Dynamic Factor Models, Cointegration, and Error Correction Mechanisms

Matteo Barigozzi[1] Marco Lippi[2] Matteo Luciani[3]

February 22, 2014

Abstract

In this paper we study Dynamic Factor Models when the factors $\mathbf{F}_t$ are $I(1)$ and singular, i.e. $\text{rank}(\mathbf{F}_t) < \dim(\mathbf{F}_t)$. By combining the classic Granger Representation Theorem with recent results by Anderson and Deistler on singular stochastic vectors, we prove that for generic values of the parameters $\mathbf{F}_t$ has an Error Correction representation with two unusual features: (i) the autoregressive matrix polynomial is *finite*, (ii) the number of error-terms is equal to the number of transitory shocks *plus* the difference between the dimension and the rank of $\mathbf{F}_t$. This result is the basis for the correct specification of an autoregressive model for $\mathbf{F}_t$. Estimation of impulse-response functions is also discussed. Results of an empirical analysis on a US quarterly database support the use of our model.

JEL subject classification: C0, C01, E0.

Key words and phrases: Dynamic Factor Models for $I(1)$ variables, Cointegration, Granger Representation Theorem.

[1]m.barigozzi@lse.ac.uk – London School of Economics and Political Science, UK.
[2]ml@lippi.ws – Einaudi Institute for Economics and Finance, Roma.
[3]matteo.luciani@ulb.ac.be – ECARES, SBS-EM, Université libre de Bruxelles and F.R.S.-F.N.R.S.

This paper has benefited from discussions with seminar participants at the 67th *ESEM* (University of Gothenburg), the JAE Conference *Forecasting Structure and Time Varying Patterns in Economics and Finance* (Erasmus University Rotterdam), the *1st Vienna Workshop on High Dimensional Time Series in Macroeconomics and Finance* (Institute for Advanced Studies), and the 2013 annual workshop on *New Developments in Econometrics and Time Series* (Université libre de Bruxelles). Matteo Luciani is *chargé de recherches* F.R.S.-F.N.R.S., and gratefully acknowledge their financial support. Of course, any errors are our responsibility.

1 Introduction

In the last fifteen years factor models have become increasingly popular in the economic literature and they are nowadays commonly used by policy institutions, such as Central Banks and Ministries. Due to the strong comovement among macroeconomic time series, these models offer a realistic (and parsimonious) representation of the data, and, moreover, they have proven successful both in forecasting (Stock and Watson, 2002a,b; Forni et al., 2005; Giannone et al., 2008; Luciani, 2014), and in structural analysis (Giannone et al., 2005; Stock and Watson, 2005; Forni et al., 2009; Forni and Gambetti, 2010; Barigozzi et al., 2013; Luciani, 2013).

Factor models are based on the idea that fluctuations in the economy are due to a few structural shocks ($\mathbf{u}_t$) affecting all the variables, and to several idiosyncratic shocks (generally of much less interest) resulting perhaps from measurement error or sectorial or regional dynamics, and influencing just one or a few variables. Therefore, each variable in the dataset (x_{it}) can be decomposed into the sum of a common (χ_{it}) and an idiosyncratic component (ξ_{it}): $x_{it} = \chi_{it} + \xi_{it}$ (Forni et al., 2000; Forni and Lippi, 2001). Typically the common component is assumed to be characterized by a linear combination of a small number of common factors ($\mathbf{F}_t$), $\chi_{it} = \boldsymbol{\lambda}_i \mathbf{F}_t$, with dynamics driven by the common shocks $\mathbf{F}_t = \mathbf{C}(L)\mathbf{u}_t$ (Forni et al., 2009).

If the factors $\mathbf{F}_t$ and the idiosyncratic terms are stationary, and hence the data are stationary as well, the factors and the loadings are estimated using principal components. The common shocks $\mathbf{u}_t$ and the impulse-response functions $\mathbf{C}(L)$ are then obtained by estimating a singular VAR on $\mathbf{F}_t$, where "singular" refers to the fact that the number of factors (dimension of $\mathbf{F}_t$) is allowed to be greater than the number of shocks (dimension of $\mathbf{u}_t$). Lastly, the impulse-response functions of the observable variables x_{it} result by applying the loadings $\boldsymbol{\lambda}_i$.

If the factors, and hence the data, are $I(1)$, the variables x_{it} are replaced by the first differences $(1-L)x_{it}$, and the same procedure can be used. With a few exceptions (Bai, 2004; Bai and Ng, 2004; Peña and Poncela, 2004), this is the common practice. However, in this setting all common shocks have a permanent effect on the levels of the non-stationary variables by construction. That is, in this setting all common shocks induce common trends, which, of course, is not the case if the $\mathbf{u}_t$ are the main sources of macroeconomic fluctuations. There is, indeed, full agreement in the macroeconomic literature that some shocks (such as technology shocks) have permanent effects, while some others (such as monetary policy shocks) have only transitory effects.

In this paper we study the case in which (i) the model is non–stationary, i.e. variables and common factors are $I(1)$, while the idiosyncratic components may or may not be $I(1)$, (ii) the vector of common factors $(1-L)\mathbf{F}_t$ is singular and has rational spectral density, i.e. it has an ARMA representation.

The main contribution of this paper is to derive the correct autoregressive representation for the vector of common factors under these assumptions. If $(1-L)\mathbf{F}_t$ is non singular, then it is well known that an autoregressive model requires also an error correction term (Engle and Granger, 1987). However, there exists large empirical evidence that the vector of common factors is singular (see, among others Giannone et al., 2005, Amengual and Watson, 2007, Bai and Ng, 2007, Forni and Gambetti, 2010, and Luciani, 2013 for the US, and Barigozzi et al., 2013 for the euro area), and this case is more complicated and studied here.

We prove that for generic values of the parameters there exists a *finite* Vector Error Correction representation for singular $I(1)$ vectors. This result is obtained by combining the Granger Representation Theorem (Engle and Granger, 1987) with the results by Anderson and Deistler (2008a,b) on stationary vectors with singular rational spectral density. Singularity plays a crucial role since this is what (generically) guarantees the existence of a *finite* autoregressive representation. Moreover, the number of error-terms is equal to the number of transitory shocks *plus* the difference between dimension and rank of $\mathbf{F}_t$, as though there were additional transitory shocks with zero loadings.

There exists also an alternative admissible representation for $\mathbf{F}_t$ consisting in assuming a VAR in levels (Engle and Granger, 1987; Sims et al., 1990). However, as argued in Phillips (1998), when the variables are cointegrated, the long-run features of the impulse-response functions are consistently estimated only if the unit roots are explicitly taken into account, that is within an Error Correction Model (see also Paruolo, 1997b).

Out of its intrinsic interest, our theorem provides a solid basis for consistent estimation of the impulse-response functions in non-stationary Dynamic Factor Models. Precisely, once the dimension of $\mathbf{F}_t$ and $\mathbf{u}_t$, as well as the vector $\mathbf{F}_t$ and the number of transitory shocks, are estimated, our result can be used to correctly specify an Error Correction Model for $\mathbf{F}_t$ and hence to recover the impulse-response functions.

Finally, we estimate a Dynamic Factor Model on a database of non-stationary US quarterly variables. Results illustrate the importance of a correct specification for the autoregressive representation of $(1-L)\mathbf{F}_t$ since using the Error Correction representation produces impulse-response functions that are consistent with standard macroeconomic theory.

The rest of the paper is organized as follows: in Section 2, we state formally the problem that we are addressing, while in Section 3 we formulate the basic assumptions and we define what a permanent and a transitory shocks are. Then, in Section 4 we present the Granger Representation Theorem for reduced–rank $I(1)$ vectors. In Section 5 we discuss estimation, while in Section 6 we present an empirical application. Finally, in Section 7 we summarize the results and we conclude.

2 Statement of the problem

Consider the Dynamic Factor Model

$$\mathbf{x}_t = \boldsymbol{\chi}_t + \boldsymbol{\xi}_t, \qquad \boldsymbol{\chi}_t = \boldsymbol{\Lambda}\mathbf{F}_t, \tag{1}$$

where: (1) the observables $\mathbf{x}_t$, the common components $\boldsymbol{\chi}_t$, and the idiosyncratic components $\boldsymbol{\xi}_t$ are n-dimensional vectors, (2) $\mathbf{F}_t$ is an r-dimensional vector of common factors, with r independent of n, (3) $\boldsymbol{\Lambda}$ is an $n \times r$ matrix. Moreover, $\mathbf{F}_t$ is driven by a q–dimensional zero–mean white noise vector process $\mathbf{u}_t$, the common shocks, with $q \leq r$.

With some exceptions, the theory of model (1) has been developed under the assumption that $\mathbf{x}_t$, $\boldsymbol{\chi}_t$, $\boldsymbol{\xi}_t$ and $\mathbf{F}_t$ are stationary, or, more precisely, that the observable variables have been reduced to stationarity by suitable transformations, first differences in particular (Bai and Ng, 2002; Stock and Watson, 2002a; Forni et al., 2009). Only a few papers have explicitly considered the consequences of non-stationarity (Bai, 2004; Bai and Ng, 2004).

Under stationarity it is usually assumed that $\mathbf{F}_t$ has a reduced–rank VAR representation:

$$\mathbf{A}(L)\mathbf{F}_t = \mathbf{R}\mathbf{u}_t, \tag{2}$$

where $\mathbf{A}(L)$ is a finite-degree $r \times r$ matrix polynomial and $\mathbf{R}$ is $r \times q$. Moreover, it is well known that in this settings $\boldsymbol{\Lambda}$ and $\mathbf{F}_t$ can be estimated by principal components, while an estimate of $\mathbf{A}(L)$, $\mathbf{R}$ and $\mathbf{u}_t$ can be obtained by standard techniques. Inversion of $\mathbf{A}(L)$ provides an estimate of the reduced-form impulse-response functions of the observables to the common shocks:

$$\mathbf{x}_t = \mathbf{\Lambda A}(L)^{-1}\mathbf{R u}_t + \boldsymbol{\xi}_t.$$

Structural shocks and structural impulse-response functions can then be obtained, respectively, as $\mathbf{w}_t = \mathbf{Q u}_t$ and $\mathbf{\Lambda A}(L)^{-1}\mathbf{R Q}^{-1}$, where the $q \times q$ matrix $\mathbf{Q}$ is determined in the same way as in Structural VARs (Forni et al., 2009).

The VAR representation (2) has a standard motivation as an approximation to an infinite autoregressive representation with exponentially declining coefficients. However, as stated above, $\mathbf{F}_t$ has reduced rank. Under reduced rank and rational spectral density for $\mathbf{F}_t$, Anderson and Deistler (2008a,b) prove that *generically* $\mathbf{F}_t$ has a finite-degree autoregressive representation, so that no approximation argument is needed to motivate (2). A formal statement of this result requires the following definitions.

Definition 1 ***(Rational reduced-rank family)*** *Assume that $r > q > 0$ and let $\mathcal{G}$ be a set of ordered couples $(\mathbf{S}(L), \mathbf{C}(L))$, where:*

(i) *$\mathbf{C}(L)$ is an $r \times q$ polynomial matrix of degree $s_1 \geq 0$.*

(ii) *$\mathbf{S}(L)$ is an $r \times r$ polynomial matrix of degree $s_2 \geq 0$. $\mathbf{S}(0) = \mathbf{I}_r$.*

(iii) *Let $\boldsymbol{\theta}$ be the vector containing the $\lambda = rq(s_1 + 1) + r^2 s_2$ coefficients of the entries of $\mathbf{C}(L)$ and $\mathbf{S}(L)$. We assume that $\boldsymbol{\theta} \in \Pi$, where Π is an open subset of $\mathbb{R}^\lambda$ and that for $\boldsymbol{\theta} \in \Pi$, if $\det(S(z)) = 0$, then $|z| > 1$.*

We say that the family of weakly stationary stochastic processes

$$\mathbf{F}_t = \mathbf{S}(L)^{-1}\mathbf{C}(L)\mathbf{u}_t, \tag{3}$$

where $\mathbf{u}_t$ is a q-dimensional white noise with non-singular variance-covariance matrix and $(\mathbf{S}(L), \mathbf{C}(L))$ belongs to $\mathcal{G}$, is a rational reduced-rank family.

Notice that (3) is the unique stationary solution of the ARMA equation

$$\mathbf{S}(L)\mathbf{Z}_t = \mathbf{C}(L)\mathbf{u}_t. \tag{4}$$

Definition 2 ***(Genericity)*** *Suppose that a statement $Q(p)$ depends on $p \in \mathcal{A}$, where $\mathcal{A}$ is an open subset of $\mathbb{R}^\mu$. Then $Q(p)$ holds generically in $\mathcal{A}$ if the subset $\mathcal{N}$ of $\mathcal{A}$ where it does not hold is nowhere dense in $\mathcal{A}$, i.e. the closure of $\mathcal{N}$ has no internal points.*

Proposition 1 ***(Anderson and Deistler)*** *(I) Suppose that $\mathbf{V}(L)$ is an $r \times q$ polynomial matrix of degree s, with $r > q$. If $\mathbf{V}(L)$ is zeroless, i.e. has rank q for all complex numbers z, then $\mathbf{V}(L)$ has a finite-degree polynomial stable left inverse, i.e. there exists a finite-degree polynomial $r \times r$ matrix $\mathbf{W}(L)$, such that (i) $\det(\mathbf{W}(z)) = 0$ implies $|z| > 1$, (ii) $\mathbf{W}(L)\mathbf{V}(L) = \mathbf{V}(0)$. (II) Let $\mathbf{F}_t = \mathbf{S}(L)^{-1}\mathbf{C}(L)\mathbf{u}_t$ be a rational reduced-rank family with parameter set Π. For generic values of the parameters in Π, $\mathbf{C}(L)$ is zeroless.*

For statements (I) and (II) of Proposition 1, see Deistler et al. (2010), Theorems 3 and 2 respectively. The genericity statement (II) is obtained by Anderson and Deistler for the parameters of the state-space representation. The same result is proved in the present paper with the coefficients of the matrix polynomials of the ARMA representation (4) as parameters, see Proposition 2.[1] Both statement (I) and our version of statement (II) are used in the proof of Proposition 3.

Now suppose that $\mathbf{F}_t$ and $\mathbf{x}_t$ are $I(1)$ and that $\boldsymbol{\xi}_t$ is either $I(1)$ or $I(0)$. Taking first differences in (1) we obtain

$$(1-L)\mathbf{x}_t = \boldsymbol{\Lambda}(1-L)\mathbf{F}_t + (1-L)\boldsymbol{\xi}_t.$$

The matrix $\boldsymbol{\Lambda}$, and $(1-L)\mathbf{F}_t$ can be estimated using standard factor techniques. However, if we are interested in impulse-response functions and therefore in an autoregressive representation for $\mathbf{F}_t$, we must face a problem that is specific to the non-stationary case, i.e. the possibility that $\mathbf{F}_t$ is cointegrated.

If $r = q$, cointegration of $\mathbf{F}_t$ has the consequence that an autoregressive model for $\mathbf{F}_t$ must take the Error Correction form (Engle and Granger, 1987). The case $r > q$ is more complicated and still unexplored. Suppose that

$$(1-L)\mathbf{F}_t = \mathbf{C}(L)\mathbf{u}_t$$

is a rational reduced-rank family, where for simplicity we set $\mathbf{S}(L) = \mathbf{I}_r$.

(i) Let us firstly observe that a reduced-rank $I(1)$ vector has at least $r-q$ cointegration vectors. In other words, denoting by c the cointegration rank of $\mathbf{F}_t$, $c \geq r-q$.

(ii) We might argue that generically $\mathbf{C}(z)$ has rank q for all z and therefore the cointegration rank of $\mathbf{F}_t$ is generically $r-q$. However, the rank of $\mathbf{C}(z)$ at $z=1$ has a special meaning, as it depends on the number of equilibrium relationships between the processes F_{ft}. Such number usually has a theoretical or behavioral motivation, so that it cannot be modified by any genericity argument. Thus we keep the result that $\mathbf{C}(z)$ is generically full rank but only for $z \neq 1$.

(iii) If $c = r-q$, then $\text{rk}(\mathbf{C}(1)) = q$ and Anderson and Deistler's result can be applied. In spite of cointegration, generically $(1-L)\mathbf{F}_t$ has a finite-degree autoregressive representation

$$\mathbf{A}(L)(1-L)\mathbf{F}_t = \mathbf{C}(0)\mathbf{u}_t. \tag{5}$$

(iv) If $c > r-q$, then no autoregressive representation exists for $(1-L)\mathbf{F}_t$, finite or infinite. However, we prove that, irrespective of whether c is equal or greater than $r-q$, generically $\mathbf{F}_t$ has a Vector Error Correction representation

$$\mathbf{A}(L)\mathbf{F}_t = \mathbf{A}^*(L)(1-L)\mathbf{F}_t + \mathbf{A}(1)\mathbf{F}_{t-1} = \mathbf{h} + \mathbf{C}(0)\mathbf{u}_t, \tag{6}$$

where the rank of $\mathbf{A}(1)$ is c, and $\mathbf{A}(L)$ and $\mathbf{A}^*(L)$ are *finite-degree* matrix polynomials.

(v) The finite-degree autoregressive representation of equation (6), unfortunately, is not necessarily unique (as it is the case with full rank vectors). For example, when $c = r-q$, (5) and (6) are two different autoregressive representations for $\mathbf{F}_t$, the first with no error terms, the second with $r-q$ error terms. However, as we discuss in Section 5 and in Appendix C, empirically this non-uniqueness does not represent a problem.

The existence of representation (6) for reduced-rank $I(1)$ vectors, where $\mathbf{A}^*(L)$ is of finite degree, is our main result. Its proof combines the Granger Representation Theorem by Engle and Granger, with Anderson and Deistler's results on singular stationary vectors with rational spectral density.[2]

[1] See also Forni et al. (2009), Forni and Lippi (2010), Forni et al. (2014).

[2] Regarding the Granger Representation Theorem we closely follow Johansen (1995), Theorem 4.5, p. 55-57.

3 Cointegration for reduced-rank vectors

3.1 Basic definitions and results

Assume that $\mathbf{S}(L)$ and $\mathbf{C}(L)$ are as in Definition 1 and that $\mathbf{C}(1) \neq \mathbf{0}$. Setting $\mathbf{U}(L) = \mathbf{S}(L)^{-1}\mathbf{C}(L)$, consider the equation

$$(1 - L)\mathbf{F}_t = \mathbf{S}(L)^{-1}\mathbf{C}(L)\mathbf{u}_t = \mathbf{U}(L)\mathbf{u}_t. \tag{7}$$

We can read (7) either as the definition of $(1 - L)\mathbf{F}_t$ or as a difference equation for $\mathbf{F}_t$. Assuming the second point of view, suppose that $u_{jt} \in L_2(\Omega, \mathcal{F}, P)$, for $j = 1, \ldots, q$, where $(\Omega, \mathcal{F}, P)$ is a probability space. Limiting the search to processes belonging to $(\Omega, \mathcal{F}, P)$, it is easily seen that all the solutions of (7) are the processes $\mathbf{F}_t^{\mathbf{W}} = \tilde{\mathbf{F}}_t + \mathbf{W}$, where $\mathbf{W}$ is an r-dimensional stochastic vector with $W_f \in L_2(\Omega, \mathcal{F}, P)$, $f = 1, \ldots, r$, and

$$\tilde{\mathbf{F}}_t = \begin{cases} \mathbf{U}(L)(\mathbf{u}_1 + \mathbf{u}_2 + \cdots + \mathbf{u}_t), \text{ for } t > 0 \\ 0, \text{ for } t = 0 \\ -\mathbf{U}(L)(\mathbf{u}_0 + \mathbf{u}_{-1} + \cdots + \mathbf{u}_{t+1}), \text{ for } t < 0. \end{cases}$$

Because by assumption $\mathbf{C}(1) \neq 0$ then a solution of (7) is $I(1)$, see Johansen (1995, p. 35). As no confusion can arise we will drop the reference to $\mathbf{W}$ and denote by $\mathbf{F}_t$ any solution of (7).

Now, because $\mathbf{S}(1)^{-1}$ is a non-singular $r \times r$ matrix, the cointegration rank of $\mathbf{F}_t$ only depends on $\mathbf{C}(1)$. Precisely, if c is the cointegration rank of $\mathbf{F}_t$, then $r > c \geq r - q$ and the matrix $\mathbf{C}(1)$ has rank $r - c$. As a consequence,

$$\mathbf{C}(1) = \boldsymbol{\xi}\boldsymbol{\eta}', \tag{8}$$

where $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ are $r \times (r - c)$ and $q \times (r - c)$, respectively, and are both of full rank $r - c$, see Lancaster and Tismenetsky (1985, p. 97, Proposition 3), notice that $q \geq r - c$. The matrix $\mathbf{C}(L)$ has the (finite) Taylor expansion

$$\mathbf{C}(L) = \mathbf{C}(1) - (1 - L)\mathbf{C}'(1) + \frac{1}{2}(1 - L)^2\mathbf{C}''(1) - \cdots$$

Gathering all terms after the second and using (8),

$$\mathbf{C}(L) = \boldsymbol{\xi}\boldsymbol{\eta}' - (1 - L)\mathbf{C}'(1) + (1 - L)^2\mathbf{C}_1(L), \tag{9}$$

where $\mathbf{C}_1(L)$ is a polynomial matrix.

Representation (9) can be used for a very convenient parameterization of $\mathbf{C}(L)$.

Definition 3 ***(Rational reduced-rank I(1) family with cointegration rank c)*** *Assume that $r > q > 0$, $r > c \geq r - q$ and let $\mathcal{G}$ be a set of couples $(\mathbf{S}(L), \mathbf{C}(L))$, where:*

(i) The matrix $\mathbf{C}(L)$ has the representation

$$\mathbf{C}(L) = \boldsymbol{\xi}\boldsymbol{\eta}' + (1 - L)\mathbf{D} + (1 - L)^2\mathbf{E}(L), \tag{10}$$

where $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ are $r \times (r - c)$ and $q \times (r - c)$ respectively, $\mathbf{D}$ is an $r \times q$ matrix and $\mathbf{E}(L)$ is an $r \times q$ matrix polynomial of degree $s_1 \geq 0$.

(ii) $\mathbf{S}(L)$ *is an* $r \times r$ *polynomial matrix of degree* s_2*.* $\mathbf{S}(0) = \mathbf{I}_r$*.*

(iii) *Let* $\boldsymbol{\theta}$ *be the vector containing the* $\lambda = (r-c)(r+q) + rq(1+s_1) + r^2 s_2$ *coefficients of the matrices in (10) and in* $\mathbf{S}(L)$*. We assume that* $\boldsymbol{\theta} \in \Pi$*, where* Π *is an open subset of* $\mathbb{R}^\lambda$ *and that for* $\boldsymbol{\theta} \in \Pi$*, (a)* $\boldsymbol{\xi}$ *and* $\boldsymbol{\eta}$ *are of full rank* $r-c$*, (b) if* $\det(\mathbf{S}(z)) = 0$ *then* $|z| > 1$*, (c)* $\mathbf{C}(0)$ *has rank* q*.*

We say that the family of I(1) processes $\mathbf{F}_t$*, such that* $(1-L)\mathbf{F}_t = \mathbf{S}(L)^{-1}\mathbf{C}(L)\mathbf{u}_t$*, where* $\mathbf{u}_t$ *is a q-dimensional white noise with non-singular variance-covariance matrix and* $(\mathbf{S}(L), \mathbf{C}(L))$ *belongs to* $\mathcal{G}$*, is a rational reduced-rank family with cointegration rank c.*

Denoting by $\boldsymbol{\xi}_\perp$ an $r \times c$ matrix whose columns are linearly independent and orthogonal to all columns of $\boldsymbol{\xi}$, the columns of $\boldsymbol{\xi}_\perp$ are a full set of independent cointegrating vectors of $\mathbf{S}(L)\mathbf{F}_t$, while the columns of $\boldsymbol{\vartheta} = \mathbf{S}(1)'\boldsymbol{\xi}_\perp$ are a full set of independent cointegrating vectors of $\mathbf{F}_t$.

3.2 Permanent and transitory shocks

Let $\mathbf{F}_t$ be a rational reduced-rank $I(1)$ family with cointegration rank c, $\mathbf{S}(L)$ and $\mathbf{C}(L)$ given as in Definition 3, and $(1-L)\mathbf{F}_t = \mathbf{S}(L)^{-1}\mathbf{C}(L)\mathbf{u}_t$.

Denote by d the number of cointegration vectors exceeding the minimum $r-q$, so that $c = r-q+d$, or $q-d = r-c$. Of course $q > d \geq 0$. Let $\boldsymbol{\eta}_\perp$ be a $q \times d$ matrix whose columns are independent and orthogonal to the columns of $\boldsymbol{\eta}$. The $q \times q$ matrix $(\boldsymbol{\eta}_\perp\ \boldsymbol{\eta})$ is invertible. We have

$$\mathbf{C}(L)\,(\boldsymbol{\eta}_\perp\ \boldsymbol{\eta}) = \left(\boldsymbol{\xi}\boldsymbol{\eta}' + (1-L)\mathbf{D} + (1-L)^2\mathbf{E}(L)\right)(\boldsymbol{\eta}_\perp\ \boldsymbol{\eta}) = \left[\mathbf{0}_{r\times d}\ \ \boldsymbol{\xi}\boldsymbol{\eta}'\boldsymbol{\eta}\right] + (1-L)\mathbf{G}(L),$$

where $\mathbf{G}(L) = (\mathbf{D} + (1-L)\mathbf{E}(L))\,(\boldsymbol{\eta}_\perp\ \boldsymbol{\eta})$. Now let $\mathbf{v}_t = [\boldsymbol{\eta}_\perp\ \boldsymbol{\eta}]^{-1}\,\mathbf{u}_t$. Let us partition $\mathbf{v}_t$ and $\mathbf{G}(L)$ as

$$\mathbf{v}_t = \begin{pmatrix}\mathbf{v}_{1t}\\ \mathbf{v}_{2t}\end{pmatrix}, \quad \mathbf{G}(L) = \begin{pmatrix}\mathbf{G}_1(L) & \mathbf{G}_2(L)\end{pmatrix},$$

where $\mathbf{v}_{1t}$ and $\mathbf{v}_{2t}$ are d-dimensional and $(q-d)$-dimensional white noise respectively, $\mathbf{G}_1(L)$ and $\mathbf{G}_2(L)$ are $r \times d$ and $r \times (q-d)$ respectively. We have

$$\mathbf{C}(L)\mathbf{u}_t = (1-L)\mathbf{G}_1(L)\mathbf{v}_{1t} + \left(\tilde{\boldsymbol{\xi}} + (1-L)\mathbf{G}_2(L)\right)\mathbf{v}_{2t}, \tag{11}$$

where $\tilde{\boldsymbol{\xi}} = \boldsymbol{\xi}\boldsymbol{\eta}'\boldsymbol{\eta}$ is a full-rank $r \times (q-d)$, i.e. $r \times (r-c)$, matrix.

All the solutions of the difference equation $(1-L)\mathbf{F}_t = \mathbf{S}(L)^{-1}\mathbf{C}(L)\mathbf{u}_t$ can be written as

$$\mathbf{F}_t = \mathbf{S}(L)^{-1}\left(\mathbf{G}_1(L)\mathbf{v}_{1t} + \mathbf{G}_2(L)\mathbf{v}_{2t} + \mathbf{T}_t + \mathbf{Z}\right),$$

where $\mathbf{Z} \in L_2(\Omega, \mathcal{F}, P)$, and

$$\mathbf{T}_t = \begin{cases}\tilde{\boldsymbol{\xi}}(\mathbf{v}_{21} + \mathbf{v}_{22} + \cdots + \mathbf{v}_{2t}), \text{ for } t > 0\\ 0, \text{ for } t = 0\\ -\tilde{\boldsymbol{\xi}}(\mathbf{v}_{20} + \mathbf{v}_{2,-1} + \cdots + \mathbf{v}_{2,t+1}), \text{ for } t < 0.\end{cases}$$

As $\tilde{\boldsymbol{\xi}}$ is full rank, we see that $\mathbf{F}_t$ is driven by the $q-d = r-c$ permanent shocks $\mathbf{v}_{2t}$, and the d temporary shocks $\mathbf{v}_{1t}$. Note that the number of permanent shocks is obtained as r minus the cointegration rank, as usual. However, the number of transitory shocks is obtained as the complement of the number of permanent shocks to q, not to r, as though $r-q$ transitory shocks had a zero coefficient.

4 ECM representations for reduced rank I(1) vectors

Let $\mathbf{F}_t$ be a rational reduced-rank $I(1)$ family with cointegration rank c with parameters in the open set $\Pi \in \mathbb{R}^\lambda$, $\mathbf{S}(L)$ and $\mathbf{C}(L)$ given as in Definition 3, $(1-L)\mathbf{F}_t = \mathbf{S}(L)^{-1}\mathbf{C}(L)\mathbf{u}_t = \mathbf{U}(L)\mathbf{u}_t$..

The matrix $\boldsymbol{\zeta} = \begin{pmatrix} \boldsymbol{\xi}'_\perp \\ \boldsymbol{\xi}' \end{pmatrix}$ is $r \times r$ and invertible. We have

$$\begin{aligned}(1-L)\boldsymbol{\zeta}\mathbf{S}(L)\mathbf{F}_t &= \left\{ \begin{pmatrix} \mathbf{0}_{c\times q} \\ \boldsymbol{\xi}'\boldsymbol{\xi}\boldsymbol{\eta}' \end{pmatrix} + (1-L)\begin{pmatrix} \boldsymbol{\xi}'_\perp\mathbf{D} \\ \boldsymbol{\xi}'\mathbf{D} \end{pmatrix} + (1-L)^2 \begin{pmatrix} \boldsymbol{\xi}'_\perp\mathbf{E}(L) \\ \boldsymbol{\xi}'\mathbf{E}(L) \end{pmatrix} \right\} \mathbf{u}_t \\ &= \begin{pmatrix} (1-L)\mathbf{I}_c & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{r-c} \end{pmatrix} \left\{ \begin{pmatrix} \boldsymbol{\xi}'_\perp\mathbf{D} \\ \boldsymbol{\xi}'\boldsymbol{\xi}\boldsymbol{\eta}' \end{pmatrix} + (1-L)\begin{pmatrix} \boldsymbol{\xi}'_\perp\mathbf{E}(L) \\ \boldsymbol{\xi}'\mathbf{D} \end{pmatrix} + (1-L)^2 \begin{pmatrix} \mathbf{0}_{c\times q} \\ \boldsymbol{\xi}'\mathbf{E}(L) \end{pmatrix} \right\} \mathbf{u}_t.\end{aligned} \tag{12}$$

Taking the first c rows,

$$(1-L)\boldsymbol{\xi}'_\perp\mathbf{S}(L)\mathbf{F}_t = (1-L)\left(\boldsymbol{\xi}'_\perp\mathbf{D} + (1-L)\boldsymbol{\xi}'_\perp\mathbf{E}(L)\right)\mathbf{u}_t.$$

We assume that the solution for $\boldsymbol{\xi}'_\perp\mathbf{S}(L)\mathbf{F}_t$ is weakly stationary with rational spectral density:

$$\boldsymbol{\xi}'_\perp\mathbf{S}(L)\mathbf{F}_t = \mathbf{k} + \left(\boldsymbol{\xi}'_\perp\mathbf{D} + (1-L)\boldsymbol{\xi}'_\perp\mathbf{E}(L)\right)\mathbf{u}_t,$$

where $\mathbf{k}$ is a c-dimensional constant, see Appendix B. Notice that since

$$\boldsymbol{\xi}'_\perp\mathbf{S}(L)\mathbf{F}_t = \boldsymbol{\xi}'_\perp\mathbf{S}(1)\mathbf{F}_t + \boldsymbol{\xi}'_\perp\mathbf{S}^*(L)(1-L)\mathbf{F}_t,$$

and $(1-L)\mathbf{F}_t$ is weakly stationary with rational spectral density, then $\boldsymbol{\xi}'_\perp\mathbf{S}(L)\mathbf{F}_t$ is weakly stationary with rational spectral density if and only if $\boldsymbol{\xi}'_\perp\mathbf{S}(1)\mathbf{F}_t = \boldsymbol{\vartheta}'\mathbf{F}_t$ is weakly stationary with rational spectral density. Therefore,

$$\begin{aligned}&\begin{pmatrix} \mathbf{I}_c & \mathbf{0} \\ \mathbf{0} & (1-L)\mathbf{I}_{r-c} \end{pmatrix}\boldsymbol{\zeta}\mathbf{S}(L)\mathbf{F}_t = \\ &\qquad \begin{pmatrix} \mathbf{k} \\ \mathbf{0}_{(r-c)\times 1} \end{pmatrix} + \left\{ \begin{pmatrix} \boldsymbol{\xi}'_\perp\mathbf{D} \\ \boldsymbol{\xi}'\boldsymbol{\xi}\boldsymbol{\eta}' \end{pmatrix} + (1-L)\begin{pmatrix} \boldsymbol{\xi}'_\perp\mathbf{E}(L) \\ \boldsymbol{\xi}'\mathbf{D} \end{pmatrix} + (1-L)^2 \begin{pmatrix} \mathbf{0}_{c\times q} \\ \boldsymbol{\xi}'\mathbf{E}(L) \end{pmatrix} \right\} \mathbf{u}_t.\end{aligned}$$

Denote by $\mathbf{M}(L)$ the matrix between curly brackets. The following statement is proved in Appendix A.

Proposition 2 *Assume that the family of I(1) processes* $\mathbf{F}_t$*, such that* $(1-L)\mathbf{F}_t = \mathbf{S}(L)^{-1}\mathbf{C}(L)\mathbf{u}_t$*, is a rational reduced-rank I(1) family with cointegration rank c and parameter set* Π*. Then, for generic values of the parameters in* Π*, the* $r \times q$ *matrix* $\mathbf{M}(z)$ *is zeroless. In particular, generically, the rank of*

$$\mathbf{M}(1) = \begin{pmatrix} \boldsymbol{\xi}'_\perp\mathbf{D} \\ \boldsymbol{\xi}'\boldsymbol{\xi}\boldsymbol{\eta}' \end{pmatrix}$$

is q.[3]

[3] For $r = q$, the statement that $\text{rank}(\mathbf{M}(1)) = q$ is equivalent to the condition that $\boldsymbol{\xi}_\perp\mathbf{C}^*\boldsymbol{\eta}_\perp$ has full rank in Johansen (1995), Theorem 5.4, p. 55.

A consequence of Proposition 2 and Proposition 1(I) is that generically there exists a finite-degree $r \times r$ polynomial matrix

$$\mathbf{N}(L) = \mathbf{I}_r + \mathbf{N}_1 L + \cdots + \mathbf{N}_p L^p,$$

for some p, such that: (i) $\mathbf{N}(L)\mathbf{M}(L) = \mathbf{M}(0)$, i.e. $\mathbf{N}(L)$ is a left inverse of $\mathbf{M}(L)$; (ii) all the roots of $\det(\mathbf{N}(L))$ lie outside the unit circle, so that $\mathbf{N}(1)$ has full rank.

In conclusion, for generic values of the parameters in Π,

$$\mathbf{A}(L)\mathbf{F}_t = \mathbf{h} + \mathbf{C}(0)\mathbf{u}_t,$$

where

$$\begin{aligned}\mathbf{A}(L) = \mathbf{I}_r + \mathbf{A}_1 L + \cdots + \mathbf{A}_P L^P &= \boldsymbol{\zeta}^{-1}\mathbf{N}(L)\begin{pmatrix}\mathbf{I}_c & \mathbf{0}\\ \mathbf{0} & (1-L)\mathbf{I}_{r-c}\end{pmatrix}\boldsymbol{\zeta}\mathbf{S}(L)\\ &= \boldsymbol{\zeta}^{-1}\mathbf{N}(L)\begin{pmatrix}\boldsymbol{\xi}'_\perp\\ (1-L)\boldsymbol{\xi}'\end{pmatrix}\mathbf{S}(L),\end{aligned}$$

with $P = p + 1 + s_2$, and

$$\mathbf{h} = \mathbf{A}(1)\begin{pmatrix}\mathbf{k}\\ \mathbf{0}_{(r-c)\times 1}\end{pmatrix}.$$

Defining

$$\boldsymbol{\alpha} = \boldsymbol{\zeta}^{-1}\mathbf{N}(1)\begin{pmatrix}\mathbf{I}_c\\ \mathbf{0}_{(r-c)\times c}\end{pmatrix},$$

and setting $\boldsymbol{\beta} = \boldsymbol{\vartheta} = \mathbf{S}(1)'\boldsymbol{\xi}_\perp$, both $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ have rank c (regarding $\boldsymbol{\alpha}$, remember that $\mathbf{N}(1)$ has full rank) and $\mathbf{A}(1) = \boldsymbol{\alpha}\boldsymbol{\beta}'$. Lastly, define $\mathbf{A}^*(L) = (1-L)^{-1}(\mathbf{A}(L) - \mathbf{A}(1)L)$. We have proved the following statement.

Proposition 3 ***(Granger Representation Theorem for reduced-rank I(1) vectors)*** *Assume that (i) $\mathbf{F}_t$ is a rational reduced-rank $I(1)$ family with cointegration rank c, parameterized in the open set $\Pi \subset \mathbb{R}^\lambda$, (ii) $\boldsymbol{\beta}'\mathbf{F}_t = \boldsymbol{\xi}'_\perp\mathbf{S}(1)\mathbf{F}_t$ is weakly stationary with rational spectral density and mean $\mathbf{k}$. Then for generic values of the parameters, $\mathbf{F}_t$ has the Error Correction representation*

$$\mathbf{A}(L)\mathbf{F}_t = \mathbf{A}^*(L)(1-L)\mathbf{F}_t + \boldsymbol{\alpha}\boldsymbol{\beta}'\mathbf{F}_{t-1} = \mathbf{h} + \mathbf{C}(0)\mathbf{u}_t, \tag{13}$$

where (1) the $r \times r$ finite-degree polynomial matrices $\mathbf{A}(L)$ and $\mathbf{A}^*(L)$, (2) the full-rank $r \times c$ matrices $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, (3) the r-dimensional constant vector $\mathbf{h}$, have been defined above.

Lastly, as shown in Appendix C, representation (13) is not unique since: (i) different Error Correction Representations can be obtained in which the number of error terms varies between d and $r - q + d$, the latter being the number chosen in Proposition 3; (ii) the left inverse of the matrix $\mathbf{M}(L)$ may be not unique.

However, non-uniqueness only affects the autoregressive representations, not the impulse-response functions. For, suppose that $\mathbf{F}_t$ fulfills both $(1-L)\mathbf{F}_t = \mathbf{S}(L)^{-1}\mathbf{C}(L)\mathbf{u}_t = \mathbf{U}(L)\mathbf{u}_t$ and

$$\mathbf{B}(L)\mathbf{F}_t = \mathbf{h} + \mathbf{C}(0)\mathbf{u}_t \tag{14}$$

where $\mathbf{B}(0) = \mathbf{I}_r$. Multiplying both sides of (14) by $(1-L)$, we have

$$\mathbf{B}(L)\mathbf{U}(L)\mathbf{u}_t = (1-L)\mathbf{C}(0)\mathbf{u}_t.$$

Because $\mathbf{u}_t$ is a non-singular q-dimensional white noise, this implies

$$\mathbf{B}(L)\mathbf{U}(L){=}(1-L)\mathbf{C}(0).$$

Now let $\mathbf{K}(L)=\mathbf{I}_r+\mathbf{K}_1L+\cdots$ be such that $\mathbf{K}(L)\mathbf{B}(L)=\mathbf{I}_r$, which is obtained by setting

$$\mathbf{K}_1=-\mathbf{B}_1,\ \mathbf{K}_2=-\mathbf{K}_1\mathbf{B}_1-\mathbf{B}_2,\ \ldots$$

We have

$$\mathbf{U}(L)=(1-L)\mathbf{K}(L)\mathbf{C}(0),$$

that is

$$\mathbf{K}_1\mathbf{C}(0)=\mathbf{U}_1+\mathbf{C}(0),\ \mathbf{K}_2\mathbf{C}(0)=\mathbf{U}_2+\mathbf{K}_1\mathbf{C}(0),\ \ldots$$

Therefore, as $\mathbf{U}(0)=\mathbf{C}(0)$, the sequence

$$\mathbf{C}(0),\ \mathbf{K}_1\mathbf{C}(0),\ \mathbf{K}_2\mathbf{C}(0),\ \ldots$$

only depends on $\mathbf{U}(L)$. Note also that, because $\mathbf{u}_t$ is non-simgular, $\mathbf{U}(L)$ is uniquely determined by $(1-L)\mathbf{F}_t$ and $\mathbf{u}_t$.

On the other hand, let

$$\mathbf{u}^*{}_t=\begin{cases}\mathbf{u}^* \text{ for } t=0\\ \mathbf{0}_q \text{ for } t\neq 0\end{cases}$$

and define $\mathbf{F}^*{}_t$ as the solution of (14) such that $\mathbf{u}_t=\mathbf{u}^*{}_t$ and $\mathbf{F}^*{}_t=\mathbf{0}_r$ for $t<0$. It is easily seen that, for $t\geq 0$,

$$\mathbf{F}^*{}_t=\mathbf{K}(L)\mathbf{C}(0)\mathbf{u}^*{}_t=\mathbf{K}_t\mathbf{C}(0)\mathbf{u}^*. \tag{15}$$

We refer to the sequence $\mathbf{\Psi}_t=\mathbf{K}_t\mathbf{C}(0)$, $t=0,1,\ldots$, as the *impulse-response* function of $\mathbf{F}_t$ with respect to $\mathbf{u}_t$. We have seen the impulse-response function of $\mathbf{F}_t$ with respect to $\mathbf{u}_t$ is independent of the particular autoregressive representation of $\mathbf{F}_t$. Replacing $\mathbf{u}_t$ with any other white noise vector $\mathbf{w}_t=\mathbf{Q}\mathbf{u}_t$, as we do when the shocks are identified according to economic considerations, produces different impulse-response functions.

Expressions $\mathbf{K}(L)=\mathbf{A}(L)^{-1}$ and $\mathbf{\Psi}(L)=\mathbf{A}(L)^{-1}\mathbf{C}(0)$ are convenient and do make sense, provided we do not forget that $\mathbf{A}(L)^{-1}$ is not a square-summable filter, and $\mathbf{A}(L)^{-1}\mathbf{C}(0)\mathbf{u}_t$ only makes sense when $\mathbf{u}_t$ has special specifications, $\mathbf{u}^*{}_t$ in particular.

Non-uniqueness of the autoregeressive representations of $\mathbf{F}_t$ and its consequences for estimation are discussed in detail in Appendix C.

5 Estimation

Let $\mathbf{x}_t$ be an n–dimensional $I(1)$ vector, with no deterministic component, described by the following Dynamic Factor Model:

$$\mathbf{x}_t = \mathbf{\Lambda}\mathbf{F}_t+\boldsymbol{\xi}_t, \tag{16}$$
$$(1-L)\mathbf{F}_t = \mathbf{S}(L)^{-1}\mathbf{C}(L)\mathbf{u}_t=\mathbf{U}(L)\mathbf{u}_t, \tag{17}$$

where $\mathbf{\Lambda}$ is $n\times r$, $\mathbf{F}_t$ is $r\times 1$, $\mathbf{u}_t$ is $q\times 1$, $\mathbf{S}(L)$ and $\mathbf{C}(L)$ are as in Definition 3, and $\mathbf{U}(L)$ is $r\times q$, with $n>r>q$.

We have shown in Proposition 3 that when $\mathbf{C}(1)$ has rank $q-d$, then the common factors $\mathbf{F}_t$ have the VECM representation (13) with impulse-response functions $\mathbf{\Psi}(L) = \mathbf{A}(L)^{-1}\mathbf{C}(0)$. Denoting by $\mathbf{\Phi}(L)$ the impulse-response functions of the observables $\mathbf{x}_t$ with respect to $\mathbf{u}_t$, we have:

$$\mathbf{\Phi}(L) = \mathbf{\Lambda}\mathbf{\Psi}(L) = \mathbf{\Lambda}\mathbf{A}(L)^{-1}\mathbf{C}(0). \tag{18}$$

In this Section we discuss estimation of $\mathbf{\Phi}(L)$. In order to do so, we first state some standard assumptions regarding the factor model, then we discuss how to determine the number of common trends, and how to estimate the levels of the common factors.

Let us denote with $\mathbf{\Sigma}_{\Delta\chi}(\theta)$, for $\theta \in [-\pi, \pi]$, the spectral density matrix of $(1-L)\boldsymbol{\chi}_t = \mathbf{\Lambda}(1-L)\mathbf{F}_t$, and with $\lambda_j^{\Delta\chi}(\theta)$ its j–th largest dynamic eigenvalue. Similarly, we denote the covariance matrix of $(1-L)\boldsymbol{\chi}_t$ as $\mathbf{\Gamma}_{\Delta\chi}$ and as $\mu_j^{\Delta\chi}$ its j–largest eigenvalue. Analogous definitions hold for other processes. Let $\tau = q-d$ the number of common trends. We then summarize the main assumptions.

(i) For any $j = 1, \dots, q$, and almost everywhere in $[-\pi, \pi]$, $\lim_{n\to\infty} \lambda_j^{\Delta\chi}(\theta) = \infty$.

(ii) For any $j = 1, \dots, r$, $\lim_{n\to\infty} \mu_j^{\Delta\chi} = \infty$.

(iii) There exists a constant M such that $\lambda_1^{\Delta\xi}(\theta) \leq M$ for any $n \in \mathbb{N}$ and θ a.e.in $[\pi, \pi]$. This implies also that $\mu_1^{\Delta\xi} \leq M$ for any $n \in \mathbb{N}$.

(iv) The loadings matrix is full–rank, i.e. $\text{rk}(\mathbf{\Lambda}) = r$.

Assumptions (i), (ii), and (iii) are taken from Forni et al. (2000), Hallin and Liška (2007), and Forni et al. (2009). In particular, assumption (iii) allows both for some degree of cross–correlation among idiosyncratic components, and for serial correlation in $(1-L)\boldsymbol{\xi}_t$, as well as for unit roots in $\boldsymbol{\xi}_t$. Moreover, under assumptions (i) and (iii) Forni et al. (2000) prove that the spectral density matrix of $(1-L)\mathbf{x}_t$ has just the q largest dynamic eigenvalues diverging as $n \to \infty$ almost everywhere in $[-\pi, \pi]$, the other $n-q$ being uniformly bounded.

Assumption (iv) ensures that the cointegration between common components is entirely due to the cointegration in the common factors, thus allowing for identification of the number of common trends as discussed below.

Lastly, in addition to assumption (i)-(iv) we also require that all spectral densities and covariance matrices are well defined. For details on the required assumptions we refer the reader to Forni et al. (2000) and Hallin and Liška (2007) for the former, and to Bai and Ng (2002) and Bai (2004) for the latter.

5.1 Determining the number of common factors, shocks, and trends

From (17), and assuming for simplicity orthonormal shocks, i.e. $\text{E}[\mathbf{u}_t\mathbf{u}_t'] = \mathbf{I}_q$, we can write the spectral density matrix of $(1-L)\boldsymbol{\chi}_t$ as:

$$\mathbf{\Sigma}_{\Delta\chi}(\theta) = \mathbf{\Lambda}\mathbf{U}(e^{-i\theta})\mathbf{U}'(e^{i\theta})\mathbf{\Lambda}', \tag{19}$$

which, since $\text{rk}(\mathbf{C}(1)) = \tau$, $\text{rk}(\mathbf{S}(1)) = r$, and $\text{rk}(\mathbf{\Lambda}) = r$ by assumption (iv), implies $\text{rk}(\mathbf{U}(1)) = \tau$ and $\text{rk}(\mathbf{\Sigma}_{\Delta\chi}(0)) = \tau$.

As a consequence of (19) and given assumption (i), $\mathbf{\Sigma}_{\Delta\chi}(\theta)$ has at most $\tau \leq q$ diverging eigenvalues at $\theta = 0$. Thus it can be proved that, as $n \to \infty$, the spectral density of $(1-L)\mathbf{x}_t$ has just its τ largest dynamic eigenvalues diverging at $\theta = 0$.

The criterion we propose to determine the number of common trends is a modified version of the criterion by Hallin and Liška (2007), who determine the total number of common shocks, q, on the basis of the asymptotic behavior of the dynamic eigenvalues at all frequencies. Our criterion, instead, determines the number of common trends by studying the behavior of the eigenvalues just at frequency zero. In particular, given a consistent estimator of the spectral density matrix of the data and of its eigenvalues $\widehat{\lambda}_j^{\Delta x}(\theta)$, the estimated number of common trends is such that

$$\widehat{\tau} = \arg\min_{\tau \in [0, \tau_{\max}]} \left[\log \left(\sum_{j=\tau+1}^{n} \widehat{\lambda}_j^{\Delta x}(0) \right) + kp(n,T) \right], \tag{20}$$

where $\tau_{\max}$ is the maximum number of common trends we allow for, $k \in [0, k_{\max}]$ is a constant which helps in tuning the penalization as explained in Hallin and Liška (2007), T is the sample size, and $p(n,T)$ is an appropriate penalty function. In Barigozzi et al. (2013a) we study in detail the asymptotic and numerical properties of the estimator given in (20), and we prove that as $n, T \to \infty$ the number of common trends is consistently estimated.

Finally, the criterion by Hallin and Liška (2007) can be applied to $(1-L)\mathbf{x}_t$ for determining $\widehat{q}$, while the number of common factors, $\widehat{r}$, can be determined by using the information criterion by Bai and Ng (2002) or Alessi et al. (2010) based on principal components of $(1-L)\mathbf{x}_t$. Consistency of the estimated number of common shocks $\widehat{q}$ and common factors $\widehat{r}$, as $n, T \to \infty$, is proved in the cited papers.[4]

An alternative way to determine the number of common trends consists in using the test of Bai and Ng (2004). Note, however, that this test is not developed for the singular case, and hence its properties in this setting are unknown. On the other hand, testing for the number of common trends or the number of cointegration relations in $\widehat{\mathbf{F}}_t$ (e.g. Johansen, 1988, 1991; Stock and Watson, 1988; Phillips and Ouliaris, 1988; Vahid and Engle, 1993) is not possible since, in addition of being developed for the non-singular case, they do not take into account that $\widehat{\mathbf{F}}_t$ is an estimated quantity. Moreover, in our setting testing for cointegration is even more complex since the cointegration rank is not unique (see Appendix C for details). We by-pass all these problems since we work directly on the observables $\mathbf{x}_t$, and since the number of common trends is unique, thus providing, once it is determined, the correct VECM representation for the common factors. Lastly, there exists also a criterion developed by Bai (2004) for the number of non-stationary factors. However, as discussed below, this criterion is based on assumptions that are too restrictive for our setting.

5.2 Estimating the common factors

In order to estimate the VECM representation (13), we need an estimate of the levels of the common factors.

Given a consistent estimator $\widehat{\mathbf{\Gamma}}_{\Delta x}$ of the covariance matrix, we can estimate the loadings, $\widehat{\mathbf{\Lambda}}$, using the eigenvectors corresponding to the r–largest eigenvalues of $\widehat{\mathbf{\Gamma}}_{\Delta x}$. An estimate of the first differences of the common factors is then obtained by taking principal components,

[4] Other criteria or tests for r are Onatski (2010), Kapetanios (2010), and Ahn and Horenstein (2013), and for q are Bai and Ng (2007), Amengual and Watson (2007), and Onatski (2009).

i.e. projecting $(1-L)\mathbf{x}_t$ onto the space spanned by $\widehat{\mathbf{\Lambda}}$. Denote such projection as $\widehat{\mathbf{f}}_t$. Bai and Ng (2004) prove that, independently of whether $\boldsymbol{\xi}_t$ is stationary or not, the common factors can then be consistently estimated as $\widehat{\mathbf{F}}_t = \sum_{t=1}^{T} \widehat{\mathbf{f}}_t$.

However, in a companion paper (Barigozzi et al., 2013b), we show that in macroeconomic databases the estimation method of Bai and Ng (2004) yields poor estimate of the common component, $\widehat{\boldsymbol{\chi}}_t = \widehat{\mathbf{\Lambda}}\widehat{\mathbf{F}}_t$ and hence we suggest a new estimation strategy. While keeping the estimate of the loadings as described above, we propose to estimate the level of the factors as

$$\widehat{\mathbf{F}}_t = (\widehat{\mathbf{\Lambda}}'\widehat{\mathbf{\Lambda}})^{-1}\widehat{\mathbf{\Lambda}}'\mathbf{x}_t, \tag{21}$$

i.e. by projecting the data $\mathbf{x}_t$ onto the space spanned by the estimated loadings. We show in Barigozzi et al. (2013b) that compared to the other available methods, our estimator performs satisfactorily well on both macroeconomic and different simulated datasets where we allow for both stationary and non–stationary idiosyncratic components. In particular, as $n, T \to \infty$, we can prove that (21) estimates consistently the space spanned by the true common factors. Finally, notice that, in order to identify the scale of the loadings we have to impose a normalization on $\widehat{\mathbf{\Lambda}}$. For example, we can define $\widehat{\mathbf{\Lambda}}$ to be given by $\sqrt{n}$ times the normalized eigenvectors of the covariance matrix (see Bai and Ng, 2002).

Another possible estimator is given by Bai (2004) who suggests to estimate $\mathbf{F}_t$ based on principal components applied to a consistent estimator of the covariance matrix of $\mathbf{x}_t$, i.e. of the data in levels. Such procedure is consistent only under the assumption that the idiosyncratic component is stationary (see also Peña and Poncela, 2004). However, this assumption might be plausible in some database but it is unrealistic in macroeconomic databases as it implies cointegration among all the variables in the panel. This can be easily seen with an example. Suppose we have a model with just one factor: $\mathbf{x}_t = \mathbf{\Lambda} F_t + \boldsymbol{\xi}_t$, where $F_t \sim I(1)$ and $\boldsymbol{\xi}_t \sim I(0)$. Let us take a linear combination of the i-th and the j-th variable, say $z_t = x_{it} - \beta x_{jt}$. Then we can also write $z_t = (\lambda_i - \beta\lambda_j)F_t + (\xi_{it} - \beta\xi_{jt})$. If we take $\beta = \lambda_i/\lambda_j$, we get $z_t = \xi_{it} - \beta\xi_{jt}$, which is stationary since $\xi_{it}, \xi_{jt} \sim I(0)$. Hence x_{it}, and x_{jt} are cointegrated. This example can be generalized to prove that if $\boldsymbol{\xi}_t \sim I(0)$, then any group of $r+1$ variables is cointegrated.

5.3 Estimating the impulse-response functions

Once we have estimates of the number of common factors, common shocks and common trends, and of the factors and their loadings, we can estimate the VECM representation (13). The procedure presented here is analogous to the one proposed by Forni et al. (2009), with the exception of the Error Correction term.

First, we estimate the VECM in (13) on $\widehat{\mathbf{F}}_t$ with $\widehat{r} - \widehat{\tau}$ cointegration relations, by using one of the available methods (Engle and Granger, 1987; Johansen, 1988, 1991; Stock and Watson, 1993; Phillips, 1995). We thus obtain $\widehat{\mathbf{A}}^*(L)$, $\widehat{\boldsymbol{\alpha}}$, $\widehat{\boldsymbol{\beta}}$, $\widehat{\mathbf{h}}$, and an $r \times 1$ vector of residuals $\widehat{\mathbf{v}}_t$. Then, we estimate $\widehat{\mathbf{C}}(0)$ as the matrix of eigenvectors corresponding to the q largest eigenvalues of the sample covariance matrix of the residuals, $\widehat{\mathbf{\Gamma}}_{\mathbf{v}} = \widehat{\mathbf{v}}_t'\widehat{\mathbf{v}}_t/T$.

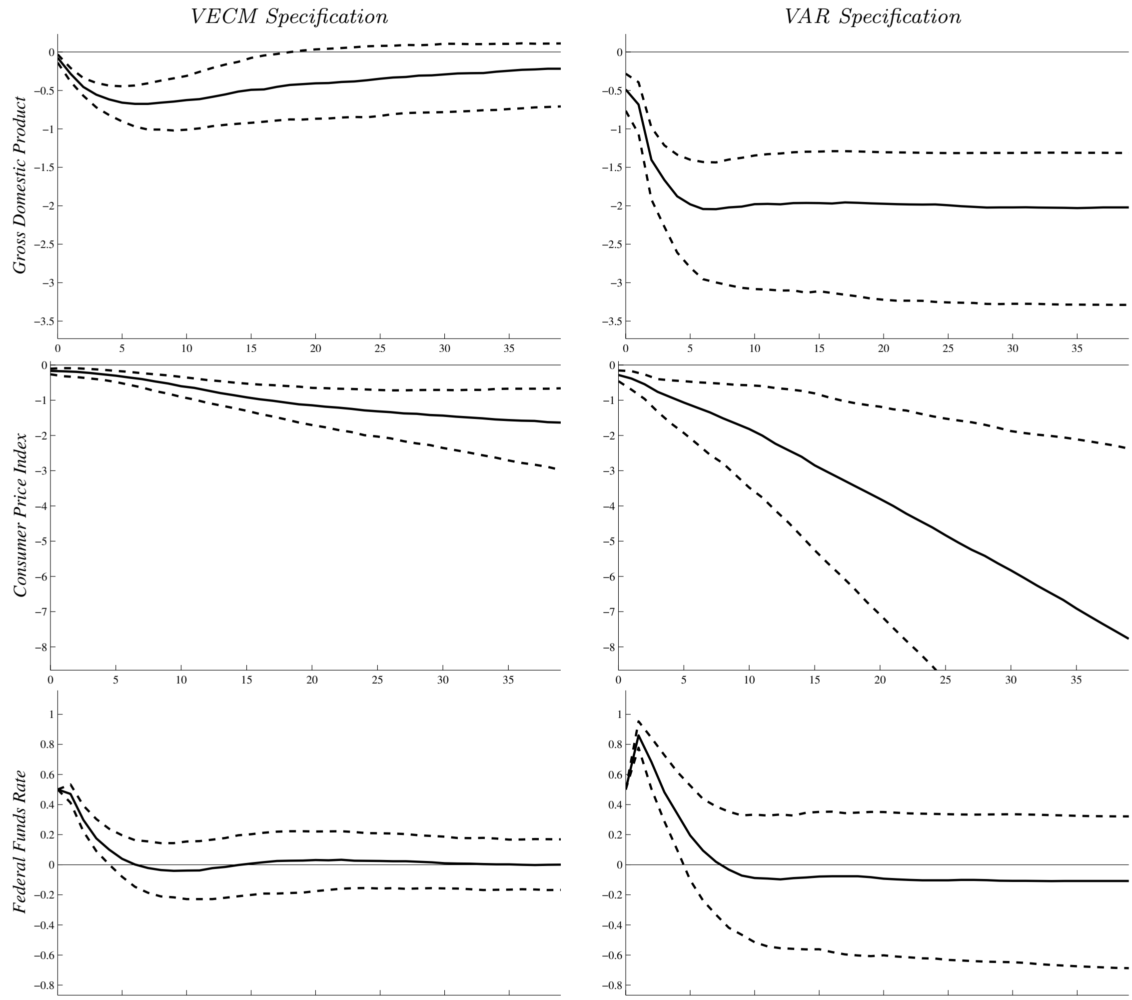
The estimated impulse-response functions are then given by (18), i.e. $\widehat{\mathbf{\Phi}}(L) = \widehat{\mathbf{\Lambda}}(\widehat{\mathbf{A}}(L))^{-1}$ $\widehat{\mathbf{C}}(0)$, where $\widehat{\mathbf{A}}(L) = (1-L)\widehat{\mathbf{A}}^*(L) + \widehat{\mathbf{A}}(1)$ with $\widehat{\mathbf{A}}(1) = \widehat{\boldsymbol{\alpha}}\widehat{\boldsymbol{\beta}}'$.

6 Empirical Application

The empirical analysis is carried out on a panel of 103 quarterly series from 1960Q3 to 2012Q4 describing the US economy. The variables cover 12 different categories: Industrial Production, Consumer Price Indexes, Producer Price Indexes, Monetary Aggregates, Banking, GDP and its Components, Housing Sector, Productivity and Costs, Interest Rates, Employment and Population, Survey, and Financial Markets.[5] All the variables in the dataset are detrended with OLS in order to remove the deterministic components.

The model is estimated as explained in the previous Section. In particular, we set $\widehat{r} = 7$, $\widehat{q} = 3$, and $\widehat{\tau} = 1$. Moreover, it is well known that in (18) impulse–response functions and common shocks are identified only up to multiplication by a $q \times q$ invertible matrix. Hence, in order to achieve identification, it is necessary to impose economically meaningful restrictions analogous to those imposed in Structural VAR analysis. In particular, the Structural VECM was first introduced in King et al. (1991) and then developed in Paruolo (1997a), Gonzalo and Ng (2001), Vlaar (2004), Omtzigt and Paruolo (2005), and Pagan and Pesaran (2008) among others. In this Section we provide results using two different identification schemes.

Figure 1: Impulse Response Functions to a Monetary Policy Shock



In each plot, solid and dotted lines represent, respectively, the estimated impulse–response and the 68% bootstrap confidence bands. The monetary policy shock is normalized so that at impact it raises the Federal Funds rate by 50 basis points. The confidence band are computed in the standard way as in (Efron and Tibshirani, 1993).

[5]The complete list of variables and transformations is reported in Appendix D.

First, we identify a monetary policy shock using sign restrictions on the first three lags, i.e. by assuming that a shock increases the Federal Funds Rate, while it decreases GDP and the Consumer Prices Index. In Figure 1 we present the impulse–response functions for selected variables of interests together with 68% bootstrap confidence bands. For a description of the sign restrictions and of the bootstrap algorithm we refer the reader to Barigozzi et al. (2013) and Luciani (2013). The left column of Figure 1 shows impulse–response functions obtained by estimating a VECM on the common factors as in (13), while the right column shows the impulse–responses obtained by estimating (13) without the Error Correction term, i.e. estimating a VAR in first differences. Results show that when considering also the presence of temporary shocks (left column), the reaction of the Federal Funds Rate does not change, whereas the responses of GDP and the Consumer Price Index change substantially. In particular, GDP decreases as far as -0.65% but then reverts to zero, while the Consumer Price Index does not revert to zero but it stabilizes around -2% rather than exhibiting the explosive behavior estimated with the VAR specification.

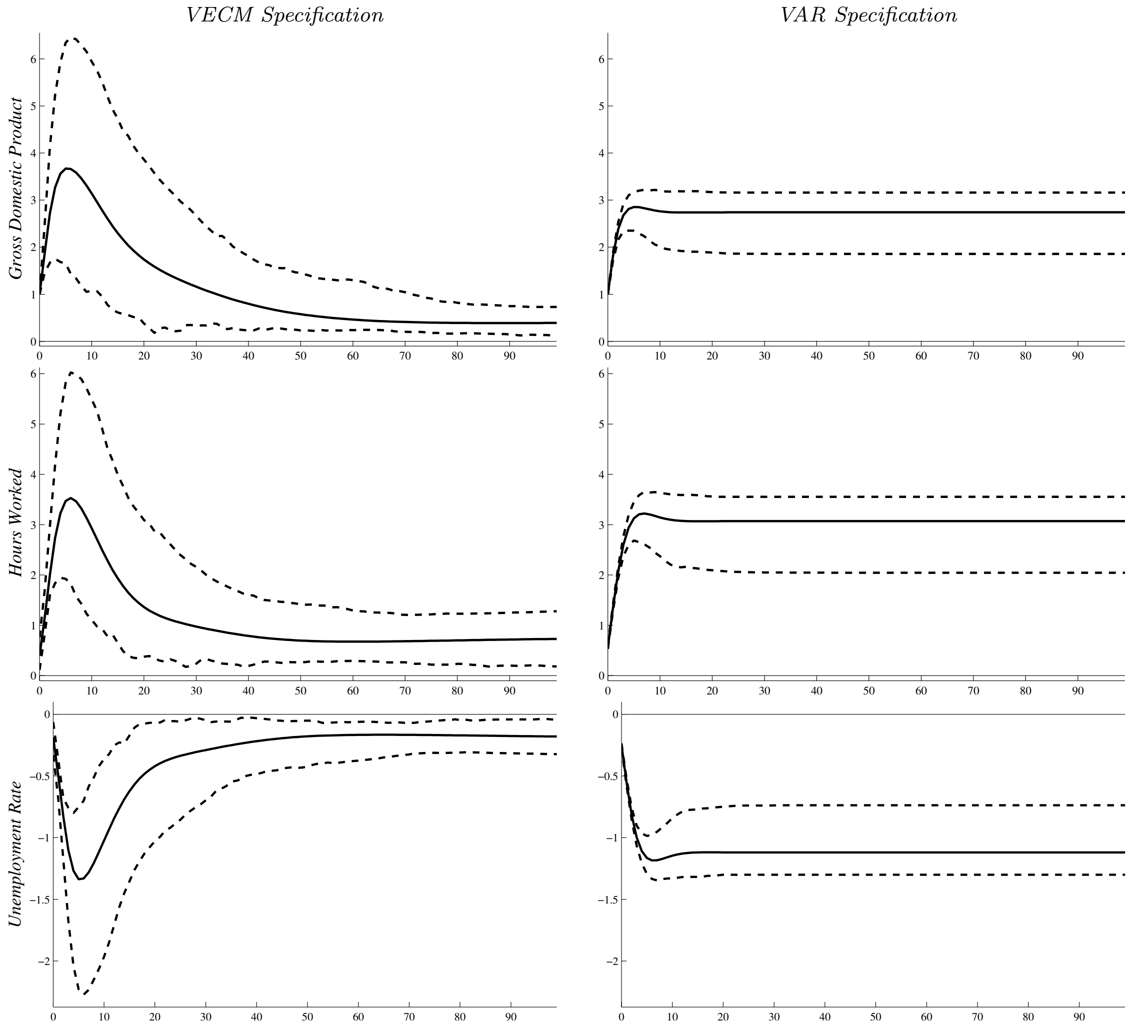
Second, we identify a technology shock by long–run restrictions as in Blanchard and Quah (1989), that is by assuming that the other two (transitory) shocks have no long–run effects on real variables. In Figure 2 we present the impulse–response functions for GDP, Hours Worked, Unemployment Rate together with 68% bootstrap confidence bands. Results for the VECM specification (left column) show that all variables have a hump shaped response, with a maximum between six and seven quarters after the shock and then reverting without, however, reaching the steady state. The response of GDP is positive and consistent with a process of diffusion of technology as described for example in Lippi and Reichlin (1994), and it has also been found by Dedola and Neri (2007) and Smets and Wouters (2007). With respect to hours worked, there is a large debate in the macroeconomic literature on whether hours should increase –as predicted by a Real Business Cycle model– or should decrease –as predicted by a New Keynesian model– after a positive technology shock. The empirical evidence is mixed with some authors suggesting that it decreases (Galì, 1999; Francis and Ramey, 2005) and some others pointing towards the opposite direction (Christiano et al., 2003; Dedola and Neri, 2007). Our results are consistent with predictions of a Real Business Cycle model.

To sum up, in this Section we have shown the importance of a correct specification for the autoregressive representation of $(1-L)\mathbf{F}_t$. Our exercise shows that using the Error Correction representation derived in Proposition 3 produces impulse-response functions that are consistent with standard macroeconomic theory, whereas estimating a VAR for the factors in first differences, which is the usual practice with Dynamic Factor Models, not necessarily does. In particular, due to cumulation of impulse-responses, when adopting a VAR specification the estimated responses are unrealistically high and persistent.

7 Summary and conclusions

In this paper we studied non-stationary Dynamic Factor Models. To this end, we linked the literature on stationary Dynamic Factor Models (Stock and Watson, 2005; Forni et al., 2009) with the literature on cointegration (Engle and Granger, 1987; Johansen, 1988, 1991), common trends and common cycles (Stock and Watson, 1988; Vahid and Engle, 1993; Lippi and Reichlin, 1994), and singular stochastic processes (Anderson and Deistler, 2008a,b). In particular, we analyze the autoregressive representation of the $I(1)$ singular vector $\mathbf{F}_t$, the relationship between cointegration rank and the number transitory shocks, and estimation of

Figure 2: Impulse Response Functions to a Technology Shock



In each plot, solid and dotted lines represent, respectively, the estimated impulse–response and the 68% bootstrap confidence bands. The permanent shock is normalized so that at impact it raises GDP of 1%. The confidence band are computed as in Hall (1992). This procedure is recommended by Brüggemann (2006) since, when imposing long-run restrictions, the standard percentile bootstrap interval is less informative about the sign of the impulse-response function.

the impulse-response functions.

Under the assumption that the common factors have rational spectral density, we prove that for generic values of the parameters there exists a *finite* Error Correction representation for $(1-L)\mathbf{F}_t$, where the number of error terms is equal to the number of transitory shocks, plus the difference between the dimension of $\mathbf{F}_t$ and the dimension of $\mathbf{u}_t$, i.e. the vector of common shocks. Although presented in the Factor Model context, our results hold for any singular vector with rational spectral density.

We use this result to construct impulse-response functions. As shown in Stock and Watson (2005) and Forni et al. (2009), identifying restrictions used in Structural VAR analysis can be applied to the identification of structural shocks and impulse-response functions in stationary Dynamic Factor Models (see Giannone et al., 2005; Stock and Watson, 2005; Forni et al., 2009; Forni and Gambetti, 2010; Barigozzi et al., 2013; Luciani, 2013, for applications). Here we study the case in which the common factors are $I(1)$ and cointegrated. Once the autoregressive representation is correctly specified, the identifying restrictions of the Structural VECM analysis can be applied to $I(1)$ Dynamic Factor Models.

Results of an empirical analysis on a US quarterly database illustrate the importance of a correct specification for the autoregressive representation of $(1-L)\mathbf{F}_t$. Our exercise shows that our approach produces impulse-response functions that are consistent with standard macroeconomic theory, whereas the usual practice of estimating a VAR on $(1-L)\mathbf{F}_t$ not necessarily does.

References

Ahn, S. C. and A. R. Horenstein (2013). Eigenvalue ratio test for the number of factors. *Econometrica 81*, 1203–1227.

Alessi, L., M. Barigozzi, and M. Capasso (2010). Improved penalization for determining the number of factors in approximate static factor models. *Statistics and Probability Letters 80*, 1806–1813.

Amengual, D. and M. W. Watson (2007). Consistent estimation of the number of dynamic factors in a large N and T panel. *Journal of Business and Economic Statistics 25*, 91–96.

Anderson, B. D. and M. Deistler (2008a). Generalized linear dynamic factor models–a structure theory. *IEE Conference on Decision and Control*.

Anderson, B. D. and M. Deistler (2008b). Properties of zero-free transfer function matrices. *SICE Journal of Control, Measurement and System Integration 1*, 284Ð292.

Bai, J. (2004). Estimating cross-section common stochastic trends in nonstationary panel data. *Journal of Econometrics 122*, 137–183.

Bai, J. and S. Ng (2002). Determining the number of factors in approximate factor models. *Econometrica 70*, 191–221.

Bai, J. and S. Ng (2004). A PANIC attack on unit roots and cointegration. *Econometrica 72*, 1127–1177.

Bai, J. and S. Ng (2007). Determining the number of primitive shocks in factor models. *Journal of Business and Economic Statistics 25*, 52–60.

Barigozzi, M., A. M. Conti, and M. Luciani (2013). Do euro area countries respond asymmetrically to the common monetary policy? *Oxford Bulletin of Economics and Statistics*. forthcoming.

Barigozzi, M., M. Lippi, and M. Luciani (2013a). Determining the number of common trends and common cycles in non-stationary dynamic factor models. Université libre de Bruxelles.

Barigozzi, M., M. Lippi, and M. Luciani (2013b). Estimating I(1) dynamic factor models. Université libre de Bruxelles.

Blanchard, O. J. and D. Quah (1989). The dynamic effects of aggregate demand and supply disturbance. *The American Economic Reiview 79*, 655–673.

Brüggemann, R. (2006). Finite sample properties of impulse response intervals in SVECMs with long–run identifying restrictions. Discussion Papers 2006-021, SFB 649.

Christiano, L. J., M. Eichenbaum, and R. Vigfusson (2003). What happens after a technology shock? International Finance Discussion Papers 768, Board of Governors of the Federal Reserve System (U.S.).

Dedola, L. and S. Neri (2007). What does a technology shock do? a var analysis with model-based sign restrictions. *Journal of Monetary Economics 54*, 512–549.

Deistler, M., B. D. Anderson, A. Filler, C. Zinner, and W. Chen (2010). Generalized linear dynamic factor models: An approach *via* singular autoregressions. *European Journal of Control*, 211–224.

Efron, B. and R. Tibshirani (1993). *An Introduction to the Bootstrap.* Chapman & Hall.

Engle, R. F. and C. W. J. Granger (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica 55*, 251–76.

Forni, M. and L. Gambetti (2010). The dynamic effects of monetary policy: A structural factor model approach. *Journal of Monetary Economics 57*, 203–216.

Forni, M., D. Giannone, M. Lippi, and L. Reichlin (2009). Opening the Black Box: Structural Factor Models versus Structural VARs. *Econometric Theory 25*, 1319–1347.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2000). The Generalized Dynamic Factor Model: Identification and Estimation. *The Review of Economics and Statistics 82*, 540–554.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2005). The Generalized Dynamic Factor Model: One Sided Estimation and Forecasting. *Journal of the American Statistical Association 100*, 830–840.

Forni, M., M. Hallin, M. Lippi, and P. Zaffaroni (2014). Dynamic factor models with infinite-dimensional factor spaces: one-sided representations. *Journal of Econometrics,* forthcoming.

Forni, M. and M. Lippi (2001). The Generalized Dynamic Factor Model: Representation Theory. *Econometric Theory 17*, 1113–1141.

Forni, M. and M. Lippi (2010). The Unrestricted Dynamic Factor Model: One-sided Representation Results. *Journal of Econometrics 163*, 23–28.

Francis, N. and V. A. Ramey (2005). Is the technology-driven real business cycle hypothesis dead? shocks and aggregate fluctuations revisited. *Journal of Monetary Economics 52*, 1379–1399.

Galì, J. (1999). Technology, employment, and the business cycle: Do technology shocks explain aggregate fluctuations? *American Economic Review 199*, 249Ð71.

Giannone, D., L. Reichlin, and L. Sala (2005). Monetary policy in real time. In M. Gertler and K. Rogoff (Eds.), *NBER Macroeconomics Annual 2004.* MIT Press.

Giannone, D., L. Reichlin, and D. Small (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics 55*, 665–676.

Gonzalo, J. and S. Ng (2001). A systematic framework for analyzing the dynamic e!ects of permanent and transitory shocks. *Journal of Economic Dynamics and Control 25*, 1527–1546.

Hall, P. (1992). *The Bootstrap and Edgeworth Expansion.* New York: Springer.

Hallin, M. and R. Liška (2007). Determining the number of factors in the general dynamic factor model. *Journal of the American Statistical Association 102*, 603–617.

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

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica 59*, 1551–80.

Johansen, S. (1995). *Likelihood-based inference in cointegrated vector autoregressive models* (First ed.). Oxford: Oxford University Press.

Kapetanios, G. (2010). A testing procedure for determining the number of factors in approximate factor models with large datasets. *Journal of Business and Economic Statistics 28*, 397–409.

King, R., C. Plosser, J. H. Stock, and M. W. Watson (1991). Stochastic trends and economic fluctuations. *American Economic Review 81*, 819–840.

Lancaster, P. and M. Tismenetsky (1985). *The theory of matrices* (Second ed.). New York: Academic Press.

Lippi, M. and L. Reichlin (1994). Common and uncommon trends and cycles. *European Economic Review 38*, 624–635.

Luciani, M. (2013). Monetary policy and the housing market: A structural factor analysis. *Journal of Applied Econometrics*. forthcoming.

Luciani, M. (2014). Forecasting with approximate dynamic factor models: the role of *non-pervasive* shocks. *International Journal of Forecasting 30*, 20–29.

Omtzigt, P. and P. Paruolo (2005). Impact factors. *Journal of Econometrics 128*, 31–68.

Onatski, A. (2009). Testing hypotheses about the number of factors in large factor models. *Econometrica 77*, 1447–1479.

Onatski, A. (2010). Determining the number of factors from empirical distribution of eigenvalues. *Review of Economics and Statistics 92*, 1004–1016.

Pagan, A. R. and M. H. Pesaran (2008). Econometric analysis of structural systems with permanent and transitory shocks. *Journal of Economic Dynamics and Control 32*, 3376–3395.

Paruolo, P. (1997a). Asymptotic inference on the moving average impact matrix in cointegrated $I(1)$ VAR systems. *Econometric Theory 13*, 79–118.

Paruolo, P. (1997b). Standard errors for the long run variance matrix. *Econometric Theory 13*, 305–306.

Peña, D. and P. Poncela (2004). Nonstationary dynamic factor analysis. *Journal of Statistical Planning and Inference 136*, 1237–1257.

Phillips, P. C. (1995). Fully modified least squares and vector autoregression. *Econometrica 63*, 1023–1078.

Phillips, P. C. (1998). Impulse response and forecast error variance asymptotics in nonstationary VARs. *Journal of Econometrics 83*, 21–56.

Phillips, P. C. and S. Ouliaris (1988). Testing for cointegration using principal components methods. *Journal of Economic Dynamics and Control 12*, 205–230.

Sims, C., J. H. Stock, and M. W. Watson (1990). Inference in linear time series models with some unit roots. *Econometrica 58*, 113–144.

Smets, F. and R. Wouters (2007). Shocks and frictions in US business cycles: A Bayesian DSGE approach. *American Economic Review 97*, 586–606.

Stock, J. H. and M. W. Watson (1988). Testing for common trends. *Journal of the American Statistical Association 83*, 1097–1107.

Stock, J. H. and M. W. Watson (1993). A simple estimator of cointegrating vectors in higher order integrated systems. *Econometrica 61*, 783–820.

Stock, J. H. and M. W. Watson (2002a). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association 97*, 1167–1179.

Stock, J. H. and M. W. Watson (2002b). Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics 20*, 147–162.

Stock, J. H. and M. W. Watson (2005). Implications of dynamic factor models for VAR analysis. Working Paper 11467, NBER.

Vahid, F. and R. F. Engle (1993). Common trends and common cycles. *Journal of Applied Econometrics 8*, 341–360.

van der Waerden, B. L. (1953). *Modern Algebra* (Second ed.), Volume I. New York: Frederick Ungar.

Vlaar, P. J. G. (2004). On the asymptotic distribution of impulse response functions with long-run restrictions. *Econometric Theory 20*, 891–903.

Appendix

A Proof of Proposition 2

Remark 1 *Suppose that the statement $S(p)$, depending on a vector $p \in \Pi$, is equivalent to a set of polynomial equations for the parameters, for example the statement that rank$(\mathbf{M}(1)) < q$. Statement $S(p)$ is true either for a nowhere dense subset of Π or for the whole Π. Thus, if the statement is false for one point in Π, it is true for a nowhere dense subset of Π. Moreover, $S(p)$ can be obviously extended to any $p \in \mathbb{R}^\lambda$ and, as Π is an open subset of $\mathbb{R}^\lambda$, if the statement is false for one point in $\mathbb{R}^\lambda$, it is true for a nowhere dense subset of $\mathbb{R}^\lambda$ and therefore of Π.*

Remark 2 *Consider the polynomials*

$$A(z) = a_0 z^n + a_1 z^{n-1} + \cdots + a_n, \quad B(z) = b_0 z^m + b_1 z^{m-1} + \cdots + a_m.$$

The resultant of A and B is defined as a polynomial

$$R(a_0, a_1, \ldots, a_n; b_0, b_1, \ldots, b_m),$$

see van der Waerden (1953), pp. 83-4. The resultant vanishes if and only if either (i) $a_0 = 0$ and $b_0 = 0$, or (ii) $a_0 \neq 0$ or $b_0 \neq 0$, and A and B have a common root. Now suppose that the coefficients a_i and b_j are polynomial functions of $p \in \Pi$. Then, by Remark 1, if there exists one point $\tilde{p} \in \Pi$ (or $\tilde{p} \in \mathbb{R}^\lambda$) such that $a_0(\tilde{p}) \neq 0$ or $b_0(\tilde{p}) \neq 0$, and $R(\tilde{p}) \neq 0$, then generically A and B have no roots in common.

Starting with

$$\mathbf{C}(z) = \boldsymbol{\xi}\boldsymbol{\eta}' + (1-z)\mathbf{D} + (1-z)^2\mathbf{E}(z),$$

we obtain, see Section 4,

$$\begin{aligned}
\boldsymbol{\zeta}\mathbf{C}(z) &= \begin{pmatrix} (1-z)\mathbf{I}_c & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{r-c} \end{pmatrix} \left\{ \begin{pmatrix} \boldsymbol{\xi}'_\perp \mathbf{D} \\ \boldsymbol{\xi}'\boldsymbol{\xi}\boldsymbol{\eta}' \end{pmatrix} + (1-L) \begin{pmatrix} \boldsymbol{\xi}'_\perp \mathbf{E}(z) \\ \boldsymbol{\xi}'\mathbf{D} \end{pmatrix} + (1-z)^2 \begin{pmatrix} \mathbf{0}_{c\times q} \\ \boldsymbol{\xi}'\mathbf{E}(z) \end{pmatrix} \right\} \\
&= \begin{pmatrix} (1-z)\mathbf{I}_c & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{r-c} \end{pmatrix} \mathbf{M}(z).
\end{aligned}$$

With no loss of generality we can assume that $r = q+1$. We denote by $\mathbf{M}_1(z)$ and $\mathbf{M}_2(z)$ the $q \times q$ matrices obtained by dropping the first and the last row of $\mathbf{M}(z)$ respectively. The degrees of the polynomials $\det(\mathbf{M}_1(z))$ and $\det(\mathbf{M}_2(z))$ are $d_1 = (q-d)(s_1+2) + d(s_1+1)$ and $d_2 = (q-d-1)(s_1+2) + (d+1)(s_1+1)$ respectively. Let Q_1 and Q_2 be the leading coefficients of $\det(\mathbf{M}_1(z))$ and $\det(\mathbf{M}_2(z))$ respectively and R_M their resultant.

Let us now specify define a subfamily of $\mathbf{M}(z)$, denoted by $\mathbf{M}_\phi(z)$:

$$\mathbf{K} = \begin{pmatrix} \mathbf{0}_{1\times(q-d)} & 1 & \mathbf{0}_{1\times d} \end{pmatrix}, \quad \mathbf{H} = \begin{pmatrix} \mathbf{0}_{d\times(q+1-d)} & \mathbf{I}_d \end{pmatrix}, \quad \mathbf{E}_\phi(z) = \begin{pmatrix} \mathbf{E}_{\phi 1}(z) \\ \mathbf{E}_{\phi 2}(z) \\ \mathbf{E}_{\phi 3}(z) \end{pmatrix},$$

$$\boldsymbol{\eta}'_\phi = \begin{pmatrix} \mathbf{0}_{(q-d)\times d} & \mathbf{I}_{q-d} \end{pmatrix}, \quad \boldsymbol{\xi}_\phi = \begin{pmatrix} \mathbf{I}_{q-d} \\ \mathbf{0}_{c\times(q-d)} \end{pmatrix}, \quad \boldsymbol{\xi}'_{\phi\perp} = \begin{pmatrix} \mathbf{K} \\ \mathbf{H} \end{pmatrix}, \quad \mathbf{D}_\phi = \begin{pmatrix} \mathbf{H}' & \mathbf{0}_{(q+1)\times(q-d)} \end{pmatrix},$$

where $\mathbf{E}_{\phi 1}(z)$ is $(q-d)\times q$, $\mathbf{E}_{\phi 2}(z)$ is $1\times q$ and $\mathbf{E}_{\phi 3}(z)$ is $d\times q$. Moreover:

$$\mathbf{E}_{\phi 1}(z) = \begin{pmatrix} & k_1(z) & h_1(z) & \cdots & 0 \\ \mathbf{0}_{(q-d)\times d} & & \ddots & \ddots & \\ & & & & h_{q-d-1}(z) \\ & 0 & \cdots & & k_{q-d}(z) \end{pmatrix}$$

$$\mathbf{E}_{\phi 2}(z) = \begin{pmatrix} e(z) & \mathbf{0}_{1\times(q-1)} \end{pmatrix}$$

$$\mathbf{E}_{\phi 3}(z) = \begin{pmatrix} f_1(z) & g_1(z) & \cdots & 0 & \\ & \ddots & \ddots & & \mathbf{0}_{d\times(q-d-1)} \\ 0 & \cdots & f_d(z) & g_d(z) & \end{pmatrix},$$

all the polynomials above being of degree s_1.

We have:

$$\mathbf{M}_{\phi}(z) = \begin{pmatrix} \mathbf{0}_{1\times d} & \mathbf{0}_{1\times(q-d)} \\ \mathbf{I}_d & \mathbf{0}_{d\times(q-d)} \\ \mathbf{0}_{(q-d)\times d} & \mathbf{I}_{q-d} \end{pmatrix} + (1-z)\begin{pmatrix} \mathbf{E}_{\phi 2}(z) \\ \mathbf{E}_{\phi 3}(z) \\ \mathbf{0}_{(q-d)\times q} \end{pmatrix} + (1-z)^2 \begin{pmatrix} \mathbf{0}_{1\times q} \\ \mathbf{0}_{d\times q} \\ \mathbf{E}_{\phi 1}(z) \end{pmatrix}$$

The matrices $\mathbf{M}_{\phi 1}(z)$ and $\mathbf{M}_{\phi 2}(z)$ are upper- whereas the second is lower-triangular respectively. We have

$$\begin{aligned} \det(\mathbf{M}_{\phi 1}(z)) &= [1+(1-z)f_1(z)]\cdots[(1+(1-z)f_d(z)] \\ &\quad \times [1+(1-z)^2 k_1(z)]\cdots[1+(1-z)^2 k_{q-d}(z)] \\ \det(\mathbf{M}_{\phi 2}(z)) &= (1-z)^{2q-d-1} e(z)[g_1(z)\cdots g_d(z)][h_1(z)\cdots h_{q-d-1}(z)] \end{aligned}$$

Now:

(i) The leading coefficient of $\det(\mathbf{M}_{\phi 1}(z))$, call it $Q_{\phi 1}$, corresponding to z^{d_1}, is the product of the leading coefficients of the polynomials $k_j(z)$, $j=1,\ldots,(q-d)$ and $f_i(z)$, $i=1,\ldots,d$. Trivially, for generic values of the parameters of $\mathbf{M}_{\phi}(z)$, $Q_{\phi 1}\neq 0$.

(ii) Suppose that $\tilde{z}$ is a root $\det(\mathbf{M}_1(z))$. $\tilde{z}$ must be a root of one of the factors of $\det(\mathbf{M}_{\phi 1}(z))$, for example of $1+(1-z)f_1(z)$. Because the polynomials $\det(\mathbf{M}_{\phi 1}(z))$ and $\det(\mathbf{M}_{\phi 2}(z))$ have no parameters in common, it is easy to see that there exist parameters such that $\tilde{z}$ is not a root of any of the factors of $\det(\mathbf{M}_{\phi 2}(z))$. Because the roots of a polynomial are continuous functions of the coefficients and because the parameters vector belongs to an open set, there exists a vector of parameters, call it $\boldsymbol{\vartheta}_{\phi}$, such that (1) the leading coefficient of $\det(\mathbf{M}_{\phi 1}(z))$ is not zero and (2) $\det(\mathbf{M}_{\phi 1}(z))$ and $\det(\mathbf{M}_{\phi 2}(z))$ have no roots in common. Thus for $\boldsymbol{\vartheta}_{\phi}$ we have $Q_{\phi 1}\neq 0$ and $R_{M_{\phi}}\neq 0$. As both the leading coefficient and the resultant are polynomials in the parameters $\boldsymbol{\vartheta}$, by Remarks 1 $Q_1\neq 0$ and $R_M\neq 0$ generically in Π. This implies, by Remark 2, that $\mathbf{M}(z)$ is generically zeroless. Q.E.D.

B Stationary solutions of $(1-L)\mathbf{y}_t = (1-L)\mathbf{z}_t$

Assume that $\mathbf{z}_t$ is a g-dimensional, weakly stationary process with the moving-average representation $\mathbf{z}_t = \mathbf{M}(L)\mathbf{v}_t$, where (i) $\mathbf{v}_t$ is an s-dimensional white noise and $s \leq g$, (ii) v_{jt} belongs to $L_2(\Omega, \mathcal{F}, P)$, for $j = 1, 2, \ldots, k$, (iii) $\mathbf{M}(L)$ is a $g \times s$ square-summable matrix. Moreover, assume that z_{kt} and v_{jt}, for $t \in \mathbb{Z}$, $k = 1, 2, \ldots, g$, $j = 1, 2, \ldots, s$, span the same subspace of $L_2(\Omega, \mathcal{F}, P)$. Assumption (iii) holds, for example, if $\mathbf{z}_t = \mathbf{M}(L)\mathbf{v}_t$ is a fundamental representation of $\mathbf{z}_t$. Moreover, suppose that $\mathbf{y}_t$ fulfills

$$(1-L)\mathbf{y}_t = (1-L)\mathbf{z}_t \tag{22}$$

and $y_{jt} \in L_2(\Omega, \mathcal{F}, P)$ for $j = 1, 2, \ldots, g$. Because $\mathbf{z}_t$ trivially fulfills (22),

$$\mathbf{y}_t = \mathbf{K} + \mathbf{z}_t$$

where $\mathbf{K}$ is a g-dimensional stochastic variable belonging to $L_2(\Omega, \mathcal{F}, P)$. Let

$$\mathbf{K} = \mathbf{N}(L)\mathbf{v}_0 + \mathbf{H} = \mathbf{V} + \mathbf{H}.$$

be the orthogonal projection of $\mathbf{K}$ on the space spanned by $\mathbf{v}_k$, $k \in \mathbb{Z}$. In general the filter $\mathbf{N}(L)$ is two-sided. We have

$$\mathrm{E}(\mathbf{y}_t\mathbf{y}_t') = \mathrm{E}(\mathbf{V}\mathbf{V}') + \mathrm{E}(\mathbf{H}\mathbf{H}') + \mathrm{E}(\mathbf{z}_t\mathbf{z}_t') + \mathrm{E}(\mathbf{V}\mathbf{z}_t') + \mathrm{E}(\mathbf{z}_t\mathbf{V}').$$

The last two terms tend to zero when t tends to infinity (by the same argument used to prove that the covariances of a moving average tend to zero as the lag tends to infinity). Therefore, if $\mathbf{y}_t$ is weakly stationary they must be zero for all $t \in \mathbb{Z}$. This implies that $\mathbf{V}$ is orthogonal to $\mathbf{z}_t$ for $t \in \mathbb{Z}$ and therefore to $\mathbf{v}_t$ for $t \in \mathbb{Z}$. As $\mathbf{V}$ is an average of $\mathbf{v}_t$, $\mathbf{V} = 0$. In conclusion, all the stationary solutions of (22) are $\mathbf{y}_t = \mathbf{K} + \mathbf{z}_t$ with $\mathbf{K}$ orthogonal to $\mathbf{z}_t$ for all $t \in \mathbb{Z}$. Lastly, $\mathbf{y}_t$ has a spectral density if and only if $\mathbf{K}$ is a constant, in that case the spectral densities of $\mathbf{y}_t$ and $\mathbf{z}_t$ coincide. This proves the statement in the last paragraph of Section 3.2.

C Non uniqueness

In Proposition 3 we prove that a singular $I(1)$ vectors has a finite Error Correction representation with $r - q + d$ error corrections. Unfortunately, this representation is not unique. There are two type of non uniqueness here: the first is related with the number of error correction, while the second is related with the degree of the autoregressive polynomial.

C.1 Alternative representations with different numbers of error terms

Let, for simplicity, $\mathbf{S}(L) = \mathbf{I}_r$ and consider the following example, with $r = 3$, $q = 2$, $c = 2$, so that $d = 1$:

$$\boldsymbol{\xi}' = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}$$
$$\boldsymbol{\eta}' = \begin{pmatrix} 1 & 2 \end{pmatrix}$$
$$\boldsymbol{\xi}_\perp' = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix}$$

We have,

$$(1-L)\begin{pmatrix}\boldsymbol{\xi}'_{\perp}\\ \boldsymbol{\xi}'\end{pmatrix}\mathbf{F}_t=\begin{pmatrix}1-L & 0 & 0\\ 0 & 1-L & 0\\ 0 & 0 & 1\end{pmatrix}\left\{\begin{pmatrix}d_{11}-d_{21} & d_{12}-d_{22}\\ d_{21}-d_{31} & d_{22}-d_{32}\\ 3 & 6\end{pmatrix}+(1-L)\mathbf{G}(L)\right\}\mathbf{u}_t,$$

where $(1-L)\mathbf{G}(L)$ gathers the second and third terms within curly brackets in the second line of (12). If the first matrix within the curly brackets has full rank, we can proceed as in Proposition 3 and obtain an EC representation with errors

$$\boldsymbol{\xi}'_{\perp}\mathbf{F}_t=\begin{pmatrix}F_{1t}-F_{2t}\\ F_{2t}-F_{3t}\end{pmatrix}.$$

However, we also have

$$(1-L)\begin{pmatrix}\boldsymbol{\xi}'_{\perp}\\ \boldsymbol{\xi}'\end{pmatrix}\mathbf{F}_t=\begin{pmatrix}1-L & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\left\{\begin{pmatrix}d_{11}-d_{21} & d_{12}-d_{22}\\ (1-L)(d_{21}-d_{31}) & (1-L)(d_{22}-d_{32})\\ 3 & 6\end{pmatrix}\right.$$
$$\left.+(1-L)\tilde{\mathbf{G}}(L)\right\}\mathbf{u}_t=\begin{pmatrix}1-L & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\tilde{\mathbf{M}}(L)\mathbf{u}_t.$$

Assuming that the matrix

$$\begin{pmatrix}d_{11}-d_{21} & d_{12}-d_{22}\\ 3 & 6\end{pmatrix}$$

is non-singular, the matrix $\tilde{\mathbf{M}}(L)$ is zeroless and has therefore a finite-degree left inverse. Proceeding as in Proposition 3, we obtain an alternative EC representation with just one error, namely $F_{1t}-F_{2t}$.

The example above can be generalized to show that generically $\mathbf{F}_t$ admits EC representations with a minimum d and a maximum $r-q+d$ of error terms. In particular, if $d=0$, in addition to an EC representation, $\mathbf{F}_t$ generically has a finite-degree autoregressive representation with no error terms, consistently with the fact that in this case $\mathbf{C}(L)$ is generically zeroless.

As we have seen at the end of Section 4, different autoregressive representations produce the same impulse-response functions. Therefore our choice of the maximum number of error terms has no consequence for our purposes.

C.2 The left inverse of $\mathbf{M}(L)$ is not unique

In this Section we show that the autoregressive polynomial is not unique. This is a standard problem when dealing with non singular vectors, which is also discussed in Forni and Lippi (2010), Forni et al. (2014).

Consider

$$(1-L)F_t=\begin{pmatrix}1+aL\\ 1+bL\end{pmatrix}u_t, \tag{23}$$

with $r=2$, $q=1$, $d=0$, $c=1$, with $a\neq b$. In this case $\mathbf{A}(L)$ is zeroless. An autoregressive representation can be obtained by elementary manipulations. Rewrite (23) as

$$\begin{aligned}(1-L)F_{1t}&=u_t+au_{t-1}\\ (1-L)F_{2t}&=u_t+bu_{t-1}\end{aligned} \tag{24}$$

Taking $(b\ -a)\mathbf{C}(L)\mathbf{u}_t$, we get

$$u_t = \frac{b(1-L)F_{1t} - a(1-L)F_{2t}}{b-a}.$$

This can be used to get rid of u_{t-1} in (24) and obtain

$$\left[\mathbf{I}_2 - \begin{pmatrix} \frac{ab}{b-a} & \frac{a^2}{b-a} \\ \frac{b^2}{b-a} & \frac{-ab}{b-a} \end{pmatrix} L\right](1-L)\mathbf{F}_t = \begin{pmatrix} 1 \\ 1 \end{pmatrix} u_t, \tag{25}$$

which is an autoregressive representation in first differences.

Model (24), slightly modified, can be used to illustrate non-uniqueness in the left inversion of $\mathbf{M}(L)$. Consider

$$\begin{aligned} (1-L)F_{1t} &= u_t + au_{t-1} \\ (1-L)F_{2t} &= u_t + bu_{t-1} \\ (1-L)F_{3t} &= u_t + cu_{t-1}. \end{aligned} \tag{26}$$

Taking any vector $\mathbf{h} = (h_1\ h_2\ h_3)$, orthogonal to $(a\ b\ c)$, we get rid of u_{t-1} in (26) and obtain an autoregressive representation in the differences. However, unlike in (24), here the vectors $\mathbf{h}$ span a 2-dimensional space, thus producing an infinite set of autoregressive representations.

In the example just above non-uniqueness can also be seen as the consequence of the fact that the three stochastic variables $F_{j,t-1}$, $j = 1, 2, 3$, are linearly dependent. Therefore, trying to project each of the F_{jt} on the space spanned by $F_{j,t-1}$, $j = 1, 2, 3$ one would find a non-invertible covariance matrix. We do not address this problem systematically in the present paper. In the empirical Section 6 we find that a "prudent" choice of the lag length keeps our estimates away from singular covariance matrices.

D Data Description and Data Treatment

No.	Series ID	Definition	Unit	F.	Source	SA	T
1	INDPRO	Industrial Production Index	2007=100	M	FED	1	2
2	IPBUSEQ	IP: Business Equipment	2007=100	M	FED	1	2
3	IPDCONGD	IP: Durable Consumer Goods	2007=100	M	FED	1	2
4	IPDMAT	IP: Durable Materials	2007=100	M	FED	1	2
5	IPNCONGD	IP: Nondurable Consumer Goods	2007=100	M	FED	1	2
6	IPNMAT	IP: nondurable Materials	2007=100	M	FED	1	2
7	CPIAUCSL	CPI: All Items	1982-84=100	M	BLS	1	3
8	CPIENGSL	CPI: Energy	1982-84=100	M	BLS	1	3
9	CPILEGSL	CPI: All Items Less Energy	1982-84=100	M	BLS	1	3
10	CPILFESL	CPI: All Items Less Food & Energy	1982-84=100	M	BLS	1	3
11	CPIUFDSL	CPI: Food	1982-84=100	M	BLS	1	3
12	CPIULFSL	CPI: All Items Less Food	1982-84=100	M	BLS	1	3
13	PPICRM	PPI: Crude Materials for Further Processing	1982=100	M	BLS	1	3
14	PPIENG	PPI: Fuels & Related Products & Power	1982=100	M	BLS	0	3
15	PPIFGS	PPI: Finished Goods	1982=100	M	BLS	1	3
16	PPIIDC	PPI: Industrial Commodities	1982=100	M	BLS	0	3
17	PPICPE	PPI: Finished Goods: Capital Equipment	1982=100	M	BLS	1	3
18	PPIACO	PPI: All Commodities	1982=100	M	BLS	0	3
19	PPIITM	PPI: Intermediate Materials	1982=100	M	BLS	1	3
20	AMBSL	St. Louis Adjusted Monetary Base	Bil. of $	M	StL	1	3
21	ADJRESSL	St. Louis Adjusted Reserves	Bil. of $	M	StL	1	3
22	CURRSL	Currency Component of M1	Bil. of $	M	FED	1	3
23	M1SL	M1 Money Stock	Bil. of $	M	FED	1	3
24	M2SL	M2 Money Stock	Bil. of $	M	FED	1	3
25	BUSLOANS	Commercial and Industrial Loans	Bil. of $	M	FED	1	2
26	CONSUMER	Consumer Loans	Bil. of $	M	FED	1	2
27	LOANINV	Bank Credit	Bil. of $	M	FED	1	2
28	LOANS	Loans and Leases in Bank Credit	Bil. of $	M	FED	1	2
29	REALLN	Real Estate Loans	Bil. of $	M	FED	1	2
30	TOTALSL	Tot. Cons. Credit Owned and Securitized	Bil. of $	M	FED	1	2
31	GDPC1	Gross Domestic Product	Bil. of Ch. 2005$	Q	BEA	1	2
32	FINSLC1	Final Sales of Domestic Product	Bil. of Ch. 2005$	Q	BEA	1	2
33	GPDIC1	Gross Private Domestic Investment	Bil. of Ch. 2005$	Q	BEA	1	2
34	SLCEC1	State & Local CE & GI	Bil. of Ch. 2005$	Q	BEA	1	2
35	PRFIC1	Private Residential Fixed Investment	Bil. of Ch. 2005$	Q	BEA	1	2
36	PNFIC1	Private Nonresidential Fixed Investment	Bil. of Ch. 2005$	Q	BEA	1	2
37	IMPGSC1	Imports of Goods & Services	Bil. of Ch. 2005$	Q	BEA	1	2
38	GCEC1	Government CE & GI	Bil. of Ch. 2005$	Q	BEA	1	2
39	EXPGSC1	Exports of Goods & Services	Bil. of Ch. 2005$	Q	BEA	1	2
40	CBIC1	Change in Private Inventories	Bil. of Ch. 2005$	Q	BEA	1	1
41	PCNDGC96	PCE: Nondurable Goods	Bil. of Ch. 2005$	Q	BEA	1	2
42	PCESVC96	PCE: Services	Bil. of Ch. 2005$	Q	BEA	1	2
43	PCDGCC96	PCE: Durable Goods	Bil. of Ch. 2005$	Q	BEA	1	2
44	PCECC96	Personal Consumption Expenditures	Bil. of Ch. 2005$	Q	BEA	1	2
45	DGIC96	National Defense Gross Investment	Bil. of Ch. 2005$	Q	BEA	1	2
46	NDGIC96	Federal Nondefense Gross Investment	Bil. of Ch. 2005$	Q	BEA	1	2
47	DPIC96	Disposable Personal Income	Bil. of Ch. 2005$	Q	BEA	1	2
48	PCECTPI	PPCE: Chain-type Price Index	2005=100	Q	BEA	1	3
49	GPDICTPI	GPDI: Chain-type Price Index	2005=100	Q	BEA	1	3
50	GDPCTPI	GDP: Chain-type Price Index	2005=100	Q	BEA	1	3
51	HOUSTMW	Housing Starts in Midwest	Thous. of Units	M	Census	1	2
52	HOUSTNE	Housing Starts in Northeast	Thous. of Units	M	Census	1	2
53	HOUSTS	Housing Starts in South	Thous. of Units	M	Census	1	2
54	HOUSTW	Housing Starts in West	Thous. of Units	M	Census	1	2
55	PERMIT	Building Permits	Thous. of Units	M	Census	1	2
56	ULCMFG	Manuf. S.: Unit Labor Cost	2005=100	Q	BLS	1	2
57	COMPRMS	Manuf. S.: Real Compensation Per Hour	2005=100	Q	BLS	1	2
58	COMPMS	Manuf. S.: Compensation Per Hour	2005=100	Q	BLS	1	2
59	HOAMS	Manuf. S.: Hours of All Persons	2005=100	Q	BLS	1	2
60	OPHMFG	Manuf. S.: Output Per Hour of All Persons	2005=100	Q	BLS	1	2
61	ULCBS	Business S.: Unit Labor Cost	2005=100	Q	BLS	1	2
62	RCPHBS	Business S.: Real Compensation Per Hour	2005=100	Q	BLS	1	2
63	HCOMPBS	Business S.: Compensation Per Hour	2005=100	Q	BLS	1	2
64	HOABS	Business S.: Hours of All Persons	2005=100	Q	BLS	1	2
65	OPHPBS	Business S.: Output Per Hour of All Persons	2005=100	Q	BLS	1	2

No.	Series ID	Definition	Unit	F.	Source	SA	T
66	MPRIME	Bank Prime Loan Rate	%	M	FED	0	1
67	FEDFUNDS	Effective Federal Funds Rate	%	M	FED	0	1
68	TB3MS	3-Month T.Bill: Secondary Market Rate	%	M	FED	0	1
69	GS1	1-Year Treasury Constant Maturity Rate	%	M	FED	0	1
70	GS3	3-Year Treasury Constant Maturity Rate	%	M	FED	0	1
71	GS10	10-Year Treasury Constant Maturity Rate	%	M	FED	0	1
72	EMRATIO	Civilian Employment-Population Ratio	%	M	BLS	1	1
73	CE16OV	Civilian Employment	Thous. of Persons	M	BLS	1	2
74	UNRATE	Civilian Unemployment Rate	%	M	BLS	1	1
75	UEMPLT5	Civilians Unemployed - Less Than 5 Weeks	Thous. of Persons	M	BLS	1	2
76	UEMP5TO14	Civilians Unemployed for 5-14 Weeks	Thous. of Persons	M	BLS	1	2
77	UEMP15T26	Civilians Unemployed for 15-26 Weeks	Thous. of Persons	M	BLS	1	2
78	UEMP27OV	Civilians Unemployed for 27 Weeks and Over	Thous. of Persons	M	BLS	1	2
79	UEMPMEAN	Average (Mean) Duration of Unemployment	Weeks	M	BLS	1	2
80	UNEMPLOY	Unemployed	Thous. of Persons	M	BLS	1	2
81	DMANEMP	All Employees: Durable goods	Thous. of Persons	M	BLS	1	2
82	NDMANEMP	All Employees: Nondurable goods	Thous. of Persons	M	BLS	1	2
83	SRVPRD	All Employees: Service-Providing Industries	Thous. of Persons	M	BLS	1	2
84	USCONS	All Employees: Construction	Thous. of Persons	M	BLS	1	2
85	USEHS	All Employees: Education & Health Services	Thous. of Persons	M	BLS	1	2
86	USFIRE	All Employees: Financial Activities	Thous. of Persons	M	BLS	1	2
87	USGOOD	All Employees: Goods-Producing Industries	Thous. of Persons	M	BLS	1	2
88	USGOVT	All Employees: Government	Thous. of Persons	M	BLS	1	2
89	USINFO	All Employees: Information Services	Thous. of Persons	M	BLS	1	2
90	USLAH	All Employees: Leisure & Hospitality	Thous. of Persons	M	BLS	1	2
91	USMINE	All Employees: Mining and logging	Thous. of Persons	M	BLS	1	2
92	USPBS	All Employees: Prof. & Business Services	Thous. of Persons	M	BLS	1	2
93	USPRIV	All Employees: Total Private Industries	Thous. of Persons	M	BLS	1	2
94	USSERV	All Employees: Other Services	Thous. of Persons	M	BLS	1	2
95	USTPU	All Employees: Trade, Trans. & Ut.	Thous. of Persons	M	BLS	1	2
96	USWTRADE	All Employees: Wholesale Trade	Thous. of Persons	M	BLS	1	2
97	OILPRICE	Spot Oil Price: West Texas Intermediate	$ per Barrel	M	DJ	0	3
98	NAPMNOI	ISM Manuf.: New Orders Index	Index	M	ISM	1	1
99	NAPMPI	ISM Manuf.: Production Index	Index	M	ISM	1	1
100	NAPMEI	ISM Manuf.: Employment Index	Index	M	ISM	1	1
101	NAPMSDI	ISM Manuf.: Supplier Deliveries Index	Index	M	ISM	1	1
102	NAPMII	ISM Manuf.: Inventories Index	Index	M	ISM	1	1
103	SP500	S&P 500 Stock Price Index	Index	D	S&P	0	2

Abbreviations

Source	Freq.	Trans.	SA
BLS=U.S. Department of Labor: Bureau of Labor Statistics	Q = Quarterly	1 = None	0 = no
BEA=U.S. Department of Commerce: Bureau of Economic Analysis	M = Monthly	2 = log	1 = yes
ISM = Institute for Supply Management	D = Daily	3 = Δ log	
Census=U.S. Department of Commerce: Census Bureau			
FED=Board of Governors of the Federal Reserve System			
StL=Federal Reserve Bank of St. Louis			

Note: All monthly and daily series are transformed into quarterly observation by simple averages