morten Nielsen for their suggestions that improved the quality of this work. The authors are also grateful to an anonymous referee for providing insightful comments. The authors would like to thank also James MacKinnon, Rocco Mosconi, Paolo Paruolo, the participants to the Third Long Memory Symposium (Aarhus 2013), the participants to the CFE'2013 conference (London 2013), and the seminar participants at Queen's University and at Bologna University for helpful comments.

References

Avarucci, M. and Velasco, C. (2009). A Wald test for the cointegration rank in nonstationary fractional systems. *Journal of Econometrics*, 151(2):178–189.

Bollerslev, T., Osterrieder, D., Sizova, N., and Tauchen, G. (2013). Risk and return: Long-run relations, fractional cointegration, and return predictability. *Journal of Financial Economics*, 108(2):409–424.

Breitung, J. and Hassler, U. (2002). Inference on the cointegration rank in fractionally integrated processes. *Journal of Econometrics*, 110(2):167–185.

Caporin, M., Ranaldo, A., and Santucci de Magistris, P. (2013). On the predictability of stock prices: A case for high and low prices. *Journal of Banking & Finance*, 37(12):5132–5146.

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

Chen, W. and Hurvich, C. (2006). Semiparametric estimation of fractional cointegrating subspaces. *Annals of Statistics*, 34:2939–2979.

Christensen, B. J. and Nielsen, M. Ø. (2006). Asymptotic normality of narrow-band least squares in the stationary fractional cointegration model and volatility forecasting. *Journal of Econometrics*, 133(1):343–371.

Franchi, M. (2010). A representation theory for polynomial cofractionality in vector autoregressive models. *Econometric Theory*, 26(04):1201–1217.

Granger, C. W. J. (1986). Developments in the study of cointegrated economic variables. *Oxford Bulletin of Economics and Statistics*, 48(3):213–28.

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

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

Johansen, S. (1995a). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press, Oxford.

Johansen, S. (1995b). A stastistical analysis of cointegration for I(2) variables. *Econometric Theory*, 11(01):25–59.

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. (2009). *Cointegration. Overview and Development*, chapter IV, pages 671–692. Springer.

Johansen, S. (2010). Some identification problems in the cointegrated vector autoregressive model. *Journal of Econometrics*, 158(2):262–273.

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

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

Łasak, K. (2010). Likelihood based testing for no fractional cointegration. *Journal of Econometrics*, 158(1):67–77.

Łasak, K. and Velasco, C. (2015). Fractional cointegration rank estimation. *Journal of Business & Economic Statistics*, 33(2):241–254.

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.

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

Shimotsu, K. and Phillips, P. C. (2005). Exact local whittle estimation of fractional integration. *Annals of Statistics*, 33(4):1890–1933.

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. University of Regensburg Working Papers in Business, Economics and Management Information Systems 471, University of Regensburg, Department of Economics.

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.

A Proofs

A.1 Proof of Proposition 2.2

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

$$\sum_{i=-1}^{k_0} \Psi_{i,0} \Delta^{d_0+ib_0} X_t = \varepsilon_t, \tag{29}$$

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

$$\sum_{i=-1}^{k} \Psi_i \Delta^{d+ib} X_t = \varepsilon_t. \tag{30}$$

It is possible to show, that, for a given k_0, m sub-models equivalent to the model in (29) 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 for

$$\begin{aligned} \Psi_{-1} &= \Psi_{-1,0} \quad \text{corresponding to} \quad d - b = d_0 - b_0 \\ \Psi_{(\ell+1)(j+1)-1} &= \Psi_{\ell,0} \qquad \text{corresponding to} \quad d + [(\ell+1)(j+1) - 1]b = d_0 + \ell b_0, \\ &\qquad\qquad \text{for} \quad \ell = 0, \ldots, k_0 \quad j = 0, 1, \ldots, m-1 \\ \Psi_s &= 0 \qquad \text{for} \quad s \neq (\ell+1)(j+1) - 1, \\ &\qquad\qquad \text{and} \quad \ell = 0, \ldots, k_0 \quad j = 0, 1, \ldots, m-1. \end{aligned} \tag{31}$$

The matrices $\Psi_{-1,0} = -\alpha_0 \beta_0'$ 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 $d_0 - b_0 = d - b$ in all equivalent sub-models. For a given $j > 0$, a system of $k_0 + 2$ equations (31) 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 $\alpha\beta' + \Gamma = \Psi_0 = 0$. It follows that 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 restrictions to be imposed on Ψ_i that satisfies the system in (31) is $\lfloor \frac{k+1}{k_0+1} \rfloor$. Hence, the number of equivalent sub-models is $m = \lfloor \frac{k+1}{k_0+1} \rfloor$.

■

A.2 Proof of Lemma 3.1

Consider two models $\mathcal{H}_k^1$ and $\mathcal{H}_k^2$ defined in $\tilde{\Theta}_{\mathcal{H}_k}$, given by

$$\sum_{j=-1}^{k} \Delta^{d_0+jb_1} \Psi_j^1 X_t = \varepsilon_t \quad \text{and} \quad \sum_{j=-1}^{k} \Delta^{d_0+jb_2} \Psi_j^2 X_t = \varepsilon_t$$

with $d_0 \geq b_1 > 0$ and $d_0 \geq b_2 > 0$. We want to prove that $\mathcal{H}_k^1$ and $\mathcal{H}_k^2$ are equal if only if $b_1 = b_2$ and $\Psi_j^1 = \Psi_j^2, j = 1, \ldots, k$ and $\Omega_1 = \Omega_2$.

Given that P_θ is Gaussian for all $\theta \in \tilde{\Theta}_{\mathcal{H}_k}$ we should check that the characteristic polynomials

$$\Pi_i(z) = \sum_{j=-1}^{k} (1-z)^{d_0+jb_i} \Psi_j^i, \quad i = 1, 2$$

are equal. They are equal if

$$(1-z)^{d_0+jb_1} = (1-z)^{d_0+jb_2} \iff (1-z)^{b_1} = (1-z)^{b_2} \iff b_1 = b_2, \quad \forall j = -1, \ldots, k$$

and

$$\Psi_j^1 = \Psi_j^2, \quad \forall j = -1, \ldots, k$$

Finally, the variance of the innovations are $\Omega_1 = \Omega_2$ by construction since the error terms ϵ_t is the same in $\mathcal{H}_k^1$ and $\mathcal{H}_k^2$. Therefore, the statistical model $\mathcal{P} = \{P_\theta : \theta \in \tilde{\Theta}_{\mathcal{H}_k}\}$ is identified.

A.3 Proof of Proposition 4.1

The unrestricted FCVAR$_{d,b}$ model is given by

$$\mathcal{H}_{r,k}: \quad \Delta^d X_t = \Pi \Delta^{d-b} L_b X_t + \sum_{j=1}^{k} \Gamma_j \Delta^{d-b} L_b X_t + \varepsilon_t, \tag{32}$$

where $0 \leq r \leq p$ is the rank of the matrix Π and k is the number of lags. The model in equation (25) can be written as

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

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

Now consider the following sets of restrictions on model (25):

$$\begin{aligned} \mathcal{H}_{p,k-1}: &\quad \Pi \text{ is a } p \times p \text{ matrix and } \Gamma_k = 0 \\ \mathcal{H}_{0,k}: &\quad \Pi{=}0. \end{aligned}$$

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

$$\sum_{i=-1}^{k-1} \tilde{\Psi}_i \Delta^{\tilde{d}+i\tilde{b}} X_t = \varepsilon_t \tag{33}$$

where $\tilde{\Psi}_{-1} = \tilde{\Pi}$, $\tilde{\Psi}_0 = I_p + \tilde{\Pi} - \sum_{i=1}^{k-1} \tilde{\Gamma}_i$ and $\tilde{\Psi}_{k-1} = (-1)^k \tilde{\Gamma}_{k-1}$. The matrices $\tilde{\Pi}$ and $\tilde{\Psi}_i$, $i = -1, ..., k-1$ define the model under the restriction $\mathcal{H}_{p,k-1}$.

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

$$\sum_{i=0}^{k} \Psi_i^* \Delta^{d^*+ib^*} X_t = \varepsilon_t, \tag{34}$$

with $\Psi^*_{-1} = 0$, $\Psi^*_0 = I_p + 0 - \sum_{i=1}^{k} \Gamma^*_i$ and $\Psi^*_k = (-1)^{k+1}\Gamma^*_k$. The matrices Ψ^*_i, $i = -1, ..., k$, define the model under the restriction $\mathcal{H}_{0,k}$.

Imposing the following set of restrictions on the matrices $\tilde{\Psi}_i$ and Ψ^*_i:

$$\begin{cases} \tilde{\Psi}_{-1} = \Psi^*_0 \\ \tilde{\Psi}_0 = \Psi^*_1 \\ \vdots \\ \tilde{\Psi}_{k-1} = \Psi^*_k, \end{cases} \tag{35}$$

it follows that the two models $\mathcal{H}_{p,k-1}$ and $\mathcal{H}_{0,k}$ are equivalent when the system

$$\begin{cases} \tilde{d} - \tilde{b} = d^* \\ \tilde{d} = d^* + b^* \\ \vdots \\ \tilde{d} + (k-1)\tilde{b} = d^* + kb^* \end{cases} \tag{36}$$

has an unique solution. Suppose that the system (36) is solved for $\tilde{d}$ and $\tilde{b}$. The unique solution in this case is $\tilde{d} = d^* + b^*$ and $\tilde{b} = b^*$, which satisfies the condition $\tilde{d} \geq \tilde{b} > 0$. Now suppose that the system (36) is solved for d^* and b^*. The unique solution in this case is $d^* = \tilde{d} - \tilde{b}$ and $b^* = \tilde{b}$, which satisfies the condition $d^* \geq b^* > 0$ if and only if $\tilde{d} \geq 2\tilde{b}$. Therefore, if $\tilde{d} \geq 2\tilde{b}$ it follows that $\mathcal{H}_{0,k} \equiv \mathcal{H}_{p,k-1}$. Hence, $\mathcal{H}_{0,k} \subset \mathcal{H}_{p,k-1}$. ■

A.4 Proof of Corollary 4.2

Using a procedure similar to that adopted in the proof of Proposition 4.1, it is straightforward to show that, when $\tilde{d} \geq 2\tilde{b}$, the model $\mathcal{H}_{s,k-1}$ with $0 < s < p$ and model $\mathcal{H}_{0,k}$ are equivalent if $\Gamma^* = I_p - \sum_{i=1}^{k} \Gamma^*_i = \Psi^*_0$ is a matrix with rank s in model (34) and the restriction $r = s$ is imposed on model (33), so that $\tilde{\Pi} = \alpha\beta'$ where α and β are $p \times s$ matrices. ■

A.5 Proof of Lemma 4.3

Consider the models $\mathcal{H}_{p,k-1}$ and $\mathcal{H}_{0,k}$ for $k = 0, 1, \ldots$ in equations (33)-(34) and impose the constraint $d = d_0$. Then,

$$\mathcal{H}_{p,k-1} : \sum_{i=-1}^{k-1} \tilde{\Psi}_i \Delta^{d_0 + i\tilde{b}} = \varepsilon_t$$

$$\mathcal{H}_{0,k} : \sum_{i=0}^{k} \Psi^*_i \Delta^{d_0 + ib^*} X_t = \varepsilon_t.$$

It follows that $\mathcal{H}_{p,k-1} \cap \mathcal{H}_{0,k} = \emptyset$ because there is no solution to the system of equations 36 when $d = d_0$ is fixed. Therefore, the nesting structure in 4.3 follows. ■

B Additional Figures

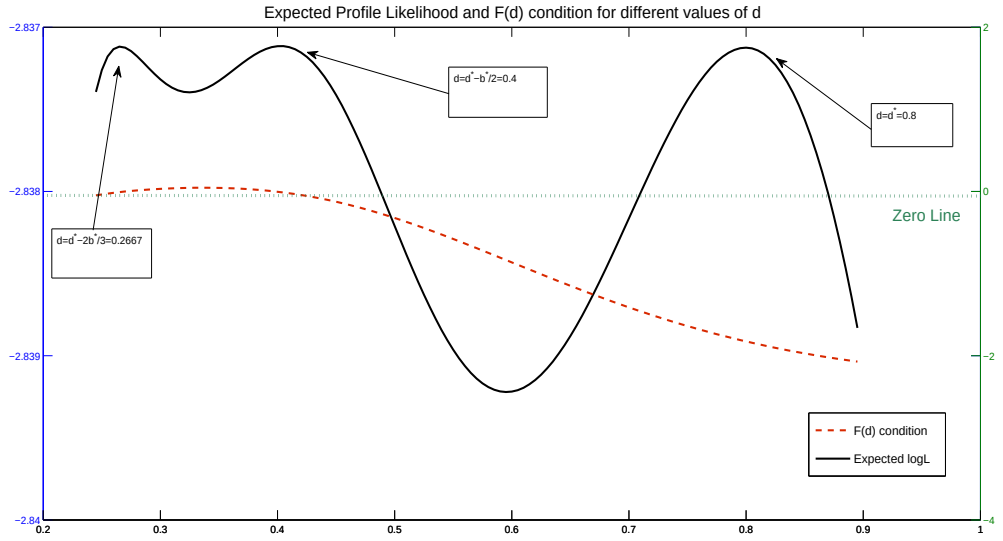


Figure B.1: Figure reports simulated values of $\bar{l}(d)$ and $\bar{\mathcal{F}}(d)$ for different values of $d \in [0.2, 1.2]$ (x-axis). The observations from the DGP are generated with $k_0 = 0$ lags and model $\mathcal{H}_k$ with $k = 2$ lags is estimated. The parameters of the DGP are $d_0 = b_0 = 0.8$, $\beta_0 = [1, -1]'$, $\alpha_0 = [-0.5, 0.5]'$.

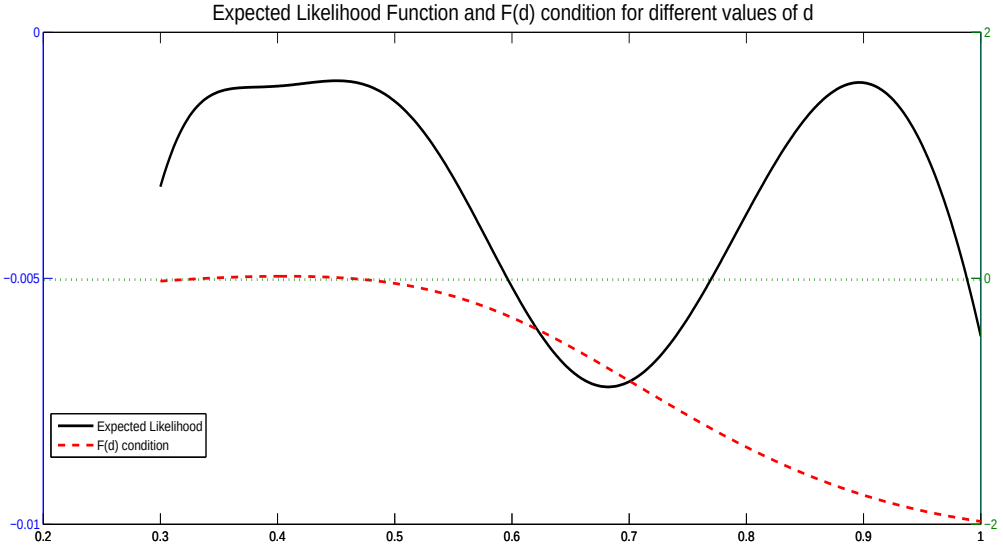


Figure B.2: Figure reports simulated values of $\bar{l}(d)$ and $\bar{\mathcal{F}}(d)$ for different values of $d \in [0.3, 0.8]$ (x-axis). The observations from the DGP are generated with $k_0 = 1$ lags and model $\mathcal{H}_k$ with $k = 3$ lags is estimated. The parameters of the DGP are $d_0 = b_0 = 0.8$, $\beta_0 = [1, -1]'$, $\alpha_0 = [-0.5, 0.5]'$,and $\Gamma_1 = \begin{bmatrix} 0.3 & -0.2 \\ 0.4 & -0.5 \end{bmatrix}$.

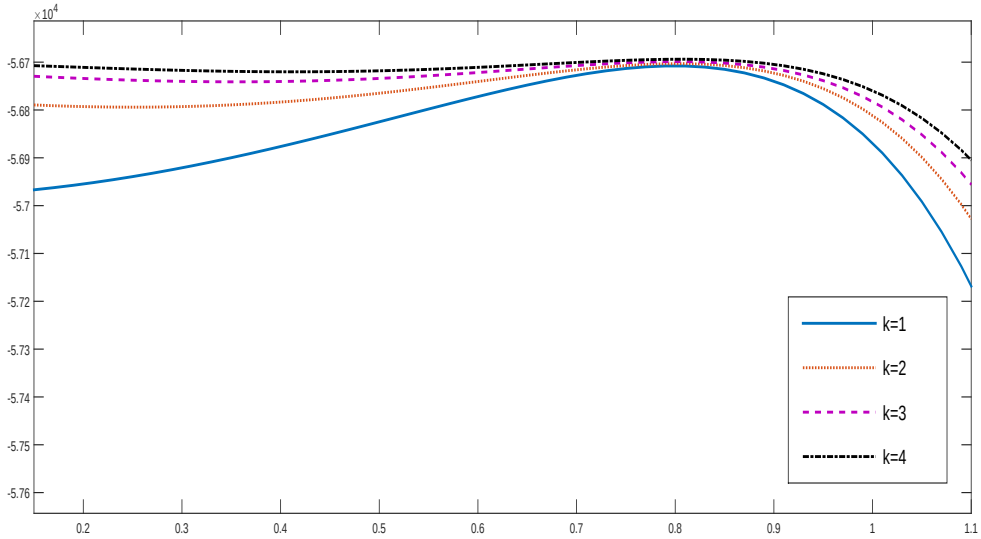
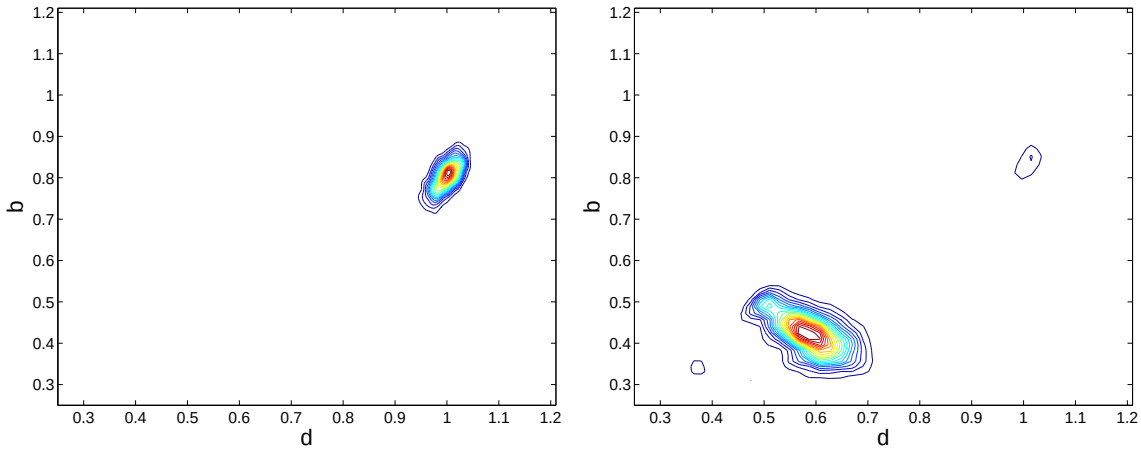


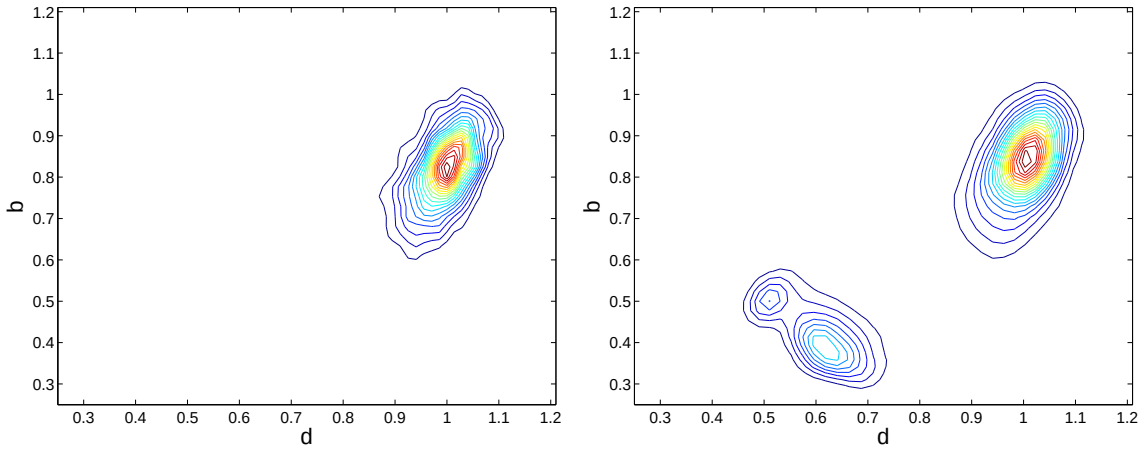
Figure B.3: Figure reports the values of the expected profile likelihood, $\bar{l}(\psi)$, for different values of $b \in [0.1, 1.1]$ (x-axis) when $d = d_0 = 1$. The sample size is $T = 20000$ and $k_0 = 0$, while $\mathcal{H}_k$ with $k = 1, 2, 3, 4$ is estimated. The parameters of the DGP are $d_0 = 1$ and $b_0 = 0.8$, $\beta_0 = [1, -1]'$, $\alpha_0 = [-0.5, 0.5]'$.



(a) Constrained

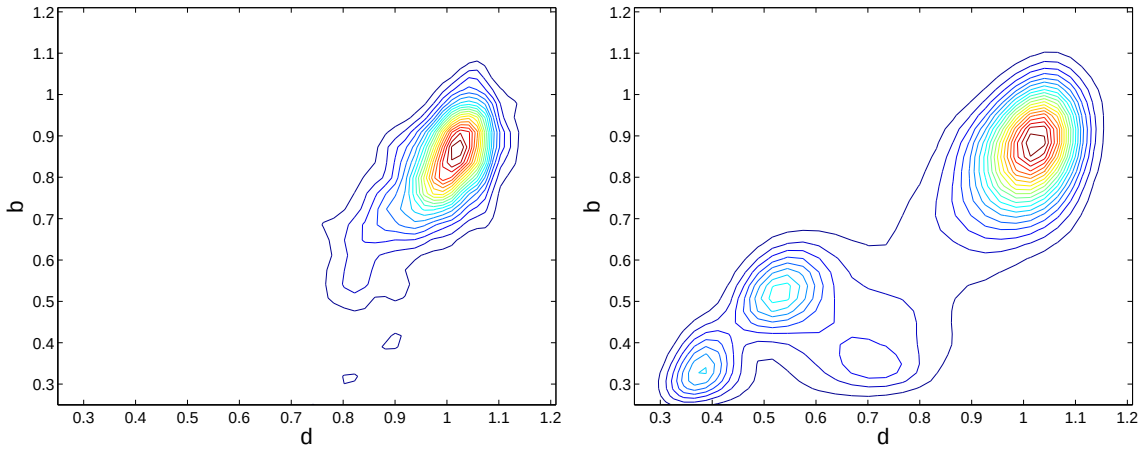
(b) Unconstrained

Figure B.4: Figure reports the contour plot of $M = 1000$ Monte Carlo estimates of the parameters d (x-axis) and b (y-axis) when a sample of $T = 10000$ observations is generated by a bivariate FCVAR$_{d,b}$ model with $k_0 = 0$, $d_0 = 1$, $b_0 = 0.8$ and the cointegration vectors given by $\beta_0 = [1, -1]'$ and $\alpha_0 = [-0.5, 0.5]'$. Model $\mathcal{H}_2$ is estimated on the data. Panel a) is relative to the estimates based on the constrained log-likelihood (13) where $\gamma = 2$ and $m = T^{0.6}$. Panel b) reports the contour plot for the unrestricted estimates.



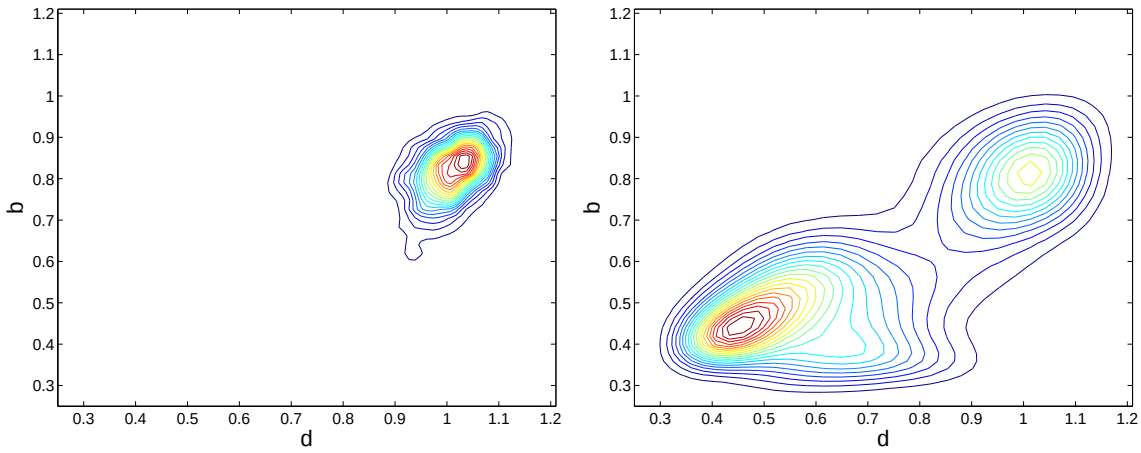
(a) Constrained (b) Unconstrained

Figure B.5: Figure reports the contour plot of $M = 1000$ Monte Carlo estimates of the parameters d (x-axis) and b (y-axis) when a sample of $T = 1000$ observations is generated by a bivariate FCVAR$_{d,b}$ model with $k_0 = 0$, $d_0 = 1$, $b_0 = 0.8$ and the cointegration vectors given by $\beta_0 = [1, -1]'$ and $\alpha_0 = [-0.5, 0.5]'$. Model $\mathcal{H}_1$ is estimated on the data. Panel a) is relative to the estimates based on the constrained log-likelihood (13) where $\gamma = 2$ and $m = T^{0.6}$. Panel b) reports the contour plot for the unrestricted estimates.



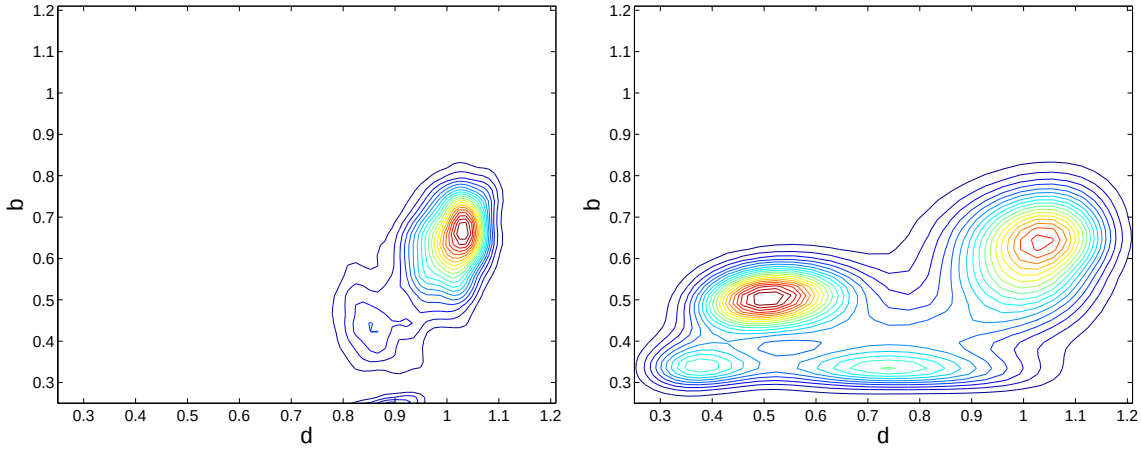
(a) Constrained (b) Unconstrained

Figure B.6: Figure reports the contour plot of $M = 1000$ Monte Carlo estimates of the parameters d (x-axis) and b (y-axis) when a sample of $T = 1000$ observations is generated by a bivariate FCVAR$_{d,b}$ model with $k_0 = 0$, $d_0 = 1$, $b_0 = 0.8$ and the cointegration vectors given by $\beta_0 = [1, -1]'$ and $\alpha_0 = [-0.5, 0.5]'$. Model $\mathcal{H}_2$ is estimated on the data. Panel a) is relative to the estimates based on the constrained log-likelihood (13) where $\gamma = 2$ and $m = T^{0.6}$. Panel b) reports the contour plot for the unrestricted estimates.



(a) Constrained (b) Unconstrained

Figure B.7: Figure reports the contour plot of $M = 1000$ Monte Carlo estimates of the parameters d (x-axis) and b (y-axis) when a sample of $T = 2500$ observations is generated by a bivariate FCVAR$_{d,b}$ model with $k_0 = 1$, $d_0 = 1$, $b_0 = 0.8$ and the cointegration vectors given by $\beta_0 = [1, -1]'$, $\alpha_0 = [-0.5, 0.5]'$ and $\Gamma_1 = \left[\begin{smallmatrix} 0.3 & -0.2 \\ -0.4 & 0.5 \end{smallmatrix}\right]$. Model $\mathcal{H}_3$ is estimated on the data. Panel a) is relative to the estimates based on the constrained log-likelihood (13) where $\gamma = 2$ and $m = T^{0.6}$. Panel b) reports the contour plot for the unrestricted estimates.



(a) Constrained (b) Unconstrained

Figure B.8: Figure reports the contour plot of $M = 1000$ Monte Carlo estimates of the parameters d (x-axis) and b (y-axis) when a sample of $T = 2500$ observations is generated by a bivariate FCVAR$_{d,b}$ model with $k_0 = 0$, $d_0 = 1$, $b_0 = 0.6$ and the cointegration vectors given by $\beta_0 = [1, -1]'$ and $\alpha_0 = [-0.5, 0.5]'$. Model $\mathcal{H}_2$ is estimated on the data. Panel a) is relative to the estimates based on the constrained log-likelihood (13) where $\gamma = 2$ and $m = T^{0.6}$. Panel b) reports the contour plot for the unrestricted estimates.

Research Papers 2013

2014-25: Matias D. Cattaneo and Michael Jansson: Bootstrapping Kernel-Based Semiparametric Estimators

2014-26: Markku Lanne, Jani Luoto and Henri Nyberg: Is the Quantity Theory of Money Useful in Forecasting U.S. Inflation?

2014-27: Massimiliano Caporin, Eduardo Rossi and Paolo Santucci de Magistris: Volatility jumps and their economic determinants

2014-28: Tom Engsted: Fama on bubbles

2014-29: Massimiliano Caporin, Eduardo Rossi and Paolo Santucci de Magistris: Chasing volatility - A persistent multiplicative error model with jumps

2014-30: Michael Creel and Dennis Kristensen: ABC of SV: Limited Information Likelihood Inference in Stochastic Volatility Jump-Diffusion Models

2014-31: Peter Christoffersen, Asger Lunde and Kasper V. Olesen: Factor Structure in Commodity Futures Return and Volatility

2014-32: Ulrich Hounyo: The wild tapered block bootstrap

2014-33: Massimiliano Caporin, Luca Corazzini and Michele Costola: Measuring the Behavioral Component of Financial Fluctuations: An Analysis Based on the S&P 500

2014-34: Morten Ørregaard Nielsen: Asymptotics for the conditional-sum-of-squares estimator in multivariate fractional time series models

2014-35: Ulrich Hounyo: Bootstrapping integrated covariance matrix estimators in noisy jump-diffusion models with non-synchronous trading

2014-36: Mehmet Caner and Anders Bredahl Kock: Asymptotically Honest Confidence Regions for High Dimensional

2014-37: Gustavo Fruet Dias and George Kapetanios: Forecasting Medium and Large Datasets with Vector Autoregressive Moving Average (VARMA) Models

2014-38: Søren Johansen: Times Series: Cointegration

2014-39: Søren Johansen and Bent Nielsen: Outlier detection algorithms for least squares time series regression

2014-40: Søren Johansen and Lukasz Gatarek: Optimal hedging with the cointegrated vector autoregressive model

2014-41: Laurent Callot and Johannes Tang Kristensen: Vector Autoregressions with Parsimoniously Time Varying Parameters and an Application to Monetary Policy

2014-42: Laurent A. F. Callot, Anders B. Kock and Marcelo C. Medeiros: Estimation and Forecasting of Large Realized Covariance Matrices and Portfolio Choice

2014-43: Paolo Santucci de Magistris and Federico Carlini: On the identification of fractionally cointegrated VAR models with the F(d) condition