Open access • Journal Article • DOI:10.1198/016214504000002050

# The Generalized Dynamic Factor Model one-sided estimation and forecasting



Mario Forni, Marc Hallin, Marco Lippi, Lucrezia Reichlin

**Published on:** 01 Sep 2005 - Journal of the American Statistical Association (Taylor & Francis)

**Topics:** Dynamic factor, Covariance, Estimator, Factor analysis and Principal component analysis

EIEF WORKING PAPER SERIES

IEF

Einaudi Institute for Economics and Finance


**EIEF Working Paper 06/11**
**March 2011**


# One-Sided Representations of Generalized Dynamic Factor Models


by

Mario Forni

(University of Modena and Reggio Emilia)

Marc Hallin

(ECARES, Université Libre de Bruxelles)

Marco Lippi

(University of Rome "La Sapienza" and EIEF)

Paolo Zaffaroni

(Imperial College London)

# One-Sided Representations of Generalized Dynamic Factor Models *

Mario Forni
Università di Modena e Reggio Emilia, CEPR and RECent
Marc Hallin
ECARES, Université Libre de Bruxelles
Marco Lippi
Università di Roma La Sapienza and EIEF
Paolo Zaffaroni
Imperial College London†

March 23, 2011. **First draft. Please do not quote.**

**Abstract.** In the present paper we study a semiparametric version of the Generalized Dynamic Factor Model introduced in Forni, Hallin, Lippi and Reichlin (2000). Precisely, we suppose that the common components have rational spectral density, while no parametric structure is assumed for the idiosyncratic components. The parametric structure assumed for the common components does not imply that the model has a static representation (though the converse implication holds), a strong restriction which is shared by most of the literature on large-dimensional dynamic factor models. We use recent results on singular stationary processes with rational spectral density, to obtain a finite autoregressive representation for the common components. We construct an estimator for the model parameters and the common shocks. Consistency and rates of convergence are obtained. An empirical section, based on US macroeconomic time series, compares estimates based on our model with those based on the usual static-representation restriction. We find convincing evidence that the latter is not supported by the data.
JEL subject classification : C0, C01, E0.
Key words and phrases : Generalized dynamic factor models. Vector processes with singular spectral density. One-sided representations for dynamic factor models. Consistency and rates for estimators of dynamic factor models.

*P. Zaffaroni gratefully acknowledges financial support from ESRC Grant RES-000-22-3219.
†Also adjunct professor with the Università di Roma La Sapienza.

# 1 Introduction

Large-dimensional factor models have been studied in a vast literature starting with Chamberlain (1983), Chamberlain and Rothschild (1983), Forni et al. (2000), Forni and Lippi (2001), Stock and Watson (2002a,b), Bai and Ng (2002), Bai (2003).

Apart for minor features, the large-dimensional factor models we are referring to are specifications of the Generalized Dynamic Factor Model (GDFM) introduced in Forni, Hallin, Lippi and Reichlin (2000). The latter is a countably infinite set of observable stationary stochastic variables $x_{it}$ that can be represented in the following way:

$$x_{it} = \chi_{it} + \xi_{it} = b_{i1}(L)u_{1t} + b_{i2}(L)u_{2t} + \cdots + b_{iq}(L)u_{qt} + \xi_{it}, \tag{1.1}$$

where $i \in \mathbb{N}$, $t \in \mathbb{Z}$, $\mathbf{u}_t = (u_{1t}\ u_{2t}\ \cdots\ u_{qt})'$ is a $q$-dimensional orthonormal unobservable white-noise vector and $b_{if}(L)$ is a square-summable filter. The basic assumptions are:
(1) $\mathbf{u}_t$ is orthogonal to $\xi_{i,t-k}$ for all $j \in \mathbb{N}$ and $k \in \mathbb{Z}$.
(2) The covariance among the variables $\xi_{it}$ is "weak". This means that some covariance among the $\xi$'s is allowed, but all sequences of weighted averages $\sum_{i=1}^{n} w_{ni}\xi_{it}$, the weights fulfilling $\lim_{n\to\infty} \sum_{i=1}^{n} w_{ni}^2 = 0$, tend to zero in mean square (the sequence of the arithmetic averages $n^{-1}\sum_{i=1}^{n} \xi_{it}$ in particular). Note that $E(\xi_{it}^2) \leq M$ for all $i$, and $E(\xi_{it}\xi_{jt}) = 0$ for all $i \neq j$, is sufficient for weak covariance but not necessary (see Section 2 for a detailed presentation and discussion). Weak covariance among the variables $\xi_{it}$, usually called the *idiosyncratic components*, implies that the covariance among the obervable variables $x_{it}$ is mainly accounted for by the *common components* $\chi_{it}$, the latter being driven by the small-dimensional vector of *common shocks* $u_{ft}$, $f = 1, 2 \ldots, q$. Weak covariance among the $\xi$'s, instead of no correlation at all, is the reason for using "generalized" in its denomination. It is a major difference with respect to the dynamic factor models studied in Sargent and Sims (1977), Geweke (1977), Quah and Sargent (1993), which are based on a finite number $n$ of equations (1.1). In those models the definition of weak covariance, which requires $n$-asymptotics, does not make sense and orthogonality among the idiosyncratic components must be assumed.

Applications of large-dimensional factor models have been:[1] (a) Forecasting, see Stock and Watson (2002a,b), Forni, Hallin, Lippi and Reichlin (2005), (b) Business cycle analysis, Altissimo et al. (2010), (c) Structural analysis of the common components, Stock and Watson (2005), Forni, Giannone, Lippi and Reichlin (2009).

The main tool employed for estimation of the components $\chi_{it}$ and $\xi_{it}$, the common shocks $\mathbf{u}_t$ and the functions $b_{if}(L)$, has been the *principal components*

[1] Out of a vast literature, we mention here a few papers that (in their published or working-paper version) have been in our opinion "representative" of important research lines in this field.

(PC) of the variables $x_{it}$, both standard and in the frequency domain. The results obtained can be summarized as follows:

(i) Most of the literature assume that *for a given* $t$, the space $\overline{\text{span}}(\chi_{it}\ i \in \mathbb{N})$ is of finite dimension $r$, where $r \geq q$. As a consequence the model can be rewritten as

$$\begin{array}{rcl} x_{it} & = & \lambda_{i1}F_{1t} + \lambda_{i2}F_{2t} + \cdots + \lambda_{ir}F_{rt} + \xi_{it} \\ \mathbf{F}_t & = & N(L)\mathbf{u}_t. \end{array} \tag{1.2}$$

Criteria to determine consistently $r$ are given in Bai and Ng (2002). The vector $\mathbf{F}_t$ and the loadings $\lambda_{ij}$ can be estimated consistently using the first $r$ standard PC's, see Stock and Watson (2002a,b), Bai and Ng (2002). Moreover, the second equation in (1.2) is usually specified as a singular VAR, so that (1.2) becomes:

$$\begin{array}{rcl} x_{it} = \lambda_{i1}F_{1t} + \lambda_{i2}F_{2t} + \cdots + \lambda_{ir}F_{rt} & + & \xi_{it} \\ (I - D_1L - D_2L^2 - \cdots - D_pL^p)\mathbf{F}_t & = & \mathbf{K}\mathbf{u}_t, \end{array} \tag{1.3}$$

where the matrices $D_j$ are $r \times r$ while $\mathbf{K}$ is $r \times q$. Under (1.3), Bai and Ng (2007), Amengual and Watson (2007) provide consistent criteria to determine $q$. Estimation of the VAR in (1.3), and therefore of $\mathbf{u}_t$, up to an orthogonal matrix, is standard.

(ii) Using the frequency-domain principal components, and without assuming the finite-dimensional structure (1.2), Forni et al. (2000) obtain an estimator of the spectral density of the common components $\chi_{it}$ and of the common components themselves. Criteria to determine $q$ without assuming (1.2) or (1.3) are obtained in Hallin and Liška (2007), Onatski (2009). However, frequency-domain principal components produce two-sided filters, so that the resulting estimator of the $\chi$'s cannot be used at the end of the sample or for prediction.

Although some papers use the spectral density estimation referred to in (ii) (see e.g. Forni et al., 2005, Altissimo et al., 2010), a finite-dimensional structure like (1.2) or (1.3) is assumed with almost no exception. In the present paper we go back to the GDFM without assuming (1.2). A strong *a priori* motivation for our research is that representation (1.2) is so restrictive that even the elementary model

$$x_{it} = \frac{1}{1 - \alpha_i L} u_t + \xi_{it}, \tag{1.4}$$

where $q = 1$, $u_t$ is a scalar white noise, and the coefficients $\alpha_i$ are drawn from a uniform distribution between, say, $-0.8$ and $0.8$, is ruled out. For, the space spanned by the components $\chi_{it}$, for a given $t$ and $i \in \mathbb{N}$, is easily seen to be infinite dimensional unless $\alpha_i$ takes on a finite number of values.

To keep the model manageable, although not assuming the restriction in (1.2) or (1.3), we suppose that the common components have *rational spectral density*, i.e. that each function $b_{if}(L)$ in (1.1) is a ratio of polynomials in $L$. Under this assumption, the dimension of $\overline{\text{span}}(\chi_{it}\ i \in \mathbb{N})$ is infinite, like in (1.4), apart from values of the coefficients of the rational functions lying in negligible subsets (subsets that are, roughly speaking, lower dimensional; see Section 2 for a formal definition).

Assuming that the functions $b_{if}(L)$ are rational, we construct one-sided estimators for the common components $\chi_{it}$, $\mathbf{u}_t$ and the corresponding functions $b_{if}(L)$. Such estimators are then applied in an empirical investigation based on US quarterly macroeconomic data. We find that our method outperforms the standard PC estimator, which is based on assumption (1.2), both for the matrices $\mathbf{A}^k(L)$ and the common components and the common shocks $\mathbf{u}_t$. Thus assumption (1.2) is not supported by the US macroeconomic dataset we use. We believe that this provides good empirical motivation for the present research. Let us give a detailed description of the construction leading to our estimator.

(A) Population results. Our assumption that the common components have rational spectral density implies the following representation:

$$\begin{aligned} \chi_{it} &= \frac{c_{i1}(L)}{d_{i1}(L)}u_{1t} + \frac{c_{i2}(L)}{d_{i2}(L)}u_{2t} + \cdots + \frac{c_{iq}(L)}{d_{iq}(L)}u_{qt} \\ c_{if}(L) &= c_{if,0} + c_{if,1}L + \cdots + c_{if,s_1}L^{s_1} \\ d_{if}(L) &= 1 + d_{if,1}L + \cdots + d_{if,s_2}L^{s_2} \end{aligned} \tag{1.5}$$

for $i \in \mathbb{N}$, $f = 1, 2, \ldots, q$ (the degrees of the polynomials are independent of $i$, but this is a minor point). Regarding the idiosyncratic components we do not make any parametric assumption, nor do we restrict the covariance among them except for weakness, as mentioned above (for example, the covariance between $\xi_{it}$ and $\xi_{jt}$ might be non-zero for all $i \neq j$). Thus our model is semiparametric. We show that for generic values of the parameters $c_{if,k}$ and $d_{if,k}$ (i.e. apart from a subset that is negligible, in a sense to be specified in Section 2), the infinite-dimensional vector

$$\boldsymbol{\chi}_t = (\chi_{1t}\ \chi_{2t}\ \cdots\ \chi_{nt}\ \cdots) \tag{1.6}$$

has an autoregressive representation with the following block structure:

$$\begin{pmatrix} \mathbf{A}^1(L) & 0 & \cdots & 0 & \cdots \\ 0 & \mathbf{A}^2(L) & \cdots & 0 & \\ & & \ddots & & \\ 0 & 0 & \cdots & \mathbf{A}^k(L) & \\ \vdots & & & & \ddots \end{pmatrix} \boldsymbol{\chi}_t = \begin{pmatrix} \mathbf{R}^1 \\ \mathbf{R}^2 \\ \vdots \\ \mathbf{R}^k \\ \vdots \end{pmatrix} \mathbf{u}_t, \tag{1.7}$$

where $\mathbf{A}^k(L)$ is a $(q+1) \times (q+1)$, *finite-length* polynomial matrix and $\mathbf{R}^k$ is $(q+1) \times q$. Denoting by $\underline{\mathbf{A}}(L)$ and $\underline{\mathbf{R}}$ the matrices on the left and right hand sides of (1.7), and defining $\mathbf{x}_t$ and $\boldsymbol{\xi}_t$ in analogy with (1.6), we have:

$$\underline{\mathbf{A}}(L)\mathbf{x}_t = \underline{\mathbf{R}}\mathbf{u}_t + \underline{\mathbf{A}}(L)\boldsymbol{\xi}_t. \tag{1.8}$$

This is a factor model for $\underline{\mathbf{A}}(L)\mathbf{x}_t$, with a static representation, playing a crucial role in the estimation of $\mathbf{u}_t$. Some features of (1.7) must be pointed out:

(i) Because the infinite-dimensional vector $\boldsymbol{\chi}_t$ is driven by the $q$-dimensional white noise $\mathbf{u}_t$, for generic values of the parameters we can invert the infinite-dimensional moving average representation (1.5) "piecewise", by partitioning the vector $\boldsymbol{\chi}_t$ into the $(q+1)$-dimensional subvectors

$$(\chi_{1t}\ \chi_{2t}\ \cdots\ \chi_{q+1,t})\ \ (\chi_{q+2,t}\ \chi_{q+3,t}\ \cdots\ \chi_{2(q+1),t})\ \ \cdots$$

(ii) For generic values of the parameters, each of the subvectors, whose dimension and rank are $(q+1)$ and $q$ respectively, has a *finite* autoregressive representation. This is an application of a general result obtained in Anderson and Deistler (2008a and b) for rational-spectrum stochastic vectors that are singular (i.e. whose dimension is larger than the rank). We contribute to this literature showing that when the dimension is equal to $q+1$ the minimum-lag autoregressive representation is generically unique.

(iii) Under the assumption that the length of the VAR matrices $\mathbf{A}^k(L)$ is bounded, the number of VAR coefficients grow at pace $n$. Each matrix $\mathbf{A}^k(L)$ is estimated independently of the previous ones.

(B) Estimation. The spectral density of the common components can be consistently estimated by using the first $q$ frequency-domain principal components (see Forni et al., 2000). Using such spectral density we obtain a consistent estimator of the autocovariance functions of the common components. The autocovariance functions are then used to estimate the matrices $\mathbf{A}^k(L)$ and $\mathbf{R}^k$. Lastly, once the matrices $\mathbf{A}^k(L)$ have been estimated we use equation (1.8) to estimate $\mathbf{u}_t$. As we have observed above, (1.8) has a static representation, so that standard principal components are the appropriate tool. However, the matrices $\mathbf{A}^k(L)$ must be replaced by their estimates, this implying considerable complications to prove consistency. For the coefficients of the matrices $\mathbf{A}^k(L)$, the entries of the matrices $\mathbf{R}^k$ and $\mathbf{u}_t$ we obtain consistency, for $n \to \infty$ and $T \to \infty$, at rate

$$\max(n^{-1/2}, \rho_T^{-1/2}),$$

where $\rho_T$ is any sequence diverging slower than $T^{2/3}/\log T$, this being the toll to be paid for using non-parametric spectral estimation. However, our empirical exercise provides some evidence that the general dynamic shape of our model can offset a lower speed of consistency, relative to the rate $T^{1/2}$, which can be obtained with model (1.3).

The body of the paper contains detailed discussion and motivation of the main assumptions. Proofs that go beyond a few lines are collected in the Appendix. The population and estimation results are presented in Section 2 and Section 3 respectively. The empirical results are presented and discussed in Section 4. Section 5 concludes.

# 2 Population assumptions and results

## 2.1 Notation

The GDFM studied in the present paper, see equation (1.1), can be alternatively thought of as (1) a double-indexed family of stochastic variables $\{x_{it},\ i \in \mathbb{N},\ t \in \mathbb{Z}\}$, (2) a family of stationary processses $\{x_{it},\ t \in \mathbb{Z}\}$ indexed by $i \in \mathbb{N}$, (3) a family of cross sections $\{x_{it},\ i \in \mathbb{N}\}$ indexed by $t \in \mathbb{Z}$, i.e. a process of infinite-dimensional stochastic vectors. We find the third option convenient and use the notation

$$\mathbf{x}_t = (x_{1t}\ x_{2t}\ \cdots\ x_{nt}\ \cdots)'.$$

Analogously for $\boldsymbol{\chi}_t$ and $\boldsymbol{\xi}_t$. Writing $\mathbf{x}_t = \boldsymbol{\chi}_t + \boldsymbol{\xi}_t$ has the obvious entry-by-entry meaning.

We also introduce infinite-dimensional matrices like $\underline{\mathbf{A}}(L)$, which is $\infty \times \infty$, or

$$\underline{\mathbf{b}}(L) = \begin{pmatrix} b_{11}(L) & b_{12}(L) & \cdots & b_{1q}(L) \\ b_{21}(L) & b_{22}(L) & \cdots & b_{2q}(L) \\ \vdots & \vdots & & \vdots \\ b_{n1}(L) & b_{n2}(L) & \cdots & b_{nq}(L) \\ \vdots & \vdots & & \vdots \end{pmatrix},$$

which is $\infty \times q$. We use such matrices in expressions like (1.8) or in $\mathbf{x}_t = \underline{\mathbf{b}}(L)\mathbf{u}_t + \boldsymbol{\xi}_t$. The reader will easily check that we never produce infinite sums of products, so that our infinite-dimensional matrices are no more than a notational convenience. All infinite-dimensional matrices are underlined, while their finite-dimensional submatrices are not. In particular, $\mathbf{A}_s(L)$ denotes the $s \times s$ upper left submatrix of $\underline{\mathbf{A}}(L)$, $\mathbf{b}_s(L)$ and $\mathbf{R}_s$ denote the $s \times q$ upper submatrices of $\underline{\mathbf{b}}(L)$ and $\underline{\mathbf{R}}$ respectively.

In Section 3 explicit reference to $s$ in $\mathbf{A}_s(L)$, $\mathbf{b}_s(L)$, $\mathbf{R}_s$, etc., is no longer necessary. Thus we switch to a fairly different and more convenient notation.

Given the infinite-dimensional process $\mathbf{y}_t = (y_{1t}\ y_{2t}\ \cdots\ y_{nt}\ \cdots)'$, we use the following notation:

(1) $\mathbf{y}_{st}$ is the $n$-dimensional process $(y_{1t}\ y_{2t}\ \cdots\ y_{st})'$.

(2) $\mathcal{H}^y = \overline{\text{span}}(y_{it},\ i \in \mathbb{N},\ t \in \mathbb{Z})$, $\quad \mathcal{H}^{y_s} = \overline{\text{span}}(y_{it},\ i \leq s,\ t \in \mathbb{Z})$.

(3) $\mathcal{H}_t^y = \overline{\text{span}}(y_{i\tau},\ i \in \mathbb{N},\ \tau \leq t)$, $\mathcal{H}_t^{y_s} = \overline{\text{span}}(y_{i\tau},\ i \leq s,\ \tau \leq t)$.

(4) Analogous notation applies if $\mathbf{y}_t$ is $n$-dimensional. For $s \leq n$, $\mathcal{H}^{y_s} = \overline{\text{span}}(y_{it},\ i \leq s,\ t \in \mathbb{Z})$, $\mathcal{H}_t^{y_s} = \overline{\text{span}}(y_{j\tau},\ j \leq s,\ \tau \leq t)$.

## 2.2 Basic assumptions

All the stochastic variables $x_{it}$, $\chi_{it}$ and $\xi_{it}$ introduced below are zero mean.

**Assumption A.1** For all $n \in \mathbb{N}$ the vector $\mathbf{x}_{nt}$ is weakly stationary and has a spectral density (an absolutely continuous spectral measure).

We denote by $\mathbf{\Sigma}_n^x(\theta)$, $\theta \in [-\pi\ \pi]$, the nested spectral density matrices of the vectors $\mathbf{x}_{nt} = (x_{1t}\ x_{2t}\ \cdots\ x_{nt})'$. The matrix $\mathbf{\Sigma}_n^x(\theta)$ is Hermitian, non-negative definite and has therefore real and non-negative eigenvalues for all $\theta \in [-\pi\ \pi]$. The $j$-th eigenvalue, in decreasing order, of $\mathbf{\Sigma}_n^x(\theta)$ is denoted by $\lambda_{nj}^x(\theta)$. Our second assumption is

**Assumption A.2** Define $\bar{\lambda}_f^x(\theta) = \sup_{n\in\mathbb{N}} \lambda_{nf}^x(\theta)$. There exists a positive integer $q$ such that: (i) $\bar{\lambda}_q^x(\theta)) = \infty$ for almost all $\theta$ in $[-\pi\ \pi]$, (ii) $\bar{\lambda}_{q+1}^x(\theta)$ is essentially bounded.

Forni and Lippi (2001) prove that

**Theorem A** Assumptions A.1 and A.2 imply that $\mathbf{x}_t$ can be represented as in (1.1), i.e.

$$\mathbf{x}_t = \boldsymbol{\chi}_t + \boldsymbol{\xi}_t = \underline{\mathbf{b}}(L)\mathbf{u}_t + \boldsymbol{\xi}_t, \tag{2.1}$$

where $\underline{\mathbf{b}}(L)$ is an $\infty \times q$ matrix of square summable filters, $\mathbf{u}_t$ is a $q$-dimensional orthonormal white noise, and
(i) $\boldsymbol{\xi}_{nt}$ fulfills Assumption A.1 and $\bar{\lambda}_1^\xi(\theta)$ (defined in the same way as $\bar{\lambda}_1^x(\theta)$) is essentially bounded.
(ii) Obviously $\boldsymbol{\chi}_t$ fulfills A.1. Moreover, $\bar{\lambda}_q^\chi(\theta) = \infty$ for almost all $\theta$ in $[\pi\ \pi]$ (note that $\bar{\lambda}_{q+1}^\chi(\theta) = 0$ a.e. in $[\pi\ \pi]$).
(iii) $\boldsymbol{\xi}_t \perp \mathbf{u}_{t-k}$ for all $k \in \mathbb{Z}$.
(iv) The integer $q$ and the components $\chi_{it}$ and $\xi_{it}$ are unique.
Viceversa, if $\mathbf{x}_t$ can be represented as in (2.1) with $\boldsymbol{\chi}_t$ and $\boldsymbol{\xi}_t$ fulfilling (i), (ii) and (iii), then $\mathbf{x}_t$ fulfills Assumptions A.1 and A.2.

An infinite-dimensional vector fulfilling (i) is called an idiosyncratic vector. Divergence of the first $q$ eigenvectors of $\boldsymbol{\chi}_{nt}$ ensures that a representation of $\boldsymbol{\chi}_t$ as a moving average of a lower-dimensional white noise is not possible.

## 2.3 Infinite-dimensional finite-rank processes

Of course uniqueness of $\boldsymbol{\chi}_t$ and $\boldsymbol{\xi}_t$ in (2.1) does not imply that $\mathbf{u}_t$ and $\underline{\mathbf{b}}(L)$ are unique. Alternative representations are $\boldsymbol{\chi}_t = [\underline{\mathbf{b}}(L)\mathbf{B}][\mathbf{B}'\mathbf{u}_t] = \underline{\mathbf{c}}(L)\mathbf{v}_t$, where $\mathbf{B}$ is a $q \times q$ orthogonal matrix, or, more in general, $\boldsymbol{\chi}_t = [\underline{\mathbf{b}}(L)\mathbf{C}(L)][(\mathbf{C}'(F)\mathbf{u}_t] = \underline{\mathbf{d}}(L)\mathbf{w}_t$, where $F = L^{-1}$ and $\mathbf{C}(L)\mathbf{C}'(F) = \mathbf{I}_q$ for $\theta$ for almost all $\theta$ in $[-\pi\ \pi]$.

Moreover, Theorem A does not ensure that $\boldsymbol{\chi}_t$ has a one-sided representation, i.e. that there exists a representation $\boldsymbol{\chi}_t = \underline{\mathbf{e}}(L)\mathbf{z}_t$ such that $\underline{\mathbf{e}}(L) = \underline{\mathbf{e}}_0 + \underline{\mathbf{e}}_1 L + \cdots$. For example, if $q = 1$ and

$$\chi_{it} = u_{t+i-1}, \tag{2.2}$$

for $i \in \mathbb{N}$, where $u_t$ is a 1-dimensional white noise, then statement (ii) of Theorem A is fulfilled, so that $\boldsymbol{\chi}_t$ is the common component of some process $\mathbf{x}_t$ fulfilling A.1 and A.2, but $\boldsymbol{\chi}_t$ has no one-sided representations (this is quite obvious, see Lemma 1).[2]

Simple conditions for the existence of one-sided representations of infinite-dimensional stochastic vectors are given in Lemmas 1 and 2.

**Definition 1** Consider the infinite dimensional process

$$\mathbf{y}_t = (y_{1t}\ y_{2t}\ \cdots\ y_{nt}\ \cdots)'.$$

Assume that $\mathbf{y}_t$ fulfills Assumption A.1. We say that $\mathbf{y}_t$ has rank $q$ if there exists $s$ such that $\text{rank}(\Sigma_n^y(\theta)) = q$, for $n \geq s$ and almost all $\theta$ in $[-\pi\ \pi]$.

**Definition 2** Let $\mathbf{y}_t$ denote an infinite-dimensional stationary stochastic vector, which has a moving average representation

$$\mathbf{y}_t = \underline{\mathbf{b}}(L)\mathbf{v}_t, \tag{2.3}$$

where $\mathbf{v}_t$ is a $q$-dimensional orthonormal white noise and $\underline{\mathbf{b}}(L)$ is an $\infty \times q$ square summable filter. We say that (2.3) is a *fundamental* representation if: (1) $\underline{\mathbf{b}}(L)$ is one-sided. (2) $\mathbf{v}_t$ belongs to $\mathcal{H}_t^y$. In that case we also say that the white noise $\mathbf{v}_t$ is fundamental for $\mathbf{y}_t$. Note that if $\mathbf{v}_t$ is fundamental for $\mathbf{y}_t$ then $\mathcal{H}_t^{v_q} = \mathcal{H}_t^y$. The same definition can be given, *mutatis mutandis,* when $\mathbf{y}_t$ is $n$-dimensional. In that case $q \leq n$.

Suppose now that $\mathbf{y}_t$ is $n$-dimensional. The following properties hold:

(A) If (2.3) is fundamental and $\mathbf{y}_t = \mathbf{c}(L)\mathbf{w}_t$ is another fundamental representation, with $\mathbf{w}_t$ orthonormal, then $\mathbf{w}_t$ has dimension $q$, $\mathbf{c}(L) = \mathbf{b}(L)\mathbf{Q}$ and $\mathbf{w}_t = \mathbf{Q}'\mathbf{v}_t$, where $\mathbf{Q}$ is a $q \times q$ orthogonal matrix (Rozanov, 1967, pp. 56-57).

[2] The possibility that $\boldsymbol{\chi}_t$ has no one-sided representations arises here from infinite dimension. This bears no relationship with non-existence of one-sided representations for finite-dimensional processes, which occurs if their spectral density is singular in a positive-measure subset of $[-\pi\ \pi]$, see e.g. Pourahmadi (2001), Theorem 10.5, p. 361.

(B) If (2.3) is fundamental then $\text{rank}(\mathbf{b}(z)) = q$ for all complex numbers $z$ such that $|z| < 1$ (Rozanov, 1967, p. 63, Remark 3). In particular, $\text{rank}(b_0) = \text{rank}(\mathbf{b}(0)) = q$.

A finite-dimensional stationary process with a spectral density does not necessarily possess a fundamental representation. For example, if the spectral density of $\mathbf{y}_t$ is singular on a positive-measure subset of $[-\pi\ \pi]$, then $\mathbf{y}_t$ has no fundamental representations (indeed, it has no one-sided representations, see footnote 2). However:

(C) If $\mathbf{y}_t$ has rational spectral density then it has fundamental representations. If $\mathbf{y}_t = \mathbf{b}(L)\mathbf{v}_t$ is one of them, $\mathbf{v}_t$ being a $q$ dimensional orthonormal white noise, then the entries of $\mathbf{b}(L)$ are rational functions of $L$ (Rozanov, 1967, Chapter I, Section 10, Hannan, 1970, pp. 62-67).

(B′) Suppose that $\mathbf{y}_t$ has rational spectral density, that $\mathbf{y}_t = \mathbf{b}(L)\mathbf{v}_t$, where $\mathbf{b}(L)$ is $n \times q$, rational, square summable and one-sided, $\mathbf{v}_t$ is a $q$-dimensional orthonormal white noise, and that $\text{rank}(\mathbf{b}(z)) = q$ for all $z$ such that $|z| < 1$. Then $\mathbf{y}_t = \mathbf{b}(L)\mathbf{v}_t$ is fundamental (Hannan, 1970, pp. 62-67).

We say that the infinite-dimensional process $\mathbf{y}_t$ has rational spectral density if $\mathbf{y}_{nt}$ has rational spectral density for all $n$.

**Lemma 1** Suppose that the infinite-dimensional process $\mathbf{y}_t$ fulfills A.1, has rational spectral density and rank $q$. The following statements are equivalent:
(i) $\mathbf{y}_t$ has a one-sided rational moving average representation

$$\mathbf{y}_t = \underline{\mathbf{b}}(L)\mathbf{v}_t, \tag{2.4}$$

(i.e. the entries of $\underline{\mathbf{b}}(L)$ are rational functions of $L$), where $\mathbf{v}_t$ is a $q$-dimensional orthonormal white noise.
(ii) There exists a positive integer $s$ such that $\mathcal{H}_t^{y_s} = \mathcal{H}_t^{y}$.

**Proof.** Assume (ii) and let

$$\mathbf{y}_{st} = \mathbf{b}_s(L)\mathbf{v}_t$$

be rational, one-sided and fundamental, so that $\mathcal{H}_t^{y_s} = \mathcal{H}_t^{v_q}$. By assumption $y_{s+k,t} \in \mathcal{H}_t^{y_s}$ and therefore $y_{s+k,t} \in \mathcal{H}_t^{v_q}$, so that

$$\begin{aligned} \mathbf{y}_{st} &= \mathbf{b}_s(L)\mathbf{v}_t \\ y_{s+k,t} &= b_{s+k}(L)\mathbf{v}_t. \end{aligned} \tag{2.5}$$

The white noise $\mathbf{v}_t$ is fundamental for $\mathbf{y}_{st}$ and is therefore fundamental for $(\mathbf{y}_{st}\ y_{s+k,t})$. Thus representation (2.5) is fundamental, so that, by (C), $b_{s+k}(L)$ must be rational. The conclusion follows. Assume now that (i) holds. We say that $\beta$ is a zero of $\underline{\mathbf{b}}(L)$ if the determinant of all the $q \times q$ submatrices of $\underline{\mathbf{b}}(\beta)$ vanish. Assume that $\alpha$ is a zero of $\underline{\mathbf{b}}(L)$ and that $|\alpha| < 1$. There exists an orthogonal $q \times q$ matrix $\mathbf{B}_\alpha$ such that all the entries of the first column of

$$\underline{\mathbf{b}}(L)\mathbf{B}_\alpha$$

vanish at $\alpha$. Defining $\boldsymbol{\gamma}_\alpha(L)$ as the $q \times q$ matrix with

$$\left(\frac{1-\alpha L}{L-\alpha} \quad 1 \ \cdots \ 1\right)$$

on the main diagonal and zero elsewhere, we have

$$\mathbf{y}_t = [\underline{\mathbf{b}}(L)\mathbf{B}_\alpha \boldsymbol{\gamma}_\alpha(L)] \left[\boldsymbol{\gamma}_{\tilde{\alpha}}(L^{-1})\tilde{\mathbf{B}}_\alpha \mathbf{v}_t\right] = \underline{\mathbf{c}}(L)\mathbf{w}_t,$$

where a tilde denotes transposition and conjugation. This is an alternative one-sided rational representation in which the multiplicity of $\alpha$ as a zero of the matrix polynomial has decreased by one. Because a zero of $\underline{\mathbf{b}}(L)$ is a zero of $\mathbf{b}_q(L)$, with a finite number of iterations we obtain a rational representation, say $\mathbf{y}_t = \underline{\mathbf{d}}(L)\mathbf{z}_t$, such that $\underline{\mathbf{d}}(L)$ has no zeros of modulus less than unity. For the same reason there exists an integer $s$ such that $\mathbf{d}_s(L)$ has no zeros of modulus less than unity. By (B$'$),

$$\mathbf{y}_{st} = \mathbf{d}_s(L)\mathbf{z}_t$$

is fundamental for $\mathbf{y}_{st}$ and therefore for $\mathbf{y}_t$. Q.E.D.

**Lemma 2** Suppose that the infinite-dimensional process $\mathbf{y}_t$ fulfills A.1, has rational spectral density and rank $q$. Then:
(i) $\mathbf{y}_t$ has a fundamental rational representation $\mathbf{y}_t = \underline{\mathbf{b}}(L)\mathbf{v}_t$ if and only if it has a one-sided representation.
(ii) If $\mathbf{y}_t = \underline{\mathbf{b}}(L)\mathbf{v}_t$ and $\mathbf{y}_t = \underline{\mathbf{c}}(L)\mathbf{w}_t$ are fundamental, with $\mathbf{v}_t$ and $\mathbf{w}_t$ $q$-dimensional and orthonormal, then $\underline{\mathbf{c}}(L) = \underline{\mathbf{b}}(L)\mathbf{Q}$ and $\mathbf{w}_t = \mathbf{Q}'\mathbf{v}_t$, where $\mathbf{Q}$ is a $q \times q$ orthogonal matrix.
(iii) If

$$\mathbf{y}_t = \underline{\mathbf{b}}(L)\mathbf{v}_t = \underline{\mathbf{b}}_0\mathbf{v}_t + \underline{\mathbf{b}}_1\mathbf{v}_{t-1} + \cdots$$

is fundamental, then $\underline{\mathbf{b}}_0$ has rank $q$.

Proof. Statement (i) is part of the proof of Lemma 1. As regards (ii), suppose that $\mathbf{y}_t = \underline{\mathbf{b}}(L)\mathbf{v}_t$ and $\mathbf{y}_t = \underline{\mathbf{c}}(L)\mathbf{w}_t$ are both fundamental. By Lemma 1 there exists $s$ such that $\mathcal{H}_t^{y_s} = H_t^y$. As a consequence both $\mathbf{v}_t$ and $\mathbf{w}_t$ belong to $H_t^{y_s}$ and therefore are fundamental for $\mathbf{y}_{st}$. This implies that $\mathbf{w}_t = \mathbf{Q}'\mathbf{v}_t$, where $\mathbf{Q}$ is orthogonal. Thus $\mathbf{y}_t = \underline{\mathbf{c}}(L)\mathbf{w}_t = [\underline{\mathbf{c}}(L)\mathbf{Q}']\mathbf{v}_t = \underline{\mathbf{b}}(L)\mathbf{v}_t$. As $\mathbf{v}_t$ is an orthonormal white noise, we have $\underline{\mathbf{c}}(L) = \underline{\mathbf{b}}(L)\mathbf{Q}$. Because $\mathbf{v}_t$ is fundamental for $\mathbf{y}_{st}$, $\mathbf{b}_s(0)$ has rank $q$, see (B), so that $\underline{\mathbf{b}}(0) = \underline{\mathbf{b}}_0$ has rank $q$. Q.E.D.

In conclusion, given the infinite-dimensional vector $\mathbf{y}_t$, assuming A.1, finite rank, rational spectral density, and the existence of a one-sided moving average representation, thus ruling out cases like (2.2), we obtain the existence of a rational fundamental representation for $\mathbf{y}_t$, which is unique up to an orthogonal matrix.

Let us now return to the infinite-dimensional vector $\boldsymbol{\chi}_t$. As we have seen, $\boldsymbol{\chi}_t$ fulfills A.1. Assume that $\boldsymbol{\chi}_t$ has rational spectral density, so that either $\text{rank}(\boldsymbol{\Sigma}_n^{\chi}(\theta)) < q$ for all $\theta \in [-\pi\ \pi]$ or $\text{rank}(\boldsymbol{\Sigma}_n^{\chi}(\theta)) = q$ for almost all $\theta$ in $[-\pi\ \pi]$. On the other hand, since $\lambda_{nq}^{\chi}(\theta)$ diverges almost everywhere in $[-\pi\ \pi]$, there exists $s$ such that $\text{rank}(\boldsymbol{\Sigma}_n^{\chi}(\theta)) = q$ for $n \geq s$ and almost all $\theta$ in $[-\pi\ \pi]$. Therefore $\boldsymbol{\chi}_t$ has rank $q$.

Assuming that $\boldsymbol{\chi}_t$ has rational spectral density and that $\mathcal{H}_t^{\chi_s} = \mathcal{H}_t^{\chi}$ for some $s$, so that cases like (2.2) cannot occur, Lemma 2 ensures that $\boldsymbol{\chi}_t$ has a rational fundamental representation. Precisely, for $i \in \mathbb{N}$,

$$\chi_{it} = \frac{c_{i1}(L)}{d_{i1}(L)}u_{1t} + \frac{c_{i2}(L)}{d_{i2}(L)}u_{2t} + \cdots + \frac{c_{iq}(L)}{d_{iq}(L)}u_{qt}, \tag{2.6}$$

where $c_{if}(L)$ and $d_{if}(L)$ are polynomials in $L$ and $\mathbf{u}_t$ is fundamental for $\boldsymbol{\chi}_t$. Representation (2.6) is unique up to an orthogonal matrix.

However, in Assumption A.3, formalized in Section 2.5, we require more than the existence of an integer $s$ such that $\mathcal{H}_t^{\chi_s} = \mathcal{H}_t^{\chi}$. Precisely, we assume that the space spanned by

$$\chi_{i_1\tau},\ \chi_{i_2\tau},\ \ldots,\ \chi_{i_{q+1},\tau},$$

for $\tau \leq t$, be equal to $\mathcal{H}_t^{\chi}$ for all $(q+1)$-tuples $i_1 < i_2 < \cdots < i_{q+1}$. Thus in (2.6) $\mathbf{u}_t$ is fundamental for any $(q+1)$-dimensional subvector of $\boldsymbol{\chi}_t$, not only for $\boldsymbol{\chi}_{st}$, for some $s$. This stronger requirement is motivated by the main result of Section 2.4. We prove that, under a quite general parameterization, the stronger condition holds generically, i.e. outside of a negligible subset, as defined in Section 2.4, of the parameter space.

## 2.4 AR representations for singular stochastic vectors

Consider a rational $n$-dimensional vector $\mathbf{y}_t$ such that

$$\begin{aligned} y_{it} &= \frac{c_{i1}(L)}{d_{i1}(L)}v_{1t} + \frac{c_{i2}(L)}{d_{i2}(L)}v_{2t} + \cdots + \frac{c_{iq}(L)}{d_{iq}(L)}v_{qt} \\ c_{if}(L) &= c_{if,0} + c_{if,1}L + \cdots + c_{if,s_1}L^{s_1} \\ d_{if}(L) &= 1 + d_{if,1}L + \cdots + d_{if,s_2}L^{s_2} \end{aligned} \tag{2.7}$$

for $i = 1, 2, \ldots, n$, $f = 1, 2, \ldots, q$. Each of the equations (2.7) is parameterized in $\mathbb{R}^{\nu}$, with $\nu = q(s_1 + s_2 + 1)$. More precisely, in the open set $\Pi \subset \mathbb{R}^{\nu}$ such that all the roots of the polynomials $d_{if}(L)$ are of modulus greater than unity. Thus the vector $\mathbf{y}_t$ is parameterized in

$$\Pi^n = \overbrace{\Pi \times \Pi \times \cdots \times \Pi}^{n},$$

which is an open subset of $\mathbb{R}^\mu$, with $\mu = n\nu$.

We are interested in the case $n > q$. Such "tall systems" have been recently studied in Anderson and Deistler (2008a and b). One of their results is that if $n > q$ then $\mathbf{y}_t$ has a *finite* autoregressive representation

$$\mathbf{A}(L)\mathbf{y}_t = \mathbf{R}\mathbf{v}_t,$$

where $\mathbf{R}$ is $n \times q$, $\mathbf{A}(L)$ is $n \times n$, for *generic* values of the parameters. Precisely, there exists a nowhere dense set $\mathcal{N} \subset \Pi^n$, i.e. a set whose closure has no interior points, such that an autoregressive representation exists for all parameter vectors in $\Pi^n - \mathcal{N}$. Of course this has the consequence that $\mathbf{v}_t$ is generically fundamental for $\mathbf{y}_t$.

To provide an intuition for this result and Proposition 1 below, let us consider the following elementary example, in which $n = 2$, $q = 1$:

$$\begin{aligned} y_{1t} &= a_1 v_t + b_1 v_{t-1} \\ y_{2t} &= a_2 v_t + b_2 v_{t-1}, \end{aligned} \tag{2.8}$$

which is parametrized in $\mathbb{R}^2 \times \mathbb{R}^2$. Outside of the nowhere dense subset in which $a_1 b_2 - a_2 b_1 = 0$, the system

$$\begin{aligned} b_1\beta_1 + b_2\beta_2 = 0 \\ a_1\beta_1 + a_2\beta_2 \neq 0 \end{aligned}$$

has the one-dimensional set of solutions $\mu(b_2 \quad - b_1)$. Thus taking the linear combination of $y_{1t}$ and $y_{2t}$ with coefficients $b_2$ and $-b_1$ we obtain

$$v_t = \frac{1}{a_1 b_2 - a_2 b_1}(b_2 y_{1t} - b_1 y_{2t}). \tag{2.9}$$

Using (2.9) to get rid of $u_{t-1}$ in (2.8), we obtain the finite autoregressive representation

$$\begin{aligned} y_{1t} &= d b_1 b_2 y_{1t-1} - d b_1^2 y_{2t-1} + a_1 v_t \\ y_{2t} &= d b_2^2 y_{1t-1} - d b_1 b_2 y_{2t-1} + a_2 v_t \end{aligned} \tag{2.10}$$

where $d = 1/(a_1 b_2 - a_2 b_1)$. Note that:

(i) If $a_1 b_2 - a_2 b_1 = 0$ no finite autoregressive representation exists. Moreover, fundamentalness of $v_t$ for $\mathbf{y}_t$ requires that the root of $a_1 + b_1 L$ (which is identical to the root of $a_2 + b_2 L$) be greater than unity in modulus.

(ii) But as soon as $a_1 b_2 - a_2 b_1 \neq 0$, the position of the root of $a_1 + b_1 L$ does not play any role for the fundamentalness of $v_t$ for $\mathbf{y}_t$.

(iii) Quite obviously $a_1 b_2 - a_2 b_1 \neq 0$ if and only if $\chi_{1t-1}$ and $\chi_{2t-1}$ are linearly independent. Therefore, generically, the projection (2.10) is unique, i.e. generically no other autoregressive representation of order one exists.

(iv) But other autoregressive representations do exist. Rewriting (2.10) as $\mathbf{y}_t = \mathbf{A}\mathbf{y}_{t-1} + \mathbf{a}v_t$ (the definition of $\mathbf{A}$ and $\mathbf{a}$ is obvious), we get $\mathbf{y}_t = \mathbf{A}^2\mathbf{y}_{t-2} + \mathbf{A}\mathbf{a}v_{t-1} + \mathbf{a}v_t$. Using (2.9) to get rid of $v_{t-1}$, we obtain another autoregressive representation, whose order is 2. Such non-uniqueness does not occur for square systems, i.e. when $n = q$.

(v) Moreover, if $n = 3$ and $y_{it} = a_i v_t + b_i v_{t-1}$, $i = 1, 2, 3$, then outside of the set in which $a_2 b_1 = a_1 b_2$ and $a_3 b_1 = a_1 b_3$, which is nowhere dense in $\mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}^2$, the system

$$\begin{aligned} b_1\beta_1 + b_2\beta_2 + b_3\beta_3 &= 0 \\ a_1\beta_1 + a_2\beta_2 + a_3\beta_3 &\neq 0 \end{aligned}$$

has a two-dimensional set of solutions. If $(\gamma_1\ \gamma_2\ \gamma_3)$ is one of them, we obtain

$$v_t = \frac{1}{a_1\gamma_1 + a_2\gamma_2 + a_3\gamma_3}(\gamma_1 y_{1t} + \gamma_2 y_{2t} + \gamma_3 y_{3t}),$$

which can be used to get rid of $v_{t-1}$, in the same way as we did in the $n = 2$ case. Thus generically $\mathbf{y}_t$ has an autoregressive representation of order one. However, the variables $y_{it-1}$, $i = 1, 2, 3$, are not linearly independent, so that such minimum-lag autoregressive representation is not unique (equivalently, unlike the $n = 2$ case, here the $\beta$'s can be chosen in a 2-dimensional set).

We show that observation (iii) generalizes to system (2.7). Precisely, if $n = q + 1$ then generically there exists only one minimum-lag autoregressive representation.

**Proposition 1** Consider an $n$-dimensional vector $\mathbf{y}_t$ with representation (2.7) and assume that $n = q + 1$. There exists a set $\mathcal{N} \subset \Pi^{q+1}$, nowhere dense in $\Pi^{q+1}$ such that if the parameter vector lies in $\Pi^{q+1} - \mathcal{N}$, then:
(a) $\mathbf{y}_t$ has a finite AR representation

$$\mathbf{A}(L)\mathbf{y}_t = \mathbf{R}\mathbf{v}_t,$$

where $\mathbf{R}$ is $(q + 1) \times q$, $\mathbf{A}(L)$ is $(q + 1) \times (q + 1)$ and has order not exceeding $S = qs_1 + q^2 s_2$. Therefore $\mathbf{v}_t$ is fundamental for $\mathbf{y}_t$.
(b) If $\mathbf{B}(L)\mathbf{y}_t = \mathbf{S}\mathbf{w}_t$ and $\mathbf{B}(L) \neq \mathbf{A}(L)$ then $\mathbf{B}(L)$ has order greater than $S$.
(c) $R_{ij} = c_{ij}(0)$. (Proof in Appendix A.)

Note that the proposition is not claiming that generically the vector $\mathbf{y}_t$ corresponding to a point in $\Pi^{q+1}$ has no non-fundamental representations. What it claims is that generically such non-fundamental representations are not parameterized in $\Pi^{q+1}$. For example, representation (2.8) is generically fundamental in

$\mathbb{R}^2 \times \mathbb{R}^2$. On the other hand, given any $a$, with $|a| > 1$, the vector $\mathbf{y}_t$ has also the representation

$$y_{it} = \left[(a_i + b_i L)\frac{1 + aL}{1 + a^{-1}L}\right]\left[\frac{1 - a^{-1}L}{1 - aL}v_t\right] = \frac{(a_i + b_i L)(1 - aL)}{1 - a^{-1}L}w_t, \qquad (2.11)$$

for $i = 1, 2$, where

$$w_t = \frac{1 - a^{-1}L}{1 - aL}v_t = -a^{-1}F\frac{1 - a^{-1}L}{1 - a^{-1}F}v_t$$

is a white noise (this is easily proved by showing that its spectral density is constant). This is a non-fundamental representation for $\mathbf{y}_t$. However, (2.11) is parameterized in $\mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}$, not $\mathbb{R}^2 \times \mathbb{R}^2$.

Now assume that $\mathbf{y}_t$ is infinite dimensional with $y_{it}$ modeled as in (2.7) for $i \in \mathbb{N}$. The vector $\mathbf{y}_t$ is parameterized in

$$\Pi^\infty = \Pi \times \Pi \times \cdots$$

We define negligible sets and genericity in $\Pi^\infty$ with respect to the product topology. Precisely, we say that a subset of $\Pi^\infty$ is negligible if it is *meagre*, i.e. the union of a countable set of nowhere dense subsets; or that a property holds generically in $\Pi^\infty$ if the subset in which it does not hold is meagre.

Define the set $\mathcal{M}_s$, for $s \geq q+1$, as the set of points in $\Pi^\infty$ such that all vectors

$$\mathbf{y}_t^{i_1,i_2,\ldots,i_{q+1}} = (y_{i_1 t}\ y_{i_2 t}\ \cdots\ y_{i_{q+1} t}),$$

with $i_1 < i_2 < \cdots < i_{q+1} \leq s$, can be represented as $\mathbf{A}(L)\mathbf{y}_t^{i_1,i_2,\ldots,i_{q+1}} = \mathbf{R}\mathbf{v}_t$, where $\mathbf{A}(L)$ is of order not greater than $S$ and unique in the sense of Proposition 1, (b). Using Proposition 1 we see that $\mathcal{N}_s = \Pi^\infty - \mathcal{N}_s$ is a nowhere dense subset in the product topology of $\Pi^\infty$, so that the set

$$\mathcal{N} = \cup_{s=q+1}^{\infty}\mathcal{N}_s,$$

being a countable union of nowhere dense subsets of $\Pi^\infty$, is a meagre subset. We can conclude that:

In $\Pi^\infty - \mathcal{N}$, thus generically in $\Pi^\infty$, all vectors

$$\mathbf{y}_t^{i_1,i_2,\ldots,i_{q+1}} = (y_{i_1 t}\ y_{i_2 t}\ \cdots\ y_{i_{q+1} t}),$$

with $i_1 < i_2 < \cdots < i_{q+1}$ (with no upper limit for $i_{q+1}$), can be represented as $A(L)\mathbf{y}_t^{i_1,i_2,\ldots,i_{q+1}} = R\mathbf{v}_t$, where $A(L)$ is of order not greater than $S$ and unique in the sense of Proposition 1, (b).[3]

[3]An analogous genericity result can be obtained if negligible subsets of $\Pi^\infty$are defined as subsets of zero measure with respect to the product measure.

Some observations are in order. Firstly, definining negligible subsets of $\Pi^\infty$ as meagre subsets has a good motivation in the fact that (i) the complement of a meagre subset of $\Pi^\infty$ is not meagre, (ii) if a subset of $\Pi^\infty$ is not meagre, obtaining it as the union of a family of nowhere dense subsets requires a more than countable family.[4]

Secondly, the family of meagre subsets of $\Pi^\infty$ is strictly wider than the family of nowhere dense subsets. In particular, the set $\mathcal{N}$ is not nowhere dense. To see this consider again the MA(1) example $y_{it} = a_i v_t + b_i v_{t-1}$, with $i \in \mathbb{N}$. Denote by

$$\mathbf{c} = (\mathbf{c}_1 \; \mathbf{c}_2 \; \cdots \; \mathbf{c}_n \; \cdots),$$

where $\mathbf{c}_i = (a_i \; b_i)$, a point in $\Pi^\infty$. Given $\mathbf{c} \in \Pi^\infty$ and a neighborhood $G$ of $\mathbf{c}$, a well-known feature of the product topology is that $G$ contains points $\mathbf{c}'$ such that, for some $s$ and all $n > s$, $\mathbf{c}'_n = \mathbf{c}'_s$. Such points obviously belong to $\mathcal{N}$. Thus $\mathcal{N}$ is meagre but dense in $\Pi^\infty$ (in the same way as the rational numbers are a meagre but dense subset of the real numbers).

Lastly, assuming that the parameter space for the polynomials $c_{ij}(L)$ and $d_{ij}(L)$ does not depend on $i$, as we do in (2.7), is convenient but not necessary. With a parameter space depending on $i$, a more general version of Proposition 1 holds as well as the meagreness result for an infinite dimensional vector $\mathbf{y}_t$. However the gain in generality does not seem to make up for the complications in the proof of Proposition 1 and the determination of the order of $\mathbf{A}(L)$.

## 2.5 Autoregressive representations for the vector $\boldsymbol{\chi}_t$

Let us now turn our attention to the vector $\boldsymbol{\chi}_t$. As we have seen, assuming that $\boldsymbol{\chi}_t$ has rational spectral density and that $\mathcal{H}_t^{\chi_s} = \mathcal{H}_t^{\chi}$ for some $s$ implies by Lemmas 1 and 2 that $\boldsymbol{\chi}_t$ has a fundamental rational representation like (2.7). The meagreness argument above motivates assuming that statements (a) and (b) hold for all $(q+1)$-dimensional subvectors of $\boldsymbol{\chi}_t$. Precisely:

**Assumption A.3** The vector $\boldsymbol{\chi}_t$ has a representation

$$\chi_{it} = \frac{c_{i1}(L)}{d_{i1}(L)} u_{1t} + \frac{c_{i2}(L)}{d_{i2}(L)} u_{2t} + \cdots + \frac{c_{iq}(L)}{d_{iq}(L)} u_{qt},$$

where

$$\begin{aligned} c_{if}(L) &= c_{if,0} + c_{if,1} L + \cdots + c_{if,s_1} L^{s_1} \\ d_{if}(L) &= 1 + d_{if,1} L + \cdots + d_{if,s_2} L^{s_2}. \end{aligned}$$

[4] Denote by $\overline{\Pi}$ the closure of $\Pi$. Then: (1) The space $\overline{\Pi}^{\,\infty}$, being the Cartesian product of a countable family of complete metric spaces, is a complete metric space, (2) In complete metric spaces the complement of a meagre subset is not meagre (Baire Category Theorem); see Dunford and Schwartz (1988), p. 32, Lemma 4, and p. 20, Theorem 9 (Baire Theorem), respectively. It is easily seen that the Baire Theorem applies as well to $\Pi^\infty$, which is an open dense subset of $\overline{\Pi}^{\,\infty}$.

for all $i \in \mathbb{N}$ and $f = 1, 2, \ldots, q$. Moreover:
(i) Each vector

$$\boldsymbol{\chi}_t^{i_1,i_2,\ldots,i_{q+1}} = (\chi_{i_1 t}\ \chi_{i_2 t}\ \cdots\ \chi_{i_{q+1} t})',$$

with $i_1 < i_2 < \cdots < i_{q+1}$, has an autoregressive representation

$$\mathbf{A}^{i_1,i_2,\ldots,i_{q+1}}(L)\boldsymbol{\chi}_t^{i_1,i_2,\ldots,i_{q+1}} = \mathbf{R}^{i_1,i_2,\ldots,i_{q+1}}\mathbf{u}_t, \tag{2.12}$$

where $\mathbf{A}^{i_1,i_2,\ldots,i_{q+1}}(L)$ is of order not greater than $S = qs_1 + q^2 s_2$. This implies that $\mathbf{u}_t$ is fundamental for all $(q+1)$-dimensional subvectors of $\boldsymbol{\chi}_t$.
(ii) If $\mathbf{B}(L)\boldsymbol{\chi}_t^{i_1,i_2,\ldots,i_{q+1}} = \mathbf{S}\mathbf{w}_t$ and $\mathbf{B}(L) \neq \mathbf{A}^{i_1,i_2,\ldots,i_{q+1}}(L)$ then $\mathbf{B}(L)$ has order greater than $S$.

An immediate consequence of Assumption A.3 is that $\boldsymbol{\chi}_t$ can be represented as in (1.7), that is

$$\mathbf{A}^1(L)\begin{pmatrix}\chi_{1t}\\ \chi_{2t}\\ \vdots\\ \chi_{q+1,t}\end{pmatrix} = \mathbf{R}^1\mathbf{u}_t, \qquad \mathbf{A}^2(L)\begin{pmatrix}\chi_{q+2,t}\\ \chi_{q+3,t}\\ \vdots\\ \chi_{2(q+1),t}\end{pmatrix} = \mathbf{R}^2\mathbf{u}_t, \qquad \ldots \tag{2.13}$$

or

$$\underline{\mathbf{A}}(L)\boldsymbol{\chi}_t = \underline{\mathbf{R}}\mathbf{u}_t,$$

where the order of the polynomial matrices $\mathbf{A}^k(L)$ does not exceed $S$. Moreover, the polynomials matrices $\mathbf{A}^k(L)$ are unique among autoregressive representations of order not greater than $S$.

Of course, any permutation of the variables produces a different $(q+1)$-blockwise autoregressive representation. Precisely, let $g : \mathbb{N} \to \mathbb{N}$ be one to one and onto, and let $\tilde{\chi}_{it} = \chi_{g(i),t}$. Assumptions A.1, A.2 and A.3 imply that $\tilde{\boldsymbol{\chi}}_t$ has a representation like (2.13), with matrices $\tilde{\mathbf{A}}^k(L)$, $\tilde{\mathbf{R}}^k$ and a white noise vector $\tilde{\mathbf{u}}_t$. Assumption A.3 implies that $\tilde{\mathbf{u}}_t = \mathbf{H}\mathbf{u}_t$, with $\mathbf{H}$ orthogonal.

It must be pointed out that neither $\mathbf{u}_t$ nor $\underline{\mathbf{R}}$ play any special role. Assumption A.3 states that there exists $\mathbf{u}_t$ such that (2.12) holds. All the white-noise vectors and matrices, corresponding to alternative representations are linked to $\mathbf{u}_t$ and $\underline{\mathbf{R}}$ by orthogonal matrices. For identification and estimation of a couple $\mathbf{u}_t^*$, $\underline{\mathbf{R}}^*$ based on economic theory see Section 3.3.

## 2.6 Constructing autoregressive representations of $\boldsymbol{\chi}_t$

Assumption A.3 ensures the existence of the autoregressive representation (2.13). We now show how (2.13), i.e. the matrices $\mathbf{A}^k(L)$ and (up to an orthogonal matrix) $\mathbf{R}^k$, can be *constructed* starting with the spectral density of the $x$'s. In Section 2.7 we discuss an additional assumption, which is necessary to estimate the white noise $\mathbf{u}_t$ (up to an othogonal matrix). The stepwise construction described below provides a population motivation for the estimation procedure studied in Section 3.

(i) The spectral density of the common components $\chi_{it}$ can be consistently estimated using a consistent estimate of the spectral density of the variables $x_{it}$ (see Section 3 below). Thus, for the present purpose, we can assume knowledge of the population spectral density of $\boldsymbol{\chi}_t$.

(ii) Denote by $\boldsymbol{\chi}_t^k$ the $k$-th of the $(q+1)$-dimensional subvectors of $\boldsymbol{\chi}_t$ (which is unobservable) appearing in (2.13), and call $\boldsymbol{\Sigma}_{jk}^{\chi}(\theta)$ the cross spectral density between $\boldsymbol{\chi}_t^j$ and $\boldsymbol{\chi}_t^k$. Then,

$$\boldsymbol{\Gamma}_{jk,s}^{\chi} = \mathrm{E}\left[\boldsymbol{\chi}_t^j \boldsymbol{\chi}_{t-s}^{k\ \prime}\right] = \int_{-\pi}^{\pi} e^{\imath s\theta} \boldsymbol{\Sigma}_{jk}^{\chi}(\theta) d\theta$$

where $\imath$ denotes the imaginary unit.

(iii) Using the autocovariance function $\boldsymbol{\Gamma}_{kk,s}^{\chi}$, we obtain the minimum lag matrix polynomial $\mathbf{A}^k(L)$ and the autocovariance function of the unobservable vectors

$$\boldsymbol{\psi}_t^1 = \mathbf{A}^1(L)\boldsymbol{\chi}_t^1, \quad \boldsymbol{\psi}_t^2 = \mathbf{A}^2(L)\boldsymbol{\chi}_t^2, \quad \ldots \tag{2.14}$$

For, let

$$\mathbf{A}^k(L) = \mathbf{I}_{q+1} - \mathbf{A}_1^k L - \cdots - \mathbf{A}_S^k L^S.$$

Define

$$\mathbf{A}^{[k]} = \begin{pmatrix}\mathbf{A}_1^k & \mathbf{A}_2^k & \cdots & \mathbf{A}_S^k\end{pmatrix}, \quad \mathbf{B}_k^{\chi} = \begin{pmatrix}\boldsymbol{\Gamma}_{kk,1}^{\chi} & \boldsymbol{\Gamma}_{kk,2}^{\chi} & \cdots & \boldsymbol{\Gamma}_{kk,S}^{\chi}\end{pmatrix},$$
$$\mathbf{C}_{jk}^{\chi} = \begin{pmatrix} \boldsymbol{\Gamma}_{jk,0}^{\chi} & \boldsymbol{\Gamma}_{jk,1}^{\chi} & \cdots & \boldsymbol{\Gamma}_{jk,S-1}^{\chi} \\ \boldsymbol{\Gamma}_{jk,-1}^{\chi} & \boldsymbol{\Gamma}_{jk,0}^{\chi} & \cdots & \boldsymbol{\Gamma}_{jk,S-2}^{\chi} \\ \vdots & & & \\ \boldsymbol{\Gamma}_{jk,-S+1}^{\chi} & \boldsymbol{\Gamma}_{jk,-S+2}^{\chi} & \cdots & \boldsymbol{\Gamma}_{jk,0}^{\chi} \end{pmatrix}. \tag{2.15}$$

We have

$$\begin{aligned} \mathbf{A}^{[k]} &= \mathbf{B}_k^{\chi}\left(\mathbf{C}_{kk}^{\chi}\right)^{-1} = \mathbf{B}_k^{\chi}\left(\mathbf{C}_{kk}^{\chi}\right)_{\mathrm{ad}} \det\left(\mathbf{C}_{kk}^{\chi}\right)^{-1}, \\ \boldsymbol{\Gamma}_{jk}^{\psi} &= \boldsymbol{\Gamma}_{jk}^{\chi} - \mathbf{A}^{[j]}\mathbf{C}_{jk}^{\chi}\mathbf{A}^{[k]} \end{aligned} \tag{2.16}$$

where $\mathbf{F}_{\mathrm{ad}}$ denotes the adjoint of the matrix $\mathbf{F}$.

(iv) Assumption A.3 implies that the vectors $\boldsymbol{\psi}_t^k$ are $(q+1)$-dimensional white noises of rank $q$, which span the same $q$-dimensional space. The covariance between $\boldsymbol{\psi}_t^j$ and $\boldsymbol{\psi}_t^k$ can be obtained as follows:

$$\boldsymbol{\Gamma}_{jk}^{\psi} = \mathrm{E}\left[\boldsymbol{\psi}_t^j \boldsymbol{\psi}_t^{k\prime}\right] = \int_{-\pi}^{\pi} e^{\imath\theta}\mathbf{A}^j(e^{-\imath\theta}\boldsymbol{\Sigma}_{jk}^{\chi}(\theta)\boldsymbol{A}^{k\prime}(e^{\imath\theta})d\theta.$$

(v) The $\infty \times \infty$ matrix $\underline{\mathbf{\Gamma}}^{\psi}$, obtained by piecing together the matrices $\mathbf{\Gamma}^{\psi}_{jk}$, is of rank $q$ (see Lemma 2, (iii)) and can therefore be represented as

$$\underline{\mathbf{\Gamma}}^{\psi} = \underline{\mathbf{S}}\,\underline{\mathbf{S}}',$$

where $\underline{\mathbf{S}}$ is an $\infty \times q$ matrix. On the other hand, $\underline{\mathbf{\Gamma}}^{\psi}$ is the covariance matrix of the right-hand side terms in (2.13), so that $\underline{\mathbf{S}} = \underline{\mathbf{R}}\mathbf{H}$, where $\mathbf{H}$ is $q \times q$ and orthogonal. In conclusion, using $\mathbf{x}_t = \boldsymbol{\chi}_t + \boldsymbol{\xi}_t$, we obtain

$$\mathbf{z}_t = \underline{\mathbf{R}}\mathbf{u}_t + \boldsymbol{\phi}_t, \tag{2.17}$$

where $\mathbf{z}_t = \underline{\mathbf{A}}(L)\mathbf{x}_t$, $\boldsymbol{\phi}_t = \underline{\mathbf{A}}(L)\boldsymbol{\xi}_t$.

## 2.7 Assumptions on the representation $\mathbf{z}_t = \underline{\mathbf{R}}\mathbf{u}_t + \boldsymbol{\phi}_t$

Equation (2.17) looks like a static representation of the form (1.2) for $\mathbf{z}_t$, with $r = q$ and $\mathbf{N}(L) = \mathbf{I}_q$. However, the properties that the $q$-th eigenvalue of the spectral density of $\mathbf{z}_t$ diverges a.e. in $[-\pi\ \pi]$, whereas the $(q+1)$-th is essentially bounded, require new assumptions.

Denote by $\lambda^{\psi}_{nj}$ the $j$-th eigenvalue of the spectral density matrix of $\boldsymbol{\psi}_{nt} = \mathbf{R}_n\mathbf{u}_t = \mathbf{A}_n(L)\chi_{nt}$ (a constant spectral density). Our assumption that $\lambda^{\chi}_{nq}(\theta)$ diverges a.e. in $[-\pi\ \pi]$ does not imply that $\lambda^{\psi}_{nq}$ diverges. This is easily seen using the MA(1) example $\chi_{it} = a_i u_t + b_i u_{t-1}$. We have

$$\lambda^{\chi}_{n1}(\theta) = \sum_{i=1}^{n} |a_i + b_i e^{-\imath\theta}|^2, \quad \lambda^{\psi}_{n1} = \sum_{i=1}^{n} a_i^2.$$

Of course the first can diverge for almost all $\theta$ in $[-\pi\ \pi]$ even when the second does not diverge. An additional assumption is therefore necessary, see Section 3.3, Assumption A.10.

Next, we need an assumption implying that $\boldsymbol{\phi}_t$ is an idiosyncratic vector, i.e. that the first eigenvalue of its spectral density is bounded. It is convenient here to assume that $n$, the number of variables, increases by blocks of size $q+1$. Thus $n = m(q+1)$, where $m$ is the number of blocks, so that $n$ and $m$ grow at the same pace. The spectral density of $\mathbf{A}_n(L)\boldsymbol{\xi}_{nt}$ is $\mathbf{A}_n(e^{-i\theta})\mathbf{\Sigma}^{\xi}_n(\theta)\mathbf{A}'_n(e^{i\theta})$. If $\mathbf{a}$ is a unit-modulus, $n$-dimensional column vector, we have

$$\mathbf{a}'\mathbf{A}_n(e^{-\imath\theta})\mathbf{\Sigma}^{\xi}_n(\theta)\mathbf{A}'_n(e^{\imath\theta})\mathbf{a} \le \lambda^{\xi}_{n1}(\theta)\left[\mathbf{a}'\mathbf{A}_n(e^{-\imath\theta})\mathbf{A}'_n(e^{\imath\theta})\mathbf{a}\right] \le \lambda^{\xi}_{n1}(\theta)\lambda^{A_n}_1(\theta),$$

where $\lambda^{A_n}_1(\theta)$ is the first eigenvalue of $\mathbf{A}_n(e^{-i\theta})\mathbf{A}'_n(e^{i\theta})$, which is Hermitian, non-negative definite. By Theorem A, $\sup_n \lambda^{\xi}_{n1}(\theta)$ is essentially bounded. Thus we have to discuss conditions under which $\sup_n \lambda^{A_n}_1(\theta)$ is essentially bounded. This is equivalent to assuming that

$$\sup_{k\in\mathbb{N}} \lambda^{A^k}_1(\theta), \tag{2.18}$$

where $\lambda^{A^k}_1(\theta)$ is the first eigenvalue of $\mathbf{A}^k(e^{-\imath\theta})\mathbf{A}^{k\prime}(e^{\imath\theta})$, is essentially bounded. Inspection of (2.16) shows that the conditions

(i) $\det(\mathbf{C}^{\chi}_{kk})$ is bounded away from zero as a function of $k \in \mathbb{N}$,

(ii) $\gamma^{\chi}_{ij,s} \leq g$ for some $g$ and all $i$, $j$ and $s$,

are sufficient to ensure that there exists a common upper bound for the moduli of the entries of $\mathbf{A}^{[k]}$, and therefore that $\lambda^{A^k}_1(\theta)$ is bounded as a function of $k$ and $\theta$. The next two assumptions take care of conditions (i) and (ii) respectively.

**Assumption A.4** There exists $d > 0$ such that $\det(\mathbf{C}^{\chi}_{kk}) \geq d$ for all $k \in \mathbb{N}$.

**Assumption A.5** There exists a real number $G$ such that

$$|\sigma^{\chi}_{ij}(\theta)| \leq G$$

for all $i \in \mathbb{N}$, $j \in \mathbb{N}$ and $\theta \in [-\pi\ \pi]$.

Under Assumptions A.4 and A.5, plus A.10, see Section 3.3, Theorem A (the viceversa part), ensures that the $q$-th eigenvalue of the spectral density matrix of $\mathbf{z}_t$ diverges a.e. in $[-\pi\ \pi]$, whereas the $(q+1)$-th is essentially bounded. Thus (2.17) is a GDFM representation for $\mathbf{z}_t$. Note that (2.17) is also a static representation, like (1.2), with $r = q$, and can therefore be used to estimate $\mathbf{v}_t$ (up to an orthogonal matrix) and $\boldsymbol{\chi}_t$ by means of standard principal components (see Section 3.3).

# 3 Estimation

## 3.1 Estimation of $\sigma^{\chi}_{ij}(\theta)$ and $\gamma^{\chi}_{ij,k}$

Explicit dependence on the index $n$ has been necessary in Section 2. From now on it is convenient to introduce a minor change in notation and drop $n$ most of the times. In particular:

(1) $\mathbf{\Sigma}^x(\theta) = \left(\sigma^x_{ij}(\theta)\right)_{i,j=1,n}$ and $\lambda^x_f(\theta)$ replace $\mathbf{\Sigma}^x_n(\theta)$ and $\lambda^x_{nf}(\theta)$ respectively.

(2) $\mathbf{\Lambda}^x(\theta)$ denotes the $q \times q$ matrix with $\lambda^x_f(\theta)$ in entry $(f, f)$ and zero elsewhere.

(3) $\mathbf{P}^x(\theta)$ denotes the $n \times q$ matrix with unit-modulus eigenvectors corresponding to the first $q$ eigenvalues on the columns. The columns and entries of $\mathbf{P}^x(\theta)$ are denoted by $\mathbf{P}^x_f(\theta)$ and $p^x_{if}(\theta)$ respectively.

(4) $\mathbf{\Sigma}^{\chi}(\theta) = \left(\sigma^{\chi}_{ij}(\theta)\right)_{i,j=1,n}$, $\lambda^{\chi}_f(\theta)$, $\mathbf{\Lambda}^{\chi}(\theta)$, $\mathbf{P}^{\chi}(\theta)$, etc. are defined in the same way.

(5) All the above matrices and scalars depend on $n$. The corresponding estimates

$$\begin{array}{cccc} \hat{\mathbf{\Sigma}}^x(\theta), & \hat{\lambda}^x_f(\theta), & \hat{\mathbf{\Lambda}}^x(\theta), & \hat{\mathbf{P}}^x(\theta) \\ \hat{\mathbf{\Sigma}}^\chi(\theta), & \hat{\lambda}^\chi_f(\theta), & \hat{\mathbf{\Lambda}}^\chi(\theta), & \hat{\mathbf{P}}^\chi(\theta) \end{array},$$

defined below, depend both on $n$ and the sample $x_{it}$, $i = 1, \dots, n$, $t = 1, \dots, T$. For short we say that they depend on $n$ and $T$.

(6) $\mathbf{A}(L)$ and $\mathbf{R}$ replace $\mathbf{A}_n(L)$ and $\mathbf{R}_n$, to denote the upper left $n \times n$ submatrix of $\underline{\mathbf{A}}(L)$, the $n \times q$ submatrix of $\underline{\mathbf{R}}$, respectively. The corresponding estimated matrices are denoted by $\hat{\mathbf{A}}(L)$ and $\hat{\mathbf{R}}$.

(7) However, to avoid confusion we keep explicit reference to $n$ in $\mathbf{x}_{nt}$, $\boldsymbol{\chi}_{nt}$, $\mathbf{z}_{nt}$ etc. This we have $\mathbf{z}_{nt} = \mathbf{A}(L)\mathbf{x}_{nt} = \mathbf{S}\mathbf{v}_t + \boldsymbol{\phi}_{nt}$, estimated vectors and matrices like $\hat{\boldsymbol{\chi}}_{nt}$, $\hat{\mathbf{z}}_{nt}$, etc.

Let $B_T$ be a sequence of positive integers,

$$\theta_h = \frac{\pi h}{B_T}, \quad h = -B_T, \ \dots, \ B_T,$$

and $\hat{\sigma}_{ij}(\theta_h)$ be the Bartlett lag-window estimate

$$\hat{\sigma}^x_{ij}(\theta_h) = \sum_{k=-B_T}^{B_T} \hat{\gamma}^x_{ij,k} w_k e^{-\imath k \theta_h}, \qquad w_k = 1 - \frac{|k|}{B_T + 1}, \tag{3.1}$$

where $\hat{\gamma}^x_{ij,k}$ is the estimate of $\gamma^x_{ij,k} = \mathrm{E}(x_{it} x_{j,t-k})$. Moreover:

**Assumption A.6** The vector $\mathbf{x}_t$ has a continuous spectral density, i.e. $\sigma^x_{ij}(\theta)$ is continuous for all $i \in \mathbb{N}$ and $j \in \mathbb{N}$.

A consequence of Assumptions A.3 and A.6 is that $\boldsymbol{\xi}_t$ has a continouous spectral density as well, so that the eigenvalue $\lambda^\xi_{n1}(\theta)$ is continuous. It is easily seen that essential boundedness of $\bar{\lambda}^\xi_1(\theta) = \sup_n \lambda^\xi_{n1}(\theta)$, see Theorem A(i), is replaced by boundedness, i.e. there exists $\lambda > 0$ such that $\lambda^\xi_{n1}(\theta) \le \lambda$ for all $n$ and $\theta$.

Liu and Wu (2010) prove that, considering the spectral densities $\sigma^x_{ii}(\theta)$,

$$\max_{|h| \le B_T} |\hat{\sigma}^x_{ii}(\theta_h) - \sigma^x_{ii}(\theta_h)|^2 = O_P\left(\frac{B_T}{T} \log T\right),$$

under fairly general conditions, including Assumption A.6 and the assumption that $B_t$ diverges at the same speed as $T^\gamma$, with $1/3 < \gamma < 1$ (see Theorems 3, 4, 5 and Remark 5). A discussion of the conditions imposed by Liu and Wu is outside the scope of the present paper. Rather, we directly assume that Liu and Wu's results hold for both the diagonal entries $\sigma^x_{ii}(\theta)$ and the cross-spectra $\sigma^x_{ij}(\theta)$, $i \neq j$. Moreover, we impose uniformity of the convergence with respect to $i$ and $j$. The present paper shows that under such conditions the matrices $\mathbf{A}^k(L)$, the common components and the common shocks can be estimated consistently, and determines consistency rates.

In a preliminary version, our assumption is the following:

**Assumption A$^*$** Let $B_T$ be the greatest integer less or equal to $dT^\gamma$, for some $d > 0$ and $1/3 < \gamma < 1$, and let $\rho_T = \frac{T}{B_T \log T}$. Given $\epsilon > 0$ there exist $\eta_\epsilon$ and $T_\epsilon$, *independent of* $i$ *and* $j$, such that

$$\mathrm{P}\left(\rho_T \max_{|h|\leq B_T} |\hat{\sigma}^x_{ij}(\theta_h) - \sigma^x_{ij}(\theta_h)|^2 > \eta_\epsilon\right) < \epsilon \tag{3.2}$$

for all $T > T_\epsilon$, $i \in \mathbb{N}$, $j \in \mathbb{N}$.

It can be convenient to rephrase (3.2) as

$$\max_{|h|\leq B_T} |\hat{\sigma}^x_{ij}(\theta_h) - \sigma^x_{ij}(\theta_h)|^2 = O_P(\rho_T^{-1}), \tag{3.3}$$

for $T \to \infty$, uniformly in $i$ and $j$. However, one must keep in mind that here "uniformity in $i$ and $j$" only means that $\eta_\epsilon$ and $T_\epsilon$ are independent of $i$ and $j$, not that the $O_P(\rho_T^{-1})$ in (3.3) is independent of $i$ and $j$. This has the consequence that (3.3) does not imply that for $T \to \infty$ and $n \to \infty$,

$$n^{-2}\sum_{i=1}^{n}\sum_{j=1}^{n} \max_{|h|\leq B_T} |\hat{\sigma}^x_{ij}(\theta_h) - \sigma^x_{ij}(\theta_h)|^2 = O_P(\rho_T^{-1}). \tag{3.4}$$

For that matter, there exist sequences such that (3.3) holds uniformly in $i$ and $j$ whereas (3.4) does not hold, see the counterexample in Appendix B. However, our counterexample exhibits an extremely complicated *ad hoc* pattern of dependence between the stochastic variables $\max_{|h|\leq B_T} |\hat{\sigma}^x_{ij}(\theta_h) - \sigma^x_{ij}(\theta_h)|^2$. We take therefore Appendix B as a motivation for our next assumption, which is (3.4), in a slightly more general form.

**Assumption A.7** Let $B_T$ be the greatest integer less or equal to $dT^\gamma$, for some $d > 0$ and $1/3 < \gamma < 1$, and let $\rho_T = \frac{T}{B_T \log T}$. Let $I_1 \subset I_2 \subset \cdots$ be a sequence of subsets of $\mathbb{N} \times \mathbb{N}$. Denote by $n_k$ the number of elements in $I_k$. Given $\epsilon$, there exists $\eta_\epsilon$, $T_\epsilon$ and $n_\epsilon$, independent of the sequence of sets $I_k$, such that

$$\mathrm{P}\left(\rho_T \frac{1}{n_k} \sum_{(i,j)\in I_k} \max_{|h|\leq B_T} |\hat{\sigma}^x_{ij}(\theta_h) - \sigma^x_{ij}(\theta_h)|^2 > \eta_\epsilon\right) < \epsilon,$$

for all $T > T_\epsilon$ and $n > n_\epsilon$. Equivalently,

$$\frac{1}{n_k} \sum_{(i,j)\in I_k} \max_{|h|\leq B_T} |\hat{\sigma}^x_{ij}(\theta_h) - \sigma^x_{ij}(\theta_h)|^2 = O_P(\rho_T^{-1}), \tag{3.5}$$

as $T \to \infty$ and $k \to \infty$, uniformly with respect to the sequence $I_k$.

In particular,

$$\frac{1}{n}\sum_{j=1}^{n}\max_{|h|\leq B_T}|\hat{\sigma}_{ij}^{x}(\theta_h)-\sigma_{ij}^{x}(\theta_h)|^2 = O_P(\rho_T^{-1}), \tag{3.6}$$

as $T\to\infty$ and $n\to\infty$, uniformly in $i$. Note that we do not impose $\lim_{k\to\infty} n_k = \infty$. Thus Assumption A.7 implies Assumption A$^*$.

Our estimate of the spectral density of $\boldsymbol{\chi}_{nt}$ is defined as in Forni, Hallin, Lippi and Reichlin (2000):

$$\hat{\mathbf{\Sigma}}^{\chi}(\theta_h) = \hat{\mathbf{P}}^{x}(\theta_h)\hat{\mathbf{\Lambda}}^{x}(\theta_h)\tilde{\hat{\mathbf{P}}}^{x}(\theta_h),$$

where $\tilde{\mathbf{F}}$ denotes the transposed, conjugate of the matrix $\mathbf{F}$.

Another assumption is necessary to prove Propositions 2 and 3.

**Assumption A.8** There exist real numbers $a_1 > b_1 > a_2 > b_2 > \cdots > a_q > b_q > 0$, and an integer $\bar{n}$ such that for $n > \bar{n}$,

$$a_s \geq \frac{\lambda_s^{\chi}(\theta)}{n} \geq b_s,$$

for $s = 1, \ldots, q$, and all $\theta \in [-\pi\ \pi]$.

We use henceforth the notation:

$$\zeta_{Tn} = \max\left(\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{\rho_T}}\right). \tag{3.7}$$

**Proposition 2** Under Assumptions A.1 through A.8, given $\epsilon > 0$ there exist $\eta_\epsilon$, $T_\epsilon$, $n_\epsilon$, such that

$$\mathrm{P}\left[\zeta_{Tn}^{-1}\max_{|h|\leq B_T}|\hat{\sigma}_{ij}^{\chi}(\theta_h)-\sigma_{ij}^{\chi}(\theta_h)| > \eta_\epsilon\right] < \epsilon,$$

for all $T > T_\epsilon$, $n > n_\epsilon$, $i\in\mathbb{N}$, $j\in\mathbb{N}$. Equivalently, for $T\to\infty$, $n\to\infty$,

$$\max_{|h|\leq B_T}|\hat{\sigma}_{ij}^{\chi}(\theta_h)-\sigma_{ij}^{\chi}(\theta_h)| = O_P\left(\zeta_{Tn}\right),$$

uniformly in $i$ and $j$ (Proof in Appendix C).

Our estimator of $\gamma_{ij,\ell}^{\chi}$, the covariance between $\chi_{it}$ and $\chi_{j,t-\ell}$, is defined as in Forni et al. (2005):

$$\hat{\gamma}_{ij,\ell}^{\chi} = \frac{\pi}{B_T}\sum_{s=-B_T+1}^{B_T} e^{\imath\ell\theta_s}\hat{\sigma}_{ij}^{\chi}(\theta_s). \tag{3.8}$$

Defining

$$\alpha_s = -\pi + (s + B_T - 1)\frac{\pi}{B_T}, \quad \beta_s = -\pi + (s + B_T)\frac{\pi}{B_T},$$

and recalling that

$$\gamma^{\chi}_{ij,\ell} = \int_{-\pi}^{\pi} e^{\imath\ell\theta}\sigma^{\chi}_{ij}(\theta)d\theta,$$

we have, using Assumption A.5,

$$\begin{aligned}
|\hat{\gamma}^{\chi}_{ij,\ell} \; -\gamma^{\chi}_{ij,\ell}| &\leq \frac{\pi}{B_T}\sum_{s=-B_T+1}^{B_T}|e^{\imath\ell\theta_s}\hat{\sigma}^{\chi}_{ij}(\theta_s) - e^{\imath\ell\theta_s}\sigma^{\chi}_{ij}(\theta_s)| \\
&\quad + \left|\frac{\pi}{B_T}\sum_{s=-B_T+1}^{B_T} e^{\imath\ell\theta_s}\sigma^{\chi}_{ij}(\theta_s) - \int_{-\pi}^{\pi} e^{\imath\ell\theta}\sigma^{\chi}_{ij}(\theta)d\theta\right| \\
&\leq \frac{\pi}{B_T}\sum_{s=-B_T+1}^{B_T}|\hat{\sigma}^{\chi}_{ij}(\theta_s) - \sigma^{\chi}_{ij}(\theta_s)| \\
&\quad + \frac{\pi}{B_T}\sum_{s=-B_T+1}^{B_T}\max_{\alpha_s\leq\theta\leq\beta_s}|e^{\imath\ell\theta_s}\sigma^{\chi}_{ij}(\theta_s) - e^{\imath\ell\theta}\sigma^{\chi}_{ij}(\theta)| \\
&\leq \pi\max_{|s|\leq B_T}|\hat{\sigma}^{\chi}_{ij}(\theta_s) - \sigma^{\chi}_{ij}(\theta_s)| \\
&\quad + \frac{\pi G}{B_T}\sum_{s=-B_T+1}^{B_T}\max_{\alpha_s\leq\theta\leq\beta_s}|e^{\imath\ell\theta_s} - e^{\imath\ell\theta}| \\
&\quad + \frac{\pi}{B_T}\sum_{s=-B_T+1}^{B_T}\max_{\alpha_s\leq\theta\leq\beta_s}|\sigma^{\chi}_{ij}(\theta_s) - \sigma^{\chi}_{ij}(\theta)|
\end{aligned} \tag{3.9}$$

Of course the function $e^{\imath\ell\theta}$ is of bounded variation. Under Assumption A.3 the functions $\sigma^{\chi}_{ij}(\theta)$ are of bounded variation as well. We assume more:

**Assumption A.9** The functions $\sigma^{\chi}_{ij}(\theta)$ and $\sigma^{\xi}_{ij}(\theta)$ are of bounded variation, uniformly in $i$ and $j$. Precisely, there exists $M$ such that for all $i$, $j$ in $\mathbb{N}$ and all

$$-\pi = \theta_0 < \theta_1 < \cdots < \theta_w = \pi,$$

we have

$$\sum_{k=1}^{w}|\sigma^{\chi}_{ij}(\theta_k) - \sigma^{\chi}_{ij}(\theta_{k-1})| \leq M, \quad \sum_{k=1}^{w}|\sigma^{\xi}_{ij}(\theta_k) - \sigma^{\xi}_{ij}(\theta_{k-1})| \leq M.$$

The same property obviously holds for $\sigma^{x}_{ij}(\theta)$.

Using Proposition 2 and Assumption A.9, we obtain:

$$|\hat{\gamma}^{\chi}_{ij,\ell} - \gamma^{\chi}_{ij,\ell}| = O_P(\zeta_{Tn}) + O(1/B_T),$$

uniformly in $i$ and $j$. Given the definition of $B_T$ and $\rho_T$, and the inequality $1/3 < \delta < 1$, we have

**Proposition 3** Under Assumptions A.1 through A.9, for each $\ell \geq 0$, given $\epsilon > 0$ there exist $\eta_\epsilon$, $T_\epsilon$ and $n_\epsilon$, such that

$$\mathrm{P}\left[\zeta_{Tn}^{-1}|\hat{\gamma}^{\chi}_{ij,\ell} - \gamma^{\chi}_{ij,\ell}| > \eta_\epsilon\right] < \epsilon,$$

for all $T > T_\epsilon$, $n > n_\epsilon$, $i \in \mathbb{N}$, $j \in \mathbb{N}$. Equivalently, for each $\ell \geq 0$, for $T \to \infty$, $n \to \infty$,

$$|\hat{\gamma}^{\chi}_{ij,k} - \gamma^{\chi}_{ij,k}| = O_P\left(\zeta_{Tn}\right), \tag{3.10}$$

uniformly in $i$ and $j$.

Note that, as the last term in (3.9) contains

$$\frac{\pi G}{B_T} \sum_{s=-B_T+1}^{B_T} \max_{\alpha_s \leq \theta \leq \beta_s} |e^{\imath \ell \theta_s} - e^{\imath \ell \theta}|,$$

(3.10) is not uniform with respect to $\ell$. However, estimation of the matrices $\mathbf{A}^k(L)$, corresponding to the blocks $\mathbf{x}^k_t$, requires only the covariances $\hat{\gamma}_{ij,\ell}$ with $\ell \leq S$, and $S$ is independent of the block.

## 3.2 Estimation of $\mathbf{A}^k(L)$ and $\mathbf{\Gamma}^{\psi}_{jk}$

The definition of the estimates $\hat{\mathbf{A}}^{[k]}$, $\hat{\mathbf{B}}^{\chi}_{kk}$, $\hat{\mathbf{C}}^{\chi}_{jk}$ and $\hat{\mathbf{\Gamma}}^{\psi}_{jk}$ is obvious, see (2.15) and (2.16). Moreover, given an $s_1 \times s_2$ matrix $\mathbf{F}$, $\|\mathbf{F}\|$ denotes the spectral norm of $\mathbf{F}$, see Appendix C for details.

**Proposition 4** Under Assumptions A.1 through A.9, given $\epsilon > 0$ there exist $\eta_\epsilon$, $T_\epsilon$ and $n_\epsilon$ such that

$$\mathrm{P}\left[\zeta_{Tn}^{-1}\|\hat{\mathbf{A}}^{[k]} - \mathbf{A}^{[k]}\| > \eta_\epsilon\right] < \epsilon$$
$$\mathrm{P}\left[\zeta_{Tn}^{-1}\|\hat{\mathbf{\Gamma}}^{\psi}_{jk} - \mathbf{\Gamma}^{\psi}_{jk}\| > \eta_\epsilon\right] < \epsilon$$

for all $T > T_\epsilon$, $n > n_\epsilon$, $k \in \mathbb{N}$ and $j \in \mathbb{N}$. Equivalently, for $T \to \infty$, $n \to \infty$,

$$\|\hat{\mathbf{A}}^{[k]} - \mathbf{A}^{[k]}\| = O_P\left(\zeta_{Tn}\right)$$
$$\|\hat{\mathbf{\Gamma}}^{\psi}_{jk} - \mathbf{\Gamma}^{\psi}_{jk}\| = O_P\left(\zeta_{Tn}\right)$$

uniformly in $k$ and $j$. (Proof in Appendix D.)

## 3.3 Estimation of $\underline{\mathbf{R}}$

Let us now consider the vector $\mathbf{z}_t$ with its representation (2.17): $\mathbf{z}_t = \underline{\mathbf{R}}\mathbf{u}_t + \boldsymbol{\phi}_t$. For $n \geq q$,

$$\mathbf{z}_{nt} = \boldsymbol{\psi}_{nt} + \boldsymbol{\phi}_{nt} = \mathbf{R}\mathbf{u}_t + \boldsymbol{\phi}_{nt}.$$

The covariance matrix of $\boldsymbol{\psi}_{nt}$ is

$$\mathbf{R}\mathbf{R}' = \mathbf{P}^{\psi}\mathbf{\Lambda}^{\psi}\mathbf{P}^{\psi\prime} = \left[\mathbf{P}^{\psi}\mathbf{\Lambda}^{\psi\frac{1}{2}}\right]\left[\mathbf{\Lambda}^{\psi\frac{1}{2}}\mathbf{P}^{\psi\prime}\right],$$

where $\mathbf{P}^\psi$ is the $n \times q$ matrix with the eigenvectors of $\mathbf{RR}'$ on the columns and $\mathbf{\Lambda}^\psi$ is $q \times q$ with the corresponding eigenvalues on the main diagonal. Thus we have the representation

$$\mathbf{z}_{nt} = \left[\mathbf{P}^\psi \mathbf{\Lambda}^{\psi\frac{1}{2}}\right] \mathbf{v}_t + \boldsymbol{\phi}_{nt} = \mathcal{R}\mathbf{v}_t + \boldsymbol{\phi}_{nt},$$

where $\mathbf{v}_t = \mathbf{H}\mathbf{u}_t$, with $\mathbf{H}$ orthogonal. Note that, given $i$ and $f$, the entry $(i, f)$ of $\mathcal{R}$ depends on $n$, so that the matrices $\mathcal{R}$ are not nested, nor is $\mathbf{v}_t$ independent of $n$. However, the product of each row of $\mathcal{R}$ by $\mathbf{v}_t$ equals the corresponding coordinate of $\boldsymbol{\psi}_{nt}$ and is therefore independent of $n$.

Our estimate of $\mathcal{R} = \mathbf{P}^\psi \mathbf{\Lambda}^{\psi\frac{1}{2}}$ is $\hat{\mathcal{R}} = \hat{\mathbf{P}}^z \hat{\mathbf{\Lambda}}^{z\frac{1}{2}}$, where $\hat{\mathbf{P}}^z$ and $\hat{\mathbf{\Lambda}}^z$ are the eigenvectors and eigenvalues, respectively, of the variance-covariance matrix of $\hat{\mathbf{z}}_{nt} = \hat{\mathbf{A}}(L)\mathbf{x}_{nt}$, that is $\mathbf{x}_{nt}$ filtered with the *estimated* matrices $\hat{\mathbf{A}}(L)$. This is the reason for the complications we have to deal with in Appendix E. The following additional assumption is necessary to prove consistency:

**Assumption A.10** There exist real numbers $h_1 > k_1 > h_2 > k_2 \cdots > h_q > k_q > 0$, and an integer $\bar{n}$ such that for $n > \bar{n}$,

$$h_s \geq \frac{\lambda_s^\psi}{n} \geq k_s,$$

for $s = 1, \ldots, q$.

**Proposition 5** Under Assumptions A.1 through A.10, for all $\epsilon > 0$ there exist $\eta_\epsilon$, $n_\epsilon$ and $T_\epsilon$, such that

$$\mathrm{P}\left(\zeta_{Tn}^{-1} \|\hat{\mathcal{R}}_i - \mathcal{R}_i \hat{\mathbf{W}}^z\| > \eta_\epsilon\right) < \epsilon,$$

for all $n > n_\epsilon$, $T > T_\epsilon$, and $i$, where $\mathcal{R}_i$ is the $i$-th row of $\mathcal{R}$, $\hat{\mathbf{W}}^z$ is a $q \times q$ diagonal matrix, depending on $n$ and $T$, whose diagonal entries equal either 1 or $-1$. Equivalently,

$$\|\hat{\mathcal{R}}_i - \mathcal{R}_i \hat{\mathbf{W}}^z\| = O_P(\zeta_{Tn}),$$

as $T \to \infty$ and $n \to \infty$, uniformly in $i$ (Proof in Appendix E).

Let us point out again that the $i$-th row of $\mathcal{R}$ depends on $n$. Therefore Proposition 5 only states that the difference between the estimated entries of $\hat{\mathcal{R}}$ and the entries of $\mathcal{R}$ converges to zero (upon sign correction), not that the estimated entries converge. Now suppose that the common shocks can be identified by means of economically meaningful statements. For example, suppose that we have good reasons to claim that the upper $q \times q$ matrix of the "structural" representation is lower triangular with positive diagonal entries (an iterative scheme for the first $q$ common components). As is well known, such conditions determine exactly one representation, denote it by $\mathbf{z}_t = \underline{\mathbf{R}}^* \mathbf{u}_t^* + \boldsymbol{\phi}_t$, or $\mathbf{z}_{nt} = \mathbf{R}^* \mathbf{u}_t^* + \boldsymbol{\phi}_t$, where the $n \times n$

matrices $\mathbf{R}^*$ are nested. In particular, starting with $\mathbf{x}_{nt} = \mathcal{R}\mathbf{v}_t + \boldsymbol{\phi}_{nt}$, there exists exactly one orthogonal matrix $\mathbf{G}(\mathcal{R})$ (actually $\mathbf{G}(\mathcal{R})$ only depends on the $q \times q$ upper submatrix of $\mathcal{R}$) such that $\mathbf{R}^* = \mathcal{R}\mathbf{G}(\mathcal{R})'$. Thus while the entries of $\mathcal{R}$ depend on $n$, the entries of $\mathcal{R}\mathbf{G}(\mathcal{R})'$ do not.

Applying the same rule to $\hat{\mathcal{R}}$ we obtain the matrices $\hat{\mathbf{R}}^* = \hat{\mathcal{R}}\mathbf{G}(\hat{\mathcal{R}})'$. It is easily seen that each entry of $\hat{\mathbf{R}}^*$ (depending on $n$ and $T$) converges to the corresponding entry of $\mathbf{R}^*$ (independent of $n$ and $T$) at rate $\zeta_{Tn}$.

An iterative identification scheme will be used in Section 4 to compare different estimates of the impulse-response functions.[5]

### 3.4 Estimation of $\mathbf{v}_t$

Our estimator of $\mathbf{v}_t$ is the projection of $\hat{\mathbf{z}}_t$ on $\hat{\mathbf{P}}^z(\hat{\mathbf{\Lambda}}^z)^{-\frac{1}{2}}$:

$$\hat{\mathbf{v}}_t = ((\hat{\mathbf{\Lambda}}^z)^{\frac{1}{2}}\hat{\mathbf{P}}^{z\prime}\hat{\mathbf{P}}^z(\hat{\mathbf{\Lambda}}^z)^{\frac{1}{2}})^{-1}(\hat{\mathbf{\Lambda}}^z)^{\frac{1}{2}}\hat{\mathbf{P}}^{z\prime}\hat{\mathbf{z}}_t = (\hat{\mathbf{\Lambda}}^z)^{-\frac{1}{2}}\hat{\mathbf{P}}^{z\prime}\hat{\mathbf{z}}_t$$

**Proposition 6** Under Assumptions A.1 through A.10, for all $\epsilon > 0$ there exist $\eta_\epsilon$, $n_\epsilon$ and $T_\epsilon$, such that

$$\mathrm{P}\left(\zeta_{Tn}^{-1}\| \left(\hat{\mathbf{v}}_t - \hat{\mathbf{W}}^z\mathbf{v}_t\right) \| > \eta_\epsilon\right) < \epsilon,$$

where $\hat{\mathbf{W}}^z$ is a $q \times q$ diagonal matrix, depending on $n$ and $T$, whose diagonal entries equal either 1 or $-1$. Equivalently

$$\|\hat{\mathbf{v}}_t - \hat{\mathbf{W}}^z\mathbf{v}_t\| = O_P(\zeta_{Tn}),$$

as $T \to \infty$ and $n \to \infty$

Proof in Appendix F.

# 4 An empirical exercise

In the present section we compare the estimation performance of the model s-tudied in the present paper, we refer to this as FHLZ, and model (1.3), whose consistency has been studied in Forni et al. (2009), referred to as FGLR. Let us recall that both models assume rational spectral density for the common components, but FGLR also assumes finite dimension for the space spanned by the variables $\chi_{it}$ for any given $t$ and $i \in \mathbb{N}$. Using a Monte Carlo simulation based on actual US macroeconomic data, we compare (i) impulse response functions and (ii) structural shocks, estimated using FHLZ and FGLR, under the same iterative identification scheme.

[5] All just-identifying rules considered in the SVAR literature can be dealt with along the same lines, see Forni, Giannone, Lippi and Reichlin, 2009,

## 4.1 Simulation design

Let us firstly illustrate the simulation design. We use two macroeconomic panels. The first is the one used in Forni and Gambetti (2010a), with 101 US quarterly series, covering the period 1959 I - 2007 IV. The second is the one used in Forni and Gambetti (2010b), which is essentially an updating of the panel used in Stock and Watson (2002a, 2002b) and Bernanke, Boivin and Eliasz (2005). It includes 112 US monthly series between March 1973 and November 2007. Details on both panels and their treatment are reported in Appendix G.

We estimate FGLR and FHLZ for both the quartely and the monthly panels:

(1) We set $q = 4$ for both panels on the basis of Hallin and Liska (2007) and Onatsky (2009).

(2) In FGLR we set $r$, the dimension of $\mathbf{F}_t$, equal to 12 for the quarterly panel and 16 for the monthly panel by using Bai and Ng (2002) IC2 criterion. Moreover, $p$, the number of lags in the VAR, see the second equation in (1.3), is set equal to 2 for both panels as suggested by the the BIC criterion.

(3) In FHLZ, the number of lags in each $(q + 1)$-dimensional VAR is chosen by the BIC criterion for each VAR. As regards the estimation of the spectral density of $\mathbf{x}_{nt}$, we set the Bartlett lag-window size to 12 for quarterly data and 30 for monthly data, large enough to retain the most important cyclical auto- and cross-correlations.

(4) Our simulations require an estimate of the idiosyncratic components. For FGLR, we take the residuals of the projection of the $x$'s on the first $r$ principal components. For FHLZ, we take the residuals of the projection of $x_{it}$ onto present, past and future values of the first $q$ dynamic principal components, i.e. we apply the two-sided filters described in Forni et al. (2000) to get an estimate of the common components and take the deviations from the $x$'s. The resulting poor end-of-sample estimation, see the Introduction, is harmless for the present purposes.

(5) Impulse-response functions and shocks are estimated by FHLZ and FGLR applying the same identification scheme. Precisely, we use a recursive identification scheme on 4 selected variables (see Section 3.2). With quarterly data we use GDP, the GDP deflator, the federal funds rate and the S-tandard & Poor's index of 500 Common Stocks; for monthly data we use Industrial Production, the CPI, the federal funds rate and the NAPM commodity price index. Similar schemes are widely used in the VAR literature on monetary policy (see e.g. Christiano, Eichenbaum and Evans, 1999).[6]

[6] We take the variables in the order specified above. The ordering however is irrelevant in the present context, since our measure of the estimation error (specified below) is invariant to the application of the same orthonormal transformations to both the target and the estimator.

(6) Regarding FHLZ, whereas the population impulse-response functions and the shocks, given the identification scheme, do not depend on the particular grouping of the variables, some dependence occurs in the estimates. This is due to the estimation errors contained in the covariances used in VAR estimation, and possibly to incorrect specification of the number of lags. We deal with this problem by averaging the impulse-response functions obtained with a number of random permutations of the variables in the panel. We find that 30 random permutations are sufficient to stabilize the averages.

Denote by $(\text{IRF}, \text{IC})_{\text{FGLR}}$ and $(\text{IRF}, \text{IC})_{\text{FHLZ}}$ the impulse-response functions and idiosyncratic components estimated using FGLR and FHLZ respectively. Using $(\text{IRF}, \text{IC})_{\text{FGLR}}$ we generate 500 artificial quarterly panels and 200 artificial monthly panels as follows. First, we produce 4 random independent standard normal shocks, filter them with the impulse-response functions and add the resulting series to get the common components (the impulse-response functions are truncated at 20 lags for quarterly data and 48 lags for monthly data. Then we add artificial idiosyncratic components obtained by block bootstrapping (without overlapping) the idiosyncratic components estimated as in (4) above. We take blocks of 19 periods for quarterly data and 51 periods for monthly data. Block bootstrapping is intended to randomize the idiosyncratic components while preserving the idiosyncratic auto- and cross-correlation structure of macroeconomic time series. The same procedure is applied to obtain the 500 quarterly and 200 monthly panels based on $(\text{IRF}, \text{IC})_{\text{FHLZ}}$.

Lastly, impulse-response functions and shocks are estimated for each artificial panel using the two competing methods, with the recursive identification scheme described above in (5):

(a) The true number of structural shocks is assumed to be known, i.e. equal to 4, for both methods.

(b) For each artificial panel we set the parameters $r$, $p$ when using FGLR, the lag-window size and the length of the $(q+1)$-dimensional VAR's, when using FHLZ, according to the criteria specified above in (2) and (3). In particular, for FHLZ, the BIC criterion is applied separately to each one of the of the $(q+1)$-dimensional VAR's, the (Bartlett) lag window size is 12 for quarterly data and 30 for monthly data (we do not report results for different values of the lag-window size since they are fairly stable within the range 2-4 years). The number of random permutations of the variables in the panel, see (6) above, was set to 30 for all experiments. Results obtained by using the HQC criterion to determine the length of the VAR's, in both methods, are also reported. On the other hand, both the AIC and the Bai and Ng (2002) IC1 criterion produce poor results, which are not reported.

(c) We also report results obtained using a grid of values for some parameters. For estimates with FGLR, we report results for $r = 6, 12$, $p = 1, 2, 3,6$ (quarterly data), $r = 8, 16$, $p = 2, 4, 6, 12$ (monthly data). For estimates with FHLZ, we report results obtained by setting the length of all the $(q + 1)$-dimensional VAR's to $1, 2, 3, 6$ (quarterly data), $2, 4, 6, 12$ (monthly data).

The estimation error for impulse-response functions and shocks is defined as the normalized sum of the squared deviations of the estimated from the "true" impulse response coefficients and shocks respectively. Precisely, let $b_{if,h}$ and $\hat{b}_{if,h}$ be the true and estimated impulse-response coefficients, respectively, of variable $i$, shock $f$, lag $h$. Then the estimation error of the impulse response functions is measured by

$$\text{MSE(irf)} = \frac{\sum_{i=1}^{n}\sum_{f=1}^{q}\sum_{h=0}^{H}\left(\hat{b}_{if,h} - b_{if,h}\right)^2}{\sum_{i=1}^{n}\sum_{f=1}^{q}\sum_{h=0}^{H} b_{if,h}^2}.$$

The truncation lag $H$ was set to 20 for quarterly data and 48 for monthly data. Similarly, the estimation error of the shocks is defined as

$$\text{MSE(shocks)} = \frac{\sum_{t=1}^{T}\sum_{f=1}^{q}\left(\hat{u}_{ft} - u_{ft}\right)^2}{\sum_{t=1}^{T}\sum_{f=1}^{q} u_{ft}^2},$$

where $u_{ft}$ and $\hat{u}_{ft}$ are the true and estimated values, respectively, of the shock $u_{ft}$ at time $t$.

## 4.2 Results

Tables 1-4 report the average and the corresponding standard deviation (slanted figures) of the estimation error across the artificial panels. Results for impulse-response functions are reported in the upper half of the tables, for shocks in the lower half. Boldface figures indicate the best result.

The first four columns report the results obtained using preassigned values for the lag in the VAR's, see (c) above, the last two columns the results obtained using the BIC or HQC criterion, see (b) above. The second and third row report results for FHLR when $r$ is set to preassigned values, see (c) above, the fourth when $r$ is chosen according to $IC2$, see (b).

Comparing the best results provides an idea of the potential performance of the competing methods, independently of model selection techniques. We see that the FHLZ outperforms FGLR even when the DGP is FGLR (Tables 1 and 3), with the exception of the impulse response functions in Table 3, upper half. However, the relevant comparison is in the last two columns, first and fourth rows. Underscored figures in the first row report results from FHLZ, with the VAR length chosen according to the BIC or the HQC criteria. Results from FGLR, with the VAR length chosen with the same criteria and $r$ by $IC2$, are reported in the fourth row. The basic outcome for the impulse-response functions does not change: FHLZ outperforms FGLR in 7 out of 8 cases, with differences that are often larger than those between boldface figures. The standard deviations show that the performance of FHLZ is almost always less volatile than FGLR, often to a large extent.

As regards the shocks, FHLZ performs uniformly better than FGLR, reducing the MSE by 13-38% for quarterly data, 28-46% for monthly data. Moreover, the relative performance does not depend on the DGP generating the data.

Table 5 quantifies the improvement of FHLZ over FGLR, by reporting the ratio FHLZ/FGLR for each one of the 8 cases.

**Table 1**: Average and standard deviation (slanted) of MSE across 500 artificial data set. Data generating Process (DGP): FGLR. Quarterly data.

| *irf* | $p = 1$ | $p = 2$ | $p = 3$ | $p = 6$ | $p = BIC$ | $p = HQC$ |
|---|---|---|---|---|---|---|
| FHLZ | 0.2494 | 0.1915 | **0.1857** | 0.2447 | <u>0.2040</u> | <u>0.1928</u> |
| | *0.0256* | *0.0274* | *0.0281* | *0.0388* | *0.0266* | *0.0281* |
| FGLR r = 6 | 0.2468 | 0.2030 | 0.2276 | 0.2937 | 0.2288 | 0.2070 |
| | *0.0490* | *0.0628* | *0.0714* | *0.0699* | *0.0604* | *0.0604* |
| FGLR r = 12 | 0.2137 | **0.1862** | 0.2163 | 0.3302 | 0.2137 | 0.1959 |
| | *0.0298* | *0.0321* | *0.0349* | *0.0445* | *0.0298* | *0.0360* |
| FGLR r = IC2 | 0.2305 | 0.1931 | 0.2190 | 0.3095 | <u>0.2226</u> | <u>0.2160</u> |
| | *0.0369* | *0.0476* | *0.0518* | *0.0764* | *0.0663* | *0.0824* |
| *shocks* | | | | | | |
| FHLZ | 0.1748 | **0.1245** | 0.1325 | 0.2491 | <u>0.1311</u> | <u>0.1362</u> |
| | *0.0290* | *0.0235* | *0.0217* | *0.0423* | *0.0239* | *0.0225* |
| FGLR r = 6 | 0.2027 | **0.1708** | 0.1951 | 0.2933 | 0.1904 | 0.1740 |
| | *0.0664* | *0.0772* | *0.0854* | *0.0817* | *0.0734* | *0.0762* |
| FGLR r = 12 | 0.1997 | 0.2041 | 0.2638 | 0.4921 | 0.1997 | 0.2057 |
| | *0.0287* | *0.0334* | *0.0351* | *0.0440* | *0.0287* | *0.0416* |
| FGLR r = IC2 | 0.1961 | 0.1784 | 0.2157 | 0.3710 | <u>0.2019</u> | <u>0.2199</u> |
| | *0.0367* | *0.0647* | *0.0834* | *0.1655* | *0.1186* | *0.1755* |

**Table 2**: Average and standard deviation (slanted) of MSE across 500 artificial data set. DGP: FHLZ. Quarterly data.

| *irf* | $p=1$ | $p=2$ | $p=3$ | $p=6$ | $p=BIC$ | $p=HQC$ |
|---|---|---|---|---|---|---|
| FHLZ | 0.1401 | 0.1186 | 0.1287 | 0.1740 | **0.1184** | 0.1280 |
| | *0.0179* | *0.0182* | *0.0184* | *0.0232* | *0.0178* | *0.0193* |
| FGLR r = 6 | 0.1651 | 0.1665 | 0.1894 | 0.2659 | 0.1651 | 0.1660 |
| | *0.0204* | *0.0232* | *0.0261* | *0.0325* | *0.0204* | *0.0210* |
| FGLR r = 12 | 0.1494 | 0.1631 | 0.1951 | 0.3149 | **0.1494** | 0.1546 |
| | *0.0205* | *0.0239* | *0.0271* | *0.0344* | *0.0205* | *0.0350* |
| FGLR r = IC2 | 0.1585 | 0.1657 | 0.1932 | 0.2914 | 0.1764 | 0.1862 |
| | *0.0200* | *0.0235* | *0.0300* | *0.0624* | *0.0732* | *0.0858* |
| *shocks* | | | | | | |
| FHLZ | 0.2376 | 0.2194 | 0.2326 | 0.3117 | **0.2141** | 0.2271 |
| | *0.0262* | *0.0248* | *0.0247* | *0.0296* | *0.0290* | *0.0245* |
| FGLR r = 6 | 0.2411 | 0.2504 | 0.2779 | 0.3714 | **0.2411** | 0.2421 |
| | *0.0250* | *0.0265* | *0.0279* | *0.0331* | *0.0250* | *0.0253* |
| FGLR r = 12 | 0.2621 | 0.3012 | 0.3581 | 0.5639 | 0.2621 | 0.2709 |
| | *0.0272* | *0.0318* | *0.0363* | *0.0446* | *0.0272* | *0.0508* |
| FGLR r = IC2 | 0.2488 | 0.2703 | 0.3107 | 0.4572 | 0.2825 | 0.3011 |
| | *0.0299* | *0.0416* | *0.0589* | *0.1412* | *0.1470* | *0.1704* |

**Table 3**: Average and standard deviation (slanted) of MSE across 200 artificial data set. DGP: FGLR. Monthly data.

| *irf* | $p=2$ | $p=4$ | $p=6$ | $p=12$ | $p=BIC$ | $p=HQC$ |
|---|---|---|---|---|---|---|
| FHLZ | 0.3003 | 0.2768 | 0.2797 | 0.3133 | 0.3012 | **0.2760** |
| | *0.0435* | *0.0383* | *0.0386* | *0.0338* | *0.0400* | *0.0364* |
| FGLR r = 8 | 0.2435 | 0.2417 | 0.2606 | 0.3274 | 0.2603 | 0.2408 |
| | *0.0919* | *0.0882* | *0.0914* | *0.0916* | *0.0955* | *0.0954* |
| FGLR r = 16 | **0.2156** | 0.2325 | 0.2649 | 0.3962 | 0.2562 | 0.2320 |
| | *0.0799* | *0.0837* | *0.0833* | *0.0795* | *0.0905* | *0.0893* |
| FGLR r = IC2 | 0.2273 | 0.2286 | 0.2523 | 0.3412 | 0.2632 | 0.2417 |
| | *0.0820* | *0.0723* | *0.0765* | *0.0893* | *0.0947* | *0.1006* |
| *shocks* | | | | | | |
| FHLZ | 0.1369 | 0.1358 | 0.1519 | 0.2351 | 0.1321 | **0.1316** |
| | *0.1004* | *0.0770* | *0.0773* | *0.0543* | *0.0891* | *0.0706* |
| FGLR r = 8 | 0.2014 | 0.2241 | 0.2639 | 0.3903 | 0.2120 | 0.2132 |
| | *0.2310* | *0.2200* | *0.2214* | *0.2128* | *0.2452* | *0.2371* |
| FGLR r = 16 | 0.1968 | 0.2734 | 0.3571 | 0.6486 | 0.1979 | 0.2022 |
| | *0.2011* | *0.2030* | *0.1930* | *0.1604* | *0.2373* | *0.2218* |
| FGLR r = IC2 | **0.1832** | 0.2169 | 0.2744 | 0.4607 | 0.2037 | 0.2232 |
| | *0.2058* | *0.1832* | *0.1895* | *0.2143* | *0.2426* | *0.2499* |

**Table 4**: Average and standard deviation (slanted) of MSE across 200 artificial data set. DGP: FHLZ. Monthly data.

| *irf* | $p = 2$ | $p = 4$ | $p = 6$ | $p = 12$ | $p = BIC$ | $p = HQC$ |
|---|---|---|---|---|---|---|
| FHLZ | 0.1226 | 0.1292 | 0.1394 | 0.1832 | 0.1228 | **0.1220** |
| | *0.0250* | *0.0243* | *0.0237* | *0.0219* | *0.0236* | *0.0214* |
| FGLR r = 8 | 0.2890 | 0.3131 | 0.3423 | 0.4040 | 0.3018 | 0.2949 |
| | *0.1064* | *0.0980* | *0.0978* | *0.0887* | *0.1129* | *0.1104* |
| FGLR r = 16 | **0.2564** | 0.2780 | 0.3148 | 0.4448 | 0.2619 | 0.2679 |
| | *0.0951* | *0.0842* | *0.0823* | *0.0718* | *0.1067* | *0.1134* |
| FGLR r = IC2 | 0.2652 | 0.2872 | 0.3208 | 0.4134 | 0.2826 | 0.2928 |
| | *0.0906* | *0.0912* | *0.0923* | *0.0849* | *0.0981* | *0.1178* |
| *shocks* | | | | | | |
| FHLZ | 0.4392 | 0.4222 | 0.4092 | 0.4474 | 0.4387 | **0.3773** |
| | *0.0503* | *0.0499* | *0.0469* | *0.0414* | *0.0694* | *0.0430* |
| FGLR r = 6 | **0.6340** | 0.6471 | 0.6723 | 0.7525 | 0.6435 | 0.6393 |
| | *0.1630* | *0.1491* | *0.1511* | *0.1414* | *0.1690* | *0.1678* |
| FGLR r = 12 | 0.6559 | 0.6742 | 0.7397 | 0.9383 | 0.6551 | 0.6603 |
| | *0.1808* | *0.1585* | *0.1594* | *0.1506* | *0.1885* | *0.1898* |
| FGLR r = IC2 | 0.6640 | 0.6765 | 0.7225 | 0.8468 | 0.6820 | 0.6935 |
| | *0.1759* | *0.1705* | *0.1677* | *0.1546* | *0.1792* | *0.1915* |

**Table 5**: Summary results: FHLZ/FGLR ratio

| | T1 irf | T2 irf | T3 irf | T4 irf | T1 sh | T2 sh | T3 sh | T4 sh |
|---|---|---|---|---|---|---|---|---|
| FHLZbest/FGLRbest | 0.997 | 0.793 | 1.280 | 0.476 | 0.729 | 0.877 | 0.718 | 0.595 |
| FHLZbic/FGLRic2bic | 0.916 | 0.671 | 1.144 | 0.435 | 0.649 | 0.758 | 0.649 | 0.643 |
| FHLZhqc/FGLRic2hqc | 0.893 | 0.687 | 1.142 | 0.417 | 0.619 | 0.754 | 0.590 | 0.544 |

The results obtained when the artificial panels are generated by $(\text{IRF}, \text{IC})_{\text{FHLZ}}$, i.e. a fairly large advantage of FHLZ over FGLR, Tables 2 and 4, provides evidence that $\text{IRF}_{\text{FHLZ}}$, the impulse-response functions estimated by FHLZ with the actual panels, do not fulfill a finite-dimension restriction, either strictly or approximately.

When the artificial panels are generated using $(\text{IRF}, \text{IC})_{\text{FGLR}}$, FHLZ is, apart from very special cases, a less parsimonious representation as compared to FGLR. In particular, if $r \geq 10$ in (1.3), the length of the $(q + 1)$-dimensional VAR's in the corresponding FHLZ representation is larger than 10 (this can be seen by elementary algebraic manipulation). On the other hand, the length of the $(q + 1)$-dimensional VAR's determined by the BIC or the HQC criterion, when estimating with FHLZ the artificial panels generated by $(\text{IRF}, \text{IC})_{\text{FGLR}}$, is either 1 or 2. Thus a quite parsimonious FHLZ specification is sufficient to account for the dynamics in the panel, this providing a possible explanation for the unexpected better performance of FHLZ even in this case.

# 5 Conclusions

An estimate of the common-components spectral density matrix $\hat{\mathbf{\Sigma}}^{\chi}$ can be easily obtained using the frequency-domain principal components of the observable variables $x_{it}$. The central idea of the present paper is that because $\hat{\mathbf{\Sigma}}^{\chi}$ has large dimension but rank $q$, a factorization of $\hat{\mathbf{\Sigma}}^{\chi}$ can be obtained piece-wise. Precisely, the factorization of $\hat{\mathbf{\Sigma}}^{\chi}$ only requires the factorization of subvectors of $\boldsymbol{\chi}_t$ of length $q+1$. Under our assumption of rational spectral density for the common components, this implies that the number of parameters to estimate grows at pace $n$, not $n^2$.

The rational spectral density assumption has also the important consequences that $\boldsymbol{\chi}_t$ has a finite autoregressive representation and that the dynamic factor model can be transformed into the static model $\mathbf{z}_t = \mathbf{R}\mathbf{v}_t + \boldsymbol{\phi}_t$, where $\mathbf{z}_t = \mathbf{A}(L)\mathbf{x}_t$. Estimators for $\mathbf{A}(L)$, $\mathbf{R}$ and $\mathbf{v}_t$ can be constructed starting with an estimate of the spectral density of the $x$'s.

Although we make use of a parametric structure for the common components, we do not assume that our dynamic factor model has a static representation like (1.3).

# References


[1] Altissimo, F., R. Cristadoro, M. Forni, M. Lippi and G. Veronese, (2010). New Eurocoin: Tracking Economic Growth in Real Time, *The Review of Economics and Statistics,* **92**, 1024-1034.

[2] Amengual, D. and M. 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.

[3] Anderson, B. and M. Deistler (2008a). Properties of Zero-free Transfer Function Matrices, *SICE Journal of Control, Measurement and System Integration* **1**, 1-9.

[4] Anderson, B. and M. Deistler (2008b). Generalized Linear Dynamic Factor Models—A Structure Theory, *2008 IEEE Conference on Decision and Control.*

[5] Bai, J. (2003). Inferential Theory for Factor Models of Large Dimensions, *Econometrica* **71**, 135-171.

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

[7] Bai, J. and S. Ng, (2007). Determining the Number of Primitive Shocks in Factor Models,*Journal of Business and Economic Statistics,***25,** 52-60.

[8] Bernanke, B. S., J. Boivin and P. Eliasz (2005). Measuring Monetary Policy: A Factor Augmented Autoregressive (FAVAR) Approach, *The Quarterly Journal of Economics* 120, 387-422.

[9] Chamberlain, G. (1983). Funds, factors, and diversification in arbitrage pricing models. *Econometrica* **51**, 1281-1304.

[10] Chamberlain, G. and M. Rothschild (1983). Arbitrage, factor structure and mean-variance analysis in large asset markets, *Econometrica* **51**, 1305-1324.

[11] Christiano, L. and J. Eichenbaum, M., Evans, C. L., (1999), Monetary policy shocks: What have we learned and to what end? *Handbook of Macroeconomics*, Chapter 02, 65-148, J. B. Taylor & M. Woodford (ed.).

[12] Dunford, N. and J. T. Schwartz, (1988). *Linear Operators*, Part I. New York: Wiley.

[13] Forni, M. and L. Gambetti, (2010a). Macroeconomic Shocks and the Business Cycle: A Structural Factor Model Approach, CEPR Discussion Paper Series no. 7692.

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

[15] 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.

[16] Forni, M., M. Hallin, M. Lippi and L. Reichlin (2005). The generalized factor model: one-sided estimation and forecasting, *Journal of the American Statistical Association* **100**, 830-40.

[17] Forni, M., D. Giannone, M. Lippi and L. Reichlin (2009). Opening the Black Box: Structural Factor Models with Large Cross-Sections, *Econometric Theory* 25, 1319-1347.

[18] Forni, M. and M. Lippi (2001). The generalized dynamic factor model: representation theory, *Econometric Theory* **17**, 1113-41.

[19] Geweke, J. (1977). The dynamic factor analysis of economic time series, in D.J. Aigner and A.S. Goldberger, Eds., *Latent Variables in Socio-Economic Models*. Amsterdam: North Holland.

[20] 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.

[21] Hannan, E.J. (1970). *Multiple Time Series,* New York: Wiley.

[22] Onatski, A. (2009). Testing Hypotheses About the Number of Factors in Large Factor Models, *Econometrica* 77, 1447-1479.

[23] Pourahmadi, M. (2001). *Foundations of Time Series Analysis and Prediction Theory*, New York: Wiley.

[24] Rozanov, Yu. (1967). *Stationary Random processes*. San Francisco: Holden Day.

[25] Sargent, T.J., C.A. Sims (1977). Business cycle modelling without pretending to have too much *a priori* economic theory, in C.A. Sims, Ed., *New Methods in Business Research*, Minneapolis: Federal Reserve Bank of Minneapolis.

[26] Stock, J.H. and M.W. Watson (2002a). Macroeconomic Forecasting Using Diffusion Indexes, *Journal of Business and Economic Statistics* 20, 147-162.

[27] Stock, J.H. and M.W. Watson (2002b). Forecasting Using Principal Components from a Large Number of Predictors, *Journal of the American Statistical Association* 97, 1167-1179.

[28] Stock, J.H., and M.W. Watson (2005). Implications of Dynamic Factor Models for VAR Analysis, NBER Working Papers no. 11467.

[29] van der Waerden, B.L. (1953). *Modern Algebra*, Vol. I. New York: Frederick Ungar.

# Appendix

## A Proof of Proposition 1

Consider first the case $s_2 = 0$, so that $\mathbf{y}_t$ is a moving average. Setting $s = s_1$, we can write:

$$\mathbf{y}_t = \mathbf{C}_0\mathbf{v}_t + \mathbf{C}_1\mathbf{v}_{t-1} + \cdots + \mathbf{C}_s\mathbf{v}_{t-s} = \mathbf{C}(L)\mathbf{v}_t, \tag{A.1}$$

Consider the stack

$$\begin{pmatrix} \mathbf{y}_{t-1} \\ \mathbf{y}_{t-2} \\ \vdots \\ \mathbf{y}_{t-S} \end{pmatrix} = \mathcal{C}_S \begin{pmatrix} \mathbf{v}_{t-1} \\ \mathbf{v}_{t-2} \\ \vdots \\ \mathbf{v}_{t-S-s} \end{pmatrix} \tag{A.2}$$

where

$$\mathcal{C}_S = \begin{pmatrix} \mathbf{C}_0 & \mathbf{C}_1 & \cdots & \mathbf{C}_s & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_0 & \cdots & \mathbf{C}_{s-1} & \mathbf{C}_s & \cdots & \mathbf{0} \\ \vdots & & & & & \ddots & \\ \mathbf{0} & \mathbf{0} & \cdots & & & \cdots & \mathbf{C}_s \end{pmatrix}. \tag{A.3}$$

The matrix $\mathcal{C}_S$ has $(q+1)S$ rows and $qS + qs$ columns. Setting $S = sq$, $\mathcal{C}_S$ is square. If $\mathcal{C}_S$ is non singular, then

$$\begin{pmatrix} \mathbf{v}_{t-1} \\ \mathbf{v}_{t-2} \\ \vdots \\ \mathbf{v}_{t-S-s} \end{pmatrix} = \mathcal{C}_S^{-1} \begin{pmatrix} \mathbf{y}_{t-1} \\ \mathbf{y}_{t-2} \\ \vdots \\ \mathbf{y}_{t-S-s} \end{pmatrix}.$$

Substituting into (A.1),

$$\mathbf{y}_t = \mathbf{A}_1\mathbf{y}_{t-1} + \cdots + \mathbf{A}_S\mathbf{y}_{t-s} + \mathbf{C}_0\mathbf{v}_t. \tag{A.4}$$

Obviously (A.4) is the orthogonal projection of $\mathbf{y}_t$ on its past (as $\mathbf{v}_t$ is orthogonal to $\mathbf{y}_{t-k}$ for $k > 0$). Moreover, non singularity of $\mathcal{C}_S$ implies that the entries of te left hand side of (A.2) are are linearly independent. As a consequence, (A.4) is the unique AR representation of $\mathbf{y}_t$ of order less or equal to $S$, up to a transformation of $\mathbf{v}_t$ by and orthogonal matrix.

It remains to prove that $\mathcal{C}_S$ is non singular for generic values of the entries of the matrices $C_j$. Note that the determinant of $\mathcal{C}_S$ is a polynomial in the parameters, and is therefore either zero for all parameters' values or generically non zero. Thus if we find a particular value of the parameters for which $\det \mathcal{C}_S \neq 0$ we can conclude that $\mathcal{C}_S$ is generically non singular.

Suppose that for a vector $\boldsymbol{\alpha} \neq 0$

$$\boldsymbol{\alpha}\mathcal{C}_S = 0. \tag{A.5}$$

The vector $\boldsymbol{\alpha}$ can be written as

$$\boldsymbol{\alpha} = (\alpha_{01} \ \cdots \ \alpha_{0,q+1}; \ \ \alpha_{11} \ \cdots \ \alpha_{1,q+1}; \ \ \cdots; \ \ \alpha_{S-1,1} \ \cdots \ \alpha_{S-1,q+1}).$$

Defining $\boldsymbol{\alpha}_i = (\alpha_{i1} \ \cdots \ \alpha_{i,q+1})$, (A.5) can be rewritten as

$$\begin{aligned} \boldsymbol{\alpha}_0(\mathbf{C}_0 + \mathbf{C}_1 L + \cdots + \mathbf{C}_s L^s) &+ \boldsymbol{\alpha}_1 L(\mathbf{C}_0 + \mathbf{C}_1 L + \cdots + \mathbf{C}_s L^s) \\ &+ \cdots + \boldsymbol{\alpha}_{S-1} L^{S-1}(\mathbf{C}_0 + \mathbf{C}_1 L + \cdots + \mathbf{C}_s L^s) \\ = (\boldsymbol{\alpha}_0 + \boldsymbol{\alpha}_1 L + \cdots &+ \boldsymbol{\alpha}_{S-1} L^{S-1})(\mathbf{C}_0 + \mathbf{C}_1 L + \cdots + \mathbf{C}_s L^s) = \mathbf{0}, \end{aligned}$$

that is

$$(\beta_1(L) \ \beta_2(L) \ \cdots \ \beta_{q+1}(L))\mathbf{C}(L) = 0, \tag{A.6}$$

where

$$\beta_j(L) = \alpha_{0j} + \alpha_{1j} L + \cdots \ + \alpha_{S-1,j} L^{S-1}$$

Thus (A.5) is equivalent to the existence of $q+1$ scalar polynomials $\beta_j(L)$, of degree $S-1$, such that $\beta_j(L) \neq 0$ for some $j$ and (A.6) holds.

Let us now construct a point in the parameter space as follows. Let $d_i(L)$, for $i = 1, \ldots, q$, be polynomials of degree $s$, such that $d_i(L)$ and $d_j(L)$ have no common roots for all $i \neq j$. Then let

$$\mathbf{C}(L) = \begin{pmatrix} d_1(L) & 0 & \cdots & 0 \\ 0 & d_2(L) & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & d_q(L) \\ 1 & 1 & \cdots & 1 \end{pmatrix}. \tag{A.7}$$

If there exist $\beta_i(L)$, $i = 1, \ldots, q+1$, not all zero, such that (A.6) holds, then for $i \leq q$,

$$d_i(L)\beta_i(L) = -\beta_{q+1}(L).$$

Therefore $\beta_{q+1}(L)$ has all the roots of the polynomials $d_i(L)$ for $i \leq q$. But then, given our assumption on the roots of the polynomials $d_i(L)$, either $\beta_i(L) = 0$ for all $i = 1, \ldots, q+1$, or the degree of $\beta_{q+1}(L)$ is at least $qs = S$. Thus $\det \mathcal{C}_S \neq 0$ in this case and therefore generically.

Let us now turn to the general rational case:

$$\mathbf{y}_t = \mathbf{E}(L)\mathbf{v}_t, \tag{A.8}$$

where

$$e_{if}(L) = \frac{c_{if}(L)}{d_{if}(L)} = \frac{c_{if,0} + c_{if,1} L + \cdots + c_{if,s_1} L^{s_1}}{1 + d_{if,1} L + \cdots + d_{if,s_2} L^{s_2}}.$$

Rewrite (A.8) as

$$\begin{pmatrix} h_1(L) & 0 & \cdots & 0 \\ 0 & h_2(L) & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & h_{q+1}(L) \end{pmatrix} \mathbf{y}_t = \mathbf{G}(L)\mathbf{v}_t,$$

where

$$\begin{aligned} h_i(L) &= \prod_{f=1}^{q} d_{if}(L) \\ g_{if}(L) &= c_{if}(L) h_i(L)/d_{if}(L) \end{aligned}$$

For generic values of the parameters the degrees of $h_j(L)$ and $g_{ik}(L)$ are $qs_2$ and $s_1 + (q-1)s_2$ respectively. Now consider the moving average $\mathbf{G}(L)\mathbf{v}_t$. Select a point in the parameter space such that:
(i) $c_{if}(L) = 0$ if $i \neq f$ and $i \leq q$, so that $g_{if}(L) = 0$ if $i \neq f$ and $i \leq q$,
(ii) $g_{ii}(L)$ has degree $s_1 + (q-1)s_2$ for $i \leq q$,
(iii) $g_{ii}(L)$ and $g_{jj}(L)$ have no roots in common for $i \neq j$,
(iv) $g_{q+1,f}(L) = 1$ for $j = 1, \ldots, q$.
This reproduces the situation in (A.7). As the entries of $\mathbf{G}(L)$ are polynomial functions of $p \in \Pi^{q+1}$, for generic values of the parameters $\mathbf{G}(L)\mathbf{v}_t$ has an autoregressive representation of order not greater than $[s_1 + (q-1)s_2]q$. This implies that, for generic values of the parameters, $\mathbf{y}_t$ has an autoregressive representation of order not greater than

$$S = qs_2 + [s_1 + (q-1)s_2]q = qs_1 + q^2 s_2.$$

To prove uniqueness we now show that generically the entries of

$$\mathbf{y}_{t-1}, \ \mathbf{y}_{t-2}, \ \ldots, \ \mathbf{y}_{t-S}$$

are linearly independent, which is equivalent to proving that if $\beta_j(L)$, $j = 1, \ldots, q+1$, are polynomials in $L$ and

$$(\beta_1(L) \ \beta_2(L) \ \cdots \ \beta_{q+1}(L))\mathbf{E}(L) = 0, \tag{A.9}$$

then for generic values of the parameters either $\beta_j(L) = 0$ for all $j$, or the degree of $\beta_j(L)$ is greater than $S - 1$ for some $j$.
(I) Let $\mathbf{E}_q(L)$ be the square submatrix of $\mathbf{E}(L)$ obtained by dropping the last row of $\mathbf{E}(L)$. Rationality of the entries of $\mathbf{E}(L)$ implies that either $\det(\mathbf{E}_q(L)) = 0$ for all $z \in \mathbb{C}$ or for a finite subset of $\mathbb{C}$. On the other hand, if $\det(\mathbf{E}_q(z)) = 0$ for all $z \in \mathbb{C}$, then the parameters fulfill a set of polynomial equations. Obviously there exist parameters such that

$$\det(\mathbf{E}_q(L)) = 0 \text{ for a finite subset of } \mathbb{C},$$

i.e. such that $\mathbf{E}_q(L)$ is a non-zero rational function. Thus for $\Pi^{q+1} - \mathcal{M}_1$, where $\mathcal{M}_1$ is a nowhere dense subset of $\Pi^{q+1}$, $\mathbf{E}_q(L)$ is a non-zero rational function, so that $[\mathbf{E}_q(L)]^{-1}$ makes sense. As a consequence, in $\Pi^{q+1} - \mathcal{M}_1$ the system of equations

$$(\rho_1(L) \ \rho_2(L) \ \cdots \ \rho_q(L))\mathbf{E}_q(L) = (e_{q+1,1}(L) \ e_{q+1,2}(L) \ \cdots \ e_{q+1,q}(L)),$$

in the unknown rational functions $\rho_j(L)$, has the unique solution

$$(\tau_1(L)\ \tau_2(L)\ \cdots\ \tau_q(L)) = (e_{q+1,1}(L)\ e_{q+1,2}(L)\ \cdots\ e_{q+1,q}(L))[\mathbf{E}_q(L)]^{-1}.$$

(II) There exists $\mathcal{M}_2$, nowhere dense in $\Pi^{q+1}$, such that in $\Pi^{q+1} - \mathcal{M}_2$

$$\det \mathbf{E}_q(L) = \frac{h(L)}{\prod_{i,j=1}^{q} d_{ij}(L)},$$

where

$$\text{Degree}(h(L)) = qs_1 + (q^2 - q)s_2.$$

(III) There exists $\mathcal{M}_3$, nowhere dense in $\Pi^{q+1}$, such that in $\Pi^{q+1} - \mathcal{M}_3$ the entry $(i, j)$ of the adjoint matrix of $\mathbf{E}_q(L)$ can be written as

$$\frac{h_{ij}(L)}{\prod_{\substack{h,k=1,\dots q \\ h\neq j,\ k\neq i}} d_{hk}(L)},$$

the degrees of numerator and denominator being

$$(q-1)s_1 + [(q-1)^2 - (q-1)]s_2 \quad \text{and} \quad (q-1)^2 s_2$$

respectively.
(IV) There exists $\mathcal{M}_4$, nowhere dense in $\Pi^{q+1}$, such that in $\Pi^{q+1} - \mathcal{M}_4$ the entries of $[\mathbf{E}_q(L)]^{-1}$ can be written as

$$\frac{h_{ij}(L) \prod_{\substack{h,j=1,\dots,q \\ h=j \text{ or } k=i}} d_{hk}(L)}{h(L)} = \frac{\tilde{h}_{ij}(L)}{h(L)},$$

where the degrees of numerator and denominator are

$$(q-1)s_1 + (q^2 - (q-1))s_2 \quad \text{and} \quad qs_1 + (q^2 - q)s_2$$

respectively.
(V) There exists $\mathcal{M}_5$, nowhere dense in $\Pi^{q+1}$, such that in $\Pi^{q+1} - \mathcal{M}_5$

$$\tau_k(L) = \sum_{i=1}^{q} \frac{c_{q+1,i}(L)\tilde{h}_{ik}(L)}{d_{q+1,i}(L)h(L)} = \frac{\sum_{i=1}^{q} c_{q+1,i}(L)\tilde{h}_{ik}(L) \prod_{\substack{j=1,\dots,q \\ j\neq i}} d_{q+1,j}(L)}{h(L)\prod_{i=1}^{q} d_{q+1,i}(L)} = \frac{\nu_k(L)}{\delta(L)},$$

where both $\nu_k(L)$ and $\delta(L)$ are polynomials of degree

$$S = qs_1 + q^2 s_2.$$

(VI) Moreover, for generic values of the parameters, $\nu_k(L)$ and $\delta(L)$ have no roots in common. To show this, let us recall that $\nu_k(L)$ and $\delta(L)$ have roots in common, when they are both of maximum degree, if and only if their resultant vanishes. The resultant is a polynomial of the coefficients of

$$\nu_k(L) = \nu_{k,S}L^S + \nu_{k,S-1}L^{S-1} + \cdots + \nu_{k,0} \quad \text{and} \quad \delta(L) = \delta_S L^S + \delta_{S-1}L^{S-1} + \cdots + \delta_0,$$

homogeneous of degree $S$ in the coefficients $\nu_{k,j}$ and $\delta_j$, containing the term

$$\nu_{k,S}^S \delta_0^S$$

(see van der Waerden, 1953, pp. 83-5). All other terms contain $\nu_{k,S}^{S-h}$, with $0 < h \leq S$. We have:

$$\begin{aligned} \nu_{k,S}^S \delta_0^S &= \left[ \sum_{i=1}^{q} c_{q+1,i,s_1} \tilde{h}_{ik,g} \prod_{\substack{j=1,\dots,q \\ j \neq i}} d_{q+1,j,s_2} \right]^S h(0)^S \\ &= c_{q+1,1,s_1}^S \left[ \tilde{h}_{1k,g}^S \prod_{j=2,\dots,q} d_{q+1,j,s_2}^S h(0)^S \right] + \cdots , \end{aligned} \tag{A.10}$$

where $g = (q-1)s_1 + (q^2 - (q-1))s_2$. Note that $h(L)$ and $\tilde{h}_{ik}(L)$ do not contain the parameters $c_{q+1,i,h}$. As a consequence, all other terms in (A.10) and in the discriminant of $\nu_k(L)$ and $\delta(L)$ contain $c_{q+1,i,s_1}^{S-h}$ with $0 < h \leq S$. Moreover, the term in square brackets in (A.10) is generically non zero. Thus the discriminant of $\nu_k(L)$ and $\delta(L)$ does not vanish everywhere in $\Pi^{q+1}$ and therefore vanishes only on a nowhere dense subset. In conclusion, there exists $\mathcal{M}_6$, nowhere dense in $\Pi^{q+1}$, such that in $\Pi^{q+1} - \mathcal{M}_6$ the discriminant of $\delta(L)$ and $\nu_k(L)$ does not vanish, for all $k = 1, 2, \dots, q$.
(VII) Lastly, if the polynomials $\beta_k(L)$ fulfill (A.9) then

$$\tau_k(L) = -\frac{\beta_k(L)}{\beta_{q+1}(L)}.$$

The results above imply that the degree of the polynomials $\beta_j(L)$ is at least $S$ for parameters belonging to $\Pi^{q+1} - \cup_{k=1}^{6} \mathcal{M}_k$. Q.E.D.

# B Discussion of Assumptions A* and A.7

Let us show that Assumption A* does not imply Assumption A.7. We construct a sequence $z_{iT}$ such that (i) $z_{iT}$ is $O_P(\rho_T)$ uniformly in $i$, (ii) $n^{-1}\sum_{i=1}^{n} z_{iT}$ is not $O_P(1)$.

Consider the interval $I = [0\ 1]$ with the uniform probability density and the following sequence of stochastic variables: $x_1(\omega) = d_1$ for all $\omega \in I$,

$$x_2(\omega) = \begin{cases} d_2 \text{ if } \omega < 1/2 \\ 0 \text{ if } \omega \geq 1/2 \end{cases} \quad x_3(\omega) = \begin{cases} 0 \text{ if } \omega < 1/2 \\ d_2 \text{ if } \omega \geq 1/2 \end{cases} \quad x_4(\omega) = \begin{cases} d_3 \text{ if } \omega < 1/4 \\ 0 \text{ if } \omega \geq 1/4 \end{cases} \quad \cdots$$

where $d_k$ is a sequence of positive real numbers. Formally, any integer $s > 1$ can be represented uniquely as

$$s = 1 + 2 + \cdots + 2^{s_1} + s_2,$$

where $0 < s_2 \leq 2^{s_1+1}$. Define: $x_1(\omega) = d_1$ for all $\omega \in I$, and, for $s > 1$,

$$x_s(\omega) = \begin{cases} d_{s_1+2} \text{ if } (s_2 - 1)2^{-s_1-1} < \omega \leq s_2 2^{-s_1-1} \\ 0 \text{ otherwise} \end{cases}$$

A special feature of $x_s$ is that $x_s = O_P(\rho_s)$ for all sequences $\rho_s$, independently of the sequence $d_k$. However, consider the average $y_S = S^{-1} \sum_{s=1}^{S} x_s$ and the subsequence $y_{S_k}$, for $S_k = 1 + 2 + \cdots + 2^k$. We have:

$$y_{S_k} = \frac{d_1 + d_2 + \cdots + d_k}{1 + 2 + \cdots + 2^k},$$

for $\omega$ a.e. in $I$, which is divergent if, for example, $d_k = d^k$, with $d > 2$. Thus, if $d_k = d^k$ and $d > 2$, $y_S$ is not $O_P(1)$ as $S \to \infty$.

Now assume $d_k = d^k$ with $d > 2$, and define $z_{iT} = x_{i+T-1}$. Of course, for $T \to \infty$, $z_{iT} = O_P(\rho_T)$, uniformly in $i$, i.e. given $\epsilon$, $\eta_\epsilon$ is independent of $i$. Then define $w_{nT} = n^{-1} \sum_{i=1}^{n} z_{iT}$ and observe that $w_{nT} = y_n - O_P(n^{-1})$, which is not $O_P(1)$ as $n \to \infty$. Thus $w_{nT}$ is not $O_P(1)$ as $n \to \infty$ and $T \to \infty$.

# C Proof of Proposition 2

The proof below closely follows Forni et al. (2009). If $\mathbf{A}$ is a complex $s \times s$ Hermitian matrix we denote by $\mu_j(\mathbf{A})$, $j = 1, 2, \ldots, s$, the (real) eigenvalues of $\mathbf{A}$ in decreasing order. Given an $s_1 \times s_2$ matrix $\mathbf{B}$, $\|\mathbf{B}\|$ denotes the spectral norm of $\mathbf{B}$: $\|\mathbf{B}\| = \sqrt{\mu_1(\tilde{\mathbf{B}}\mathbf{B})}$, which is the euclidean norm if $\mathbf{B}$ is a row matrix. We recall that if $\mathbf{A}$ is $s_1 \times s_2$ and $\mathbf{B}$ is $s_2 \times s_3$, then

$$\|\mathbf{AB}\| \leq \|\mathbf{A}\| \|\mathbf{B}\|. \tag{C.1}$$

We will make use of the following inequality. Let $\mathbf{A}$ and $\mathbf{B}$ be two $s \times s$ Hermitian matrices:

$$|\mu_j(\mathbf{A} + \mathbf{B}) - \mu_j(\mathbf{A})| \leq \sqrt{\mu_1(\mathbf{B}\tilde{\mathbf{B}})} = \|\mathbf{B}\|, \quad j = 1, \ldots, s. \tag{C.2}$$

This, also known as Weyl's inequality, is an obvious consequence of Lancaster and Tismenetsky (1985), p. 301 (see also Forni and Lippi (2001), Fact M and Forni et al., 2009, Appendix).

The proof of Proposition 2 is broken into some intermediate propositions. Let $a_1 < a_2 < \cdots < a_s$ be integers and $\mathbf{M} = \{a_1, a_2, \cdots, a_s\}$. Denote by $\mathcal{S}_\mathbf{M}$ the $n \times s$

matrix with 1 in entries $(a_j, j)$ and zero elsewhere. As most of the arguments below depend on equalities and inequalities that hold for all $\theta \in [-\pi\ \pi]$, the notation has been simplified by dropping $\theta_h$. Moreover, properties holding for $\max_{|h|\leq B_T} f(\theta_h)$, where $f$ is some function of $\theta$, are often phrased as holding for $f$ *uniformly in* $\theta$.

Lemmas 3 through 6 are proved under Assumptions A.1 through A.8.

**Lemma 3** For $T \to \infty$ and $n \to \infty$,
(i) $\max_{|h|\leq B_T} \frac{1}{n}\|\hat{\mathbf{\Sigma}}^x - \mathbf{\Sigma}^x\| = O_P\left(\frac{1}{\sqrt{\rho_T}}\right)$.
(ii) Given $s$, $\max_{|h|\leq B_T} \frac{1}{\sqrt{n}}\|\mathcal{S}'_{\mathbf{M}}\left(\hat{\mathbf{\Sigma}}^x - \mathbf{\Sigma}^x\right)\| = O_P\left(\frac{1}{\sqrt{\rho_T}}\right)$, uniformly in $\mathbf{M}$.
(iii) $\max_{|h|\leq B_T} \frac{1}{n}\|\hat{\mathbf{\Sigma}}^x - \mathbf{\Sigma}^\chi\| = O_P\left(\max\left(\frac{1}{n}, \frac{1}{\sqrt{\rho_T}}\right)\right)$.
(iv) Given $s$, $\max_{|h|\leq B_T} \frac{1}{\sqrt{n}}\|\mathcal{S}'_{\mathbf{M}}\left(\hat{\mathbf{\Sigma}}^x - \mathbf{\Sigma}^\chi\right)\| = O_P\left(\max\left(\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{\rho_T}}\right)\right)$, uniformly in $\mathbf{M}$.

Proof. We have

$$\mu_1\left((\hat{\mathbf{\Sigma}}^x - \mathbf{\Sigma}^x)(\tilde{\hat{\mathbf{\Sigma}}}^x - \tilde{\mathbf{\Sigma}}^x)\right) \leq \text{trace}\left((\hat{\mathbf{\Sigma}}^x - \mathbf{\Sigma}^x)(\tilde{\hat{\mathbf{\Sigma}}}^x - \tilde{\mathbf{\Sigma}}^x)\right) = \sum_{i=1}^n\sum_{j=1}^n |\hat{\sigma}^x_{ij} - \sigma^x_{ij}|^2.$$

Using Assumption A.7,

$$\frac{1}{n^2}\max_{|h|\leq B_T}\sum_{i=1}^n\sum_{j=1}^n |\hat{\sigma}^x_{ij} - \sigma^x_{ij}|^2 \leq \frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n \max_{|h|\leq B_T}|\hat{\sigma}^x_{ij} - \sigma^x_{ij}|^2 = O_P\left(\frac{1}{\rho_T}\right).$$

Statement (i) follows. Similarly, we have

$$\text{trace}\left(\mathcal{S}'_{\mathbf{M}}\left(\hat{\mathbf{\Sigma}}^x - \mathbf{\Sigma}^x\right)\left(\tilde{\hat{\mathbf{\Sigma}}}^x - \tilde{\mathbf{\Sigma}}^x\right)\mathcal{S}_{\mathbf{M}}\right) = \sum_{i\in\mathbf{M}}\sum_{j=1}^n |\hat{\sigma}^x_{ij} - \sigma^x_{ij}|^2.$$

Statement (ii) follows from Assumption A.7. As for (iii), observe that $\hat{\mathbf{\Sigma}}^x - \mathbf{\Sigma}^\chi = \hat{\mathbf{\Sigma}}^x - \mathbf{\Sigma}^x + \mathbf{\Sigma}^\xi$ (orthogonality of the common and idiosyncratic components at all leads and lags), so that $\frac{1}{n}\|\hat{\mathbf{\Sigma}}^x - \mathbf{\Sigma}^\chi\| \leq \frac{1}{n}\|\hat{\mathbf{\Sigma}}^x - \mathbf{\Sigma}^x\| + \frac{1}{n}\|\mathbf{\Sigma}^\xi\|$ (triangle inequality for matrix norm). Using (i) and boundedness of $\lambda^\xi_1$, the statement follows from $\frac{1}{n}\|\mathbf{\Sigma}^\xi\| = \frac{1}{n}\lambda^\xi_1 = O\left(\frac{1}{n}\right)$. Statement (iv) is obtained in a similar way, using (ii) instead of (i) and the upper bound $\frac{1}{\sqrt{n}}\lambda^\xi_1$ instead of $\frac{1}{n}\lambda^\xi_1$. Uniformity with respect to $\mathbf{M}$ in (ii) and (iv) follows form Assumption A.7. QED

**Lemma 4**
(i) $\max_{|h|\leq B_T}\left|\frac{\hat{\lambda}^x_f}{n} - \frac{\lambda^\chi_f}{n}\right| = O_P\left(\max\left(\frac{1}{n}, \frac{1}{\sqrt{\rho_T}}\right)\right)$ for $f = 1, 2, \ldots, q$.
(ii) Define:

$$\mathbf{G}^\chi = \begin{cases}\mathbf{I}_q \text{ if } \lambda^\chi_q = 0\\ \left(\frac{\mathbf{\Lambda}^\chi}{n}\right)^{-1} \text{ otherwise}\end{cases} \qquad \hat{\mathbf{G}}^x = \begin{cases}\mathbf{I}_q \text{ if } \hat{\lambda}^x_q = 0\\ \left(\frac{\hat{\mathbf{\Lambda}}^x}{n}\right)^{-1} \text{ otherwise.}\end{cases}$$

Then, for $n \to \infty$,

$$\max_{|h|\leq B_T} \left\| \frac{\mathbf{\Lambda}^\chi}{n} \right\| \quad \text{and} \quad \max_{|h|\leq B_T} \|\mathbf{G}^\chi\| \;\; \text{are} \;\; O(1).$$

For $T \to \infty$ and $n \to \infty$,

$$\max_{|h|\leq B_T} \left\| \frac{\hat{\mathbf{\Lambda}}^x}{n} \right\| \quad \text{and} \quad \max_{|h|\leq B_T} \|\hat{\mathbf{G}}^x\| \;\; \text{are} \;\; O_P(1).$$

(iii) As $T \to \infty$ and $n \to \infty$,

$$\max_{|h|\leq B_T} \left\| \frac{\hat{\mathbf{\Lambda}}^x}{n} \hat{\mathbf{G}}^x - \mathbf{I}_q \right\| = O_P\left( \max\left( \frac{1}{n}, \frac{1}{\sqrt{\rho_T}} \right) \right)$$

Proof. Setting $\mathbf{A} = \mathbf{\Sigma}^\chi$, $\mathbf{B} = \hat{\mathbf{\Sigma}}^x - \mathbf{\Sigma}^\chi$ and applying (C.2) we get $\frac{1}{n}|\hat{\lambda}_f^x - \lambda_f^\chi| \leq \frac{1}{n}\|\hat{\mathbf{\Sigma}}^x - \mathbf{\Sigma}^\chi\|$. Statement (i) follows from Lemma 3 (iii). As for statement (ii), boundedness of $\|\frac{\mathbf{\Lambda}^\chi}{n}\|$ and $\|\mathbf{G}^\chi\|$, uniformly in $\theta$, is an obvious consequence of Assumption A.8. Boundedness in probability of $\|\frac{\hat{\mathbf{\Lambda}}^x}{n}\|$ and $\|\hat{\mathbf{G}}^x\|$, uniformly in $\theta$, follows from statement (i). Statement (iii) follows from (i) and Assumption A.8. Q.E.D.

**Lemma 5** As $T \to \infty$ and $n \to \infty$,
(i) $\max_{|h|\leq B_T} \|\tilde{\mathbf{P}}^\chi \hat{\mathbf{P}}^x \frac{\hat{\mathbf{\Lambda}}^x}{n} - \frac{\mathbf{\Lambda}^\chi}{n} \tilde{\mathbf{P}}^\chi \hat{\mathbf{P}}^x\| = O_P\left(\max\left(\frac{1}{n}, \frac{1}{\sqrt{\rho_T}}\right)\right)$.
(ii) $\max_{|h|\leq B_T} \|\tilde{\hat{\mathbf{P}}}^x \mathbf{P}^\chi \tilde{\mathbf{P}}^\chi \hat{\mathbf{P}}^x - \mathbf{I}_q\| = O_P\left(\max\left(\frac{1}{n}, \frac{1}{\sqrt{\rho_T}}\right)\right)$.
(iii) There exist diagonal complex orthogonal matrices

$$\hat{\mathbf{W}}_q = \operatorname{diag}(\hat{w}_1 \; \hat{w}_2 \; \cdots \; \hat{w}_q), \quad |w_j|^2 = 1, \; j = 1, \ldots, q,$$

depending on $n$ and $T$, such that

$$\max_{|h|\leq B_T} \|\tilde{\hat{\mathbf{P}}}^x \mathbf{P}^\chi - \hat{\mathbf{W}}_q\| = O_P\left( \max\left( \frac{1}{n}, \frac{1}{\sqrt{\rho_T}} \right) \right).$$

Proof. We have $\|\tilde{\mathbf{P}}^\chi \hat{\mathbf{P}}^x \frac{\hat{\mathbf{\Lambda}}^x}{n} - \frac{\mathbf{\Lambda}^\chi}{n} \tilde{\mathbf{P}}^\chi \hat{\mathbf{P}}^x\| = \|\frac{1}{n}\tilde{\mathbf{P}}^\chi \left(\hat{\mathbf{\Sigma}}^x - \mathbf{\Sigma}^\chi\right) \hat{\mathbf{P}}^x\| \leq \frac{1}{n}\|\hat{\mathbf{\Sigma}}^x - \mathbf{\Sigma}^\chi\|$, by (C.1). Statement (i) then follows from Lemma 3 (iii). As for (ii), set
$a = \tilde{\hat{\mathbf{P}}}^x \mathbf{P}^\chi \tilde{\mathbf{P}}^\chi \hat{\mathbf{P}}^x$,
$b = [\tilde{\hat{\mathbf{P}}}^x \mathbf{P}^\chi \tilde{\mathbf{P}}^\chi \hat{\mathbf{P}}^x] \frac{\hat{\mathbf{\Lambda}}^x}{n} \hat{\mathbf{G}}^x = \tilde{\hat{\mathbf{P}}}^x P^\chi \left[ \tilde{\mathbf{P}}^\chi \hat{\mathbf{P}}^x \frac{\hat{\mathbf{\Lambda}}^x}{n} \right] \hat{\mathbf{G}}^x$,
$c = \tilde{\hat{\mathbf{P}}}^x \mathbf{P}^\chi \left[ \frac{\mathbf{\Lambda}^\chi}{n} \tilde{\mathbf{P}}^\chi \hat{\mathbf{P}}^x \right] \hat{\mathbf{G}}^x = \left[ \frac{1}{n} \tilde{\hat{\mathbf{P}}}^x \mathbf{\Sigma}^\chi \hat{\mathbf{P}}^x \right] \hat{G}^x$,
$d = \left[ \frac{1}{n} \tilde{\hat{\mathbf{P}}}^x \hat{\mathbf{\Sigma}}^x \hat{\mathbf{P}}^x \right] \hat{\mathbf{G}}^x = \frac{\hat{\mathbf{\Lambda}}^x}{n} \hat{\mathbf{G}}^x$,

$f = \mathbf{I}_q$.
We have $\|a-f\| \leq \|a-b\| + \|b-c\| + \|c-d\| + \|d-f\|$. All terms on the right hand side can be shown to be $O_P\left(\max\left(\frac{1}{n}, \rho_T\right)\right)$ uniformly in $\theta$:
The first, using (C.1), Lemma 4 (iii) and boundedness in probability of $\|\tilde{\hat{\mathbf{P}}}^x \mathbf{P}^\chi \tilde{\mathbf{P}}^\chi \hat{\mathbf{P}}^x\|$, uniformly in $\theta$.
The second, using (C.1), statement (i), boundedness in probability, uniformly in $\theta$, of $\|\tilde{\hat{\mathbf{P}}}^x \mathbf{P}^\chi\|$ and $\|\hat{\mathbf{G}}^x\|$, see Lemma 4 (ii).
The third, using (C.1) and boundedness in probability of $\|\hat{\mathbf{G}}^x\|$, uniformly in $\theta$.
The fourth, using (C.1) and Lemma 4 (iii).
As regards (iii), note that, from (i),

$$\tilde{\hat{\mathbf{P}}}^x_h \mathbf{P}^\chi_k \left(\frac{\lambda^\chi_k}{n} - \frac{\hat{\lambda}^x_h}{n}\right) = O_P\left(\max\left(\frac{1}{n}, \sqrt{\rho_T}\right)\right).$$

Assumption A.8 implies that for $h \neq k$

$$\tilde{\hat{\mathbf{P}}}^x_h \mathbf{P}^\chi_k = O_P\left(\max\left(\frac{1}{n}, \sqrt{\rho_T}\right)\right). \tag{C.3}$$

From (ii),

$$\sum_{f=1}^{q} |\tilde{\hat{\mathbf{P}}}^x_h \mathbf{P}^\chi_f|^2 - 1 = O_P\left(\max\left(\frac{1}{n}, \sqrt{\rho_T}\right)\right).$$

Using (C.3),

$$|\tilde{\hat{\mathbf{P}}}^x_h \mathbf{P}^\chi_h|^2 - 1 = \left(|\tilde{\hat{\mathbf{P}}}^x_h \mathbf{P}^\chi_h| - 1\right)\left(|\tilde{P}^\chi_h \hat{P}^x_h| + 1\right) = O_P\left(\max\left(\frac{1}{n}, \sqrt{\rho_T}\right)\right).$$

The conclusion follows. Q.E.D.

Obviously, transposing and conjugating, Lemma 5 holds for $\|\tilde{\hat{\mathbf{P}}}^x \mathbf{P}^\chi \frac{\mathbf{\Lambda}^\chi}{n} - \frac{\hat{\mathbf{\Lambda}}^x}{n} \tilde{\hat{\mathbf{P}}}^x \mathbf{P}^\chi\|$, $\|\tilde{\mathbf{P}}^\chi \hat{\mathbf{P}}^x \tilde{\hat{\mathbf{P}}}^x \mathbf{P}^\chi - \mathbf{I}_q\|$ and $\|\tilde{\mathbf{P}}^\chi \hat{\mathbf{P}}^x - \tilde{\hat{\mathbf{W}}}_q\|$.

**Lemma 6** Given $s$:

$$\max_{|h| \leq B_T} \|\mathcal{S}'_{\mathbf{M}} \left(\mathbf{P}^\chi (\mathbf{\Lambda}^\chi)^{\frac{1}{2}} \hat{\mathbf{W}}_q - \hat{\mathbf{P}}^x \left(\hat{\mathbf{\Lambda}}^x\right)^{\frac{1}{2}}\right)\| = O_P\left(\max\left(\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{\rho_T}}\right)\right), \tag{C.4}$$

uniformly in $\mathbf{M}$.

Proof. We have:

$$\begin{aligned}\|\mathcal{S}'_{\mathbf{M}} \left(\mathbf{P}^\chi (\mathbf{\Lambda}^\chi)^{\frac{1}{2}} \hat{\mathbf{W}}_q - \hat{\mathbf{P}}^x \left(\hat{\mathbf{\Lambda}}^x\right)^{\frac{1}{2}}\right)\| &\leq \|\mathcal{S}'_{\mathbf{M}} \left(\left(\sqrt{n}\mathbf{P}^\chi \hat{\mathbf{W}}_q - \sqrt{n}\hat{\mathbf{P}}^x\right)\left(\frac{\mathbf{\Lambda}^\chi}{n}\right)^{\frac{1}{2}}\right)\| \\ &+ \|\mathcal{S}'_{\mathbf{M}} \hat{\mathbf{P}}^x \left(\left(\frac{\mathbf{\Lambda}^\chi}{n}\right)^{\frac{1}{2}} - \left(\frac{\hat{\mathbf{\Lambda}}^x}{n}\right)^{\frac{1}{2}}\right)\|.\end{aligned}$$

By Lemma 4 (i), we only need to prove that

$$\|\sqrt{n}\mathcal{S}'_{\mathbf{M}}\mathbf{P}^{\chi}\hat{\mathbf{W}}_q - \sqrt{n}\mathcal{S}'_{\mathbf{M}}\hat{\mathbf{P}}^{x}\| = O_P\left(\max\left(\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{\rho_T}}\right)\right).$$

Firstly we show that

$$\|\sqrt{n}\mathcal{S}'_{\mathbf{M}}\mathbf{P}^{\chi}\| = O(1) \tag{C.5}$$

uniformly in $\theta$ and $\mathbf{M}$. Assumption 5 implies that

$$\sigma^{\chi}_{ii} = \sum_{f=1}^{q} \lambda^{\chi}_f |p^{\chi}_{if}|^2 = O(1),$$

uniformly in $\theta$, $f$ and $i$. As all the terms in the sum are positive, $\lambda^{\chi}_f |p^{\chi}_{if}|^2 = \frac{\lambda^{\chi}_f}{n}\left(n|p^{\chi}_{if}|^2\right) = O(1)$, uniformly in $\theta$, $f$ and $i$. Assumption 8 implies that $\frac{\lambda^{\chi}_f}{n}$ is bounded away from zero uniformly in $\theta$ and $f$, so that $n|p^{\chi}_{if}|^2$ must be $O(1)$, uniformly in $\theta$ and $f$ and $i$. As a consequence, the eigenvalues of $\mathcal{S}'_{\mathbf{M}}\mathbf{P}^{\chi}\tilde{\mathbf{P}}^{\chi}\mathcal{S}_{\mathbf{M}}$ are $O(1)$ uniformly in $\theta$ and $\mathbf{M}$. Now set

$a = \sqrt{n}\mathcal{S}'_{\mathbf{M}}\mathbf{P}^{\chi}\left[\hat{\mathbf{W}}_q\right]$,

$b = \sqrt{n}\mathcal{S}'_{\mathbf{M}}\mathbf{P}^{\chi}\left[\tilde{\mathbf{P}}^{\chi}\hat{\mathbf{P}}^{x}\right] = \sqrt{n}\mathcal{S}'_{\mathbf{M}}\mathbf{P}^{\chi}\left[\tilde{\mathbf{P}}^{\chi}\hat{\mathbf{P}}^{x}\frac{\hat{\mathbf{\Lambda}}^{x}}{n}\right]\left(\frac{\hat{\mathbf{\Lambda}}^{x}}{n}\right)^{-1}$,

$c = \sqrt{n}\mathcal{S}'_{\mathbf{M}}\mathbf{P}^{\chi}\left[\frac{\mathbf{\Lambda}^{\chi}}{n}\tilde{\mathbf{P}}^{\chi}\hat{\mathbf{P}}^{x}\right]\left(\frac{\hat{\mathbf{\Lambda}}^{x}}{n}\right)^{-1} = \left[\frac{1}{\sqrt{n}}\mathcal{S}'_{\mathbf{M}}\mathbf{\Sigma}^{\chi}\right]\hat{\mathbf{P}}^{x}\left(\frac{\hat{\mathbf{\Lambda}}^{x}}{n}\right)^{-1}$,

$d = \left[\frac{1}{\sqrt{n}}\mathcal{S}'_{\mathbf{M}}\hat{\mathbf{\Sigma}}^{x}\right]\hat{\mathbf{P}}^{x}\left(\frac{\hat{\mathbf{\Lambda}}^{x}}{n}\right)^{-1} = \sqrt{n}\mathcal{S}'_{\mathbf{M}}\hat{\mathbf{P}}^{x}$.

Using (C.5) and Lemma 5 (iii), we obtain $\|a-b\| = O_P(\max(\frac{1}{n}, \frac{1}{\sqrt{\rho_T}})$.

Using (C.5) and Lemma 5 (i), we obtain $\|b-c\| = O_P(\max(\frac{1}{n}, \frac{1}{\sqrt{\rho_T}})$.

Using Lemma 3 (iv), we obtain $\|c-d\| = O_P(\max(\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{\rho_T}})$. Q.E.D.

Note that the eigenvectors $\mathbf{P}^{\chi}$ are defined up to post-multiplication by a diagonal matrix with unit modulus diagonal entries. In particular, using the eigenvectors $\mathbf{\Pi}^{\chi} = \mathbf{P}^{\chi}\hat{\mathbf{W}}_q$, (C.4) would hold for $\mathbf{\Pi}^{\chi}(\mathbf{\Lambda}^{\chi})^{\frac{1}{2}} - \hat{\mathbf{P}}^{x}\left(\hat{\mathbf{\Lambda}}^{x}\right)^{\frac{1}{2}}$. We avoid introducing a new symbol and refer henceforth to the result of Lemma 6 as:

$$\max_{|h|\le B_T}\|\mathcal{S}'_{\mathbf{M}}\left(\mathbf{P}^{\chi}(\mathbf{\Lambda}^{\chi})^{\frac{1}{2}} - \hat{\mathbf{P}}^{x}\left(\hat{\mathbf{\Lambda}}^{x}\right)^{\frac{1}{2}}\right)\| = O_P\left(\max\left(\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{\rho_T}}\right)\right). \tag{C.6}$$

In the same way, Lemma 5(iii) will be referred to as:

$$\|\tilde{\hat{\mathbf{P}}}^{x}\mathbf{P}^{\chi} - \mathbf{I}_q\| = O_P\left(\max\left(\frac{1}{n}, \frac{1}{\sqrt{\rho_T}}\right)\right). \tag{C.7}$$

Lastly, Proposition 3 follows from $\hat{\mathbf{\Sigma}}^{x} = \hat{\mathbf{P}}^{x}\hat{\mathbf{\Lambda}}^{x}\tilde{\hat{\mathbf{P}}}^{x}$ and $\mathbf{\Sigma}^{\chi} = \mathbf{P}^{\chi}\mathbf{\Lambda}^{\chi}\tilde{\mathbf{P}}^{\chi}$.

# D Proof of Proposition 4

From (2.16), using (C.1),

$$\|\hat{\mathbf{A}}^{[k]} - \mathbf{A}^{[k]}\| \leq \|\hat{\mathbf{B}}^{\chi}_{kk}\| \|(\hat{\mathbf{C}}^{\chi}_{kk})^{-1} - (\mathbf{C}^{\chi}_{kk})^{-1}\| + \|\hat{\mathbf{B}}^{\chi}_{k} - \mathbf{B}^{\chi}_{k}\| \|(\mathbf{C}^{\chi}_{kk})^{-1}\|.$$

Proposition 3 implies that $\|\hat{\mathbf{B}}^{\chi}_{k} - \mathbf{B}^{\chi}_{k}\|$ is $O_P\left(\max\left(\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{\rho_T}}\right)\right)$. By Assumption A.5, $\|\mathbf{B}^{\chi}_{k}\| \leq W$ for some $W > 0$ and all $k$, so that $\|\hat{\mathbf{B}}^{\chi}_{k}\|$ is bounded in probability uniformly in $k$. By Assumptions A.5 and A.4, $\|(\mathbf{C}^{\chi}_{kk})^{-1}\| \leq W_1$ for some $W_1 > 0$ and all $k$. Observing that the entries of $(\mathbf{C}^{\chi}_{kk})^{-1}$ are rational functions of the entries of $\mathbf{C}^{\chi}_{kk}$, and that $\det(\mathbf{C}^{\chi}_{kk})$ is bounded away from zero as a function of $k$, Proposition 3 implies that $\|(\hat{\mathbf{C}}^{\chi}_{kk})^{-1} - (\mathbf{C}^{\chi}_{kk})^{-1}\|$ is $O_P\left(\max\left(\frac{1}{n}, \sqrt{\rho_T}\right)\right)$ uniformly in $k$. Thus $\|\hat{\mathbf{A}}^{[k]} - \mathbf{A}^{[k]}\| = O_P\left(\max\left(\frac{1}{n}, \sqrt{\rho_T}\right)\right)$ uniformly in $k$. Using (C.1),

$$\begin{aligned}\|\hat{\mathbf{A}}^{[j]}\hat{\mathbf{C}}^{\chi}_{jk}\hat{\mathbf{A}}^{[k]} - \mathbf{A}^{[j]}\mathbf{C}^{\chi}_{jk}\mathbf{A}^{[k]}\| \leq \|\hat{\mathbf{A}}^{[j]}\hat{\mathbf{C}}^{\chi}_{jk}\| \|\hat{\mathbf{A}}^{[k]} - \mathbf{A}^{[k]}\| + \|\hat{\mathbf{A}}^{[j]}\| \|\hat{\mathbf{C}}^{\chi}_{jk} - \mathbf{C}^{\chi}_{jk}\| \|\mathbf{A}^{[k]}\| \\ + \|\hat{\mathbf{A}}^{[j]} - \mathbf{A}^{[j]}\| \|\mathbf{C}^{\chi}_{jk}\mathbf{A}^{[k]}\|.\end{aligned}$$

The conclusion follows.

# E Proof of Proposition 5

Consider the static model

$$\mathbf{z}_{nt} = \mathcal{R}\mathbf{v}_t + \boldsymbol{\phi}_{nt}.$$

If $\mathbf{z}_{nt} = \mathbf{A}(L)\mathbf{x}_{nt}$ were observed, i.e. if the matrices $\mathbf{A}(L)$ were observed, then Proposition 5 would be easy to prove. However, we only observe $\hat{\mathbf{z}}_{nt} = \hat{\mathbf{A}}(L)\mathbf{x}_t$. Lemmas 7 through 12 prepare the following result,

$$\frac{1}{n}\|\hat{\mathbf{\Gamma}}^z - \mathbf{\Gamma}^z\| = O_P(\zeta_{Tn}), \qquad \frac{1}{\sqrt{n}}\|\mathcal{S}_{\mathbf{M}}\left(\hat{\mathbf{\Gamma}}^z - \mathbf{\Gamma}^z\right)\| = O_P(\zeta_{Tn}), \tag{E.1}$$

where $\hat{\mathbf{\Gamma}}^z$ and $\mathbf{\Gamma}^z$ are the covariance matrices of $\hat{\mathbf{z}}_{nt}$ and $\mathbf{z}_{nt}$ respectively, which is proved in Lemma 13. Starting with (E.1), which plays here the same role as Assumption A.7 for the proof of Proposition 2, we can easily prove statements that replicate in this context Lemmas 3, 4, 5 and 6, using the same arguments used in Section C, with $x$, $\chi$ and $\xi$ replaced by $z$, $\psi$ and $\phi$ respectively. Precisely:

(I) In the results corresponding to Lemma 3 we obtain the rate $\zeta_{Tn}$ for (i), (ii), (iii) and (iv). Note that no reduction from $1/n$ to $1/\sqrt{n}$ occurs between (iii) and (iv), as in Lemma 3. For, (iii) has $O_P(\zeta_{Tn}) + O(1/n) = O_P(\zeta_{Tn})$, while (iv) has $O_P(\zeta_{Tn}) + O(1/\sqrt{n}) = O_P(\zeta_{Tn})$.

(II) The same rate $\zeta_{Tn}$ is obtained for the results of Lemma 4.

(III) The same holds for Lemma 5. The orthogonal matrix in point (iii), call it again $\hat{\mathbf{W}}_q$, has either 1 or $-1$ on the diagonal. Thus $\tilde{\hat{\mathbf{W}}}_q = \hat{\mathbf{W}}_q$.

(IV) Lastly, Lemma 6 becomes

$$\|\mathcal{S}'_{\mathbf{M}}\left(\hat{\mathbf{P}}^z\left(\hat{\mathbf{\Lambda}}^z\right)^{\frac{1}{2}} - \mathbf{P}^\psi\left(\mathbf{\Lambda}^\psi\right)^{\frac{1}{2}}\hat{\mathbf{W}}_q\right)\| = O_P\left(\zeta_{Tn}\right). \tag{E.2}$$

Going over the proof of Lemma 6, we see that $\|c-d\|$ has the worst rate, whereas here $\|a-b\|$, $\|b-c\|$ and $\|c-d\|$ all have rate $O_P(\zeta_{Tn})$.

(V) Moreover, in the same way as the proof of Lemma 6 can be replicated to obtain (E.2), the proof of Lemma 8, see below, can be replicated to obtain:

$$\|\hat{\mathbf{P}}^z(\hat{\mathbf{\Lambda}}^z)^{\frac{1}{2}} - \mathbf{P}^\psi(\mathbf{\Lambda}^\psi)^{\frac{1}{2}}\hat{\mathbf{W}}_q\| = O_P\left(n^{\frac{1}{2}}\zeta_{Tn}\right). \tag{E.3}$$

**Lemma 7** (i) For $n\to\infty$ and $T\to\infty$,

$$\mid p^\chi_{if}\mid = O\left(n^{-\frac{1}{2}}\right),\ \mid \hat{p}^x_{if}\mid = O_P\left(n^{-\frac{1}{2}}\right) \tag{E.4}$$

uniformly in $\theta$, $i\in\mathbb{N}$ and $f=1,\dots,q$.
(ii) Given a positive integer $d$, for $n\to\infty$ and $T\to\infty$,

$$\frac{1}{n}\sum_{i=1}^n \mid p^\chi_{if}\mid^d = O\left(n^{-\frac{d}{2}}\right),\ \frac{1}{n}\sum_{i=1}^n \mid \hat{p}^x_{if}\mid^d = O_P\left(n^{-\frac{d}{2}}\right).$$

uniformly in $\theta$ and $f=1,\dots,q$.

Proof. The first part of (E.4) has been proved (see the proof of Lemma 6). Lemma 6 and Assumption 5 imply that

$$\hat{\sigma}^x_{ii} = \sum_{f=1}^q \hat{\lambda}^x_f|\hat{p}^x_{if}|^2 = O_P(1),$$

uniformly in $\theta$ and $i$. As all the terms in the sum are positive, $\hat{\lambda}^x_f|\hat{p}^x_{if}|^2 = \frac{\hat{\lambda}^x_f}{n}\left(n|\hat{p}^x_{if}|^2\right) = O_P(1)$, uniformly in $i$, $f$ and $\theta$. Lemma 4 and Assumption 8 imply that $\frac{\hat{\lambda}^x_f}{n}$ is $O_P(1)$ and bounded away from zero in probability uniformly in $\theta$ and $f$. The conclusion follows.
We prove (ii) by induction. Consider first $\mathbf{P}^\chi_f$. When $d=1$

$$\frac{1}{n}\sum_{i=1}^n |p^\chi_{if}| \le \left(\frac{1}{n}\sum_{i=1}^n |p^\chi_{if}|^2\right)^{\frac{1}{2}} = O(n^{-\frac{1}{2}}).$$

Assume now that $d \geq 2$ and the result holds for $d-1$. Summing by parts and using (i) of this Lemma,

$$\begin{aligned}
\tfrac{1}{n}\sum_{i=1}^{n} \mid p_{if}^{\chi} \mid^{d} &= \tfrac{1}{n}\sum_{i=1}^{n} \mid p_{if}^{\chi} \mid^{d-1} \mid p_{if}^{\chi} \mid \\
&= \frac{\mid p_{nf}^{\chi} \mid}{n} \sum_{i=1}^{n} \mid p_{if}^{\chi} \mid^{d-1} - \frac{1}{n} \sum_{i=1}^{n-1} \left( \sum_{s=1}^{i} \mid p_{sf}^{\chi} \mid^{d-1} \right) (\mid p_{i+1,f}^{\chi} \mid - \mid p_{if}^{\chi} \mid) \\
&\leq |p_{nf}^{\chi}| \left( \tfrac{1}{n} \sum_{i=1}^{n} \mid p_{if}^{\chi} \mid^{d-1} \right) = O\left( n^{-\frac{1}{2}} n^{-\frac{(d-1)}{2}} \right) = O\left( n^{-\frac{d}{2}} \right),
\end{aligned}$$

the inequality holding because without loss of generality (reordering) we can assume $\mid p_{i+1,f}^{\chi} \mid \geq \mid p_{if}^{\chi} \mid$. The same argument applies to $\hat{\mathbf{P}}_{f}^{x}$. Q.E.D.

**Lemma 8**

$$\max_{|h| \leq B_T} \left\| \mathbf{P}^{\chi} \left( \mathbf{\Lambda}^{\chi} \right)^{\frac{1}{2}} \hat{\mathbf{W}}_q - \hat{\mathbf{P}}^{x} \left( \hat{\mathbf{\Lambda}}^{x} \right)^{\frac{1}{2}} \right\| = O_P \left( \sqrt{n} \max \left( \frac{1}{n}, \frac{1}{\sqrt{\rho_T}} \right) \right). \tag{E.5}$$

Proof. The left hand side of (E.5) equals the left hand side of (C.4) when $\mathcal{S}_{\mathbf{M}}$ is replaced by $\mathbf{I}_n$. The proof goes in the same way as that of Lemma 6. Firstly, $\|\sqrt{n}\mathbf{P}^{\chi}\| = O(\sqrt{n})$. Both $\|a-b\|$ and $\|b-c\|$ are $O_P(\sqrt{n} \max \left( \frac{1}{n}, \frac{1}{\sqrt{\rho_T}} \right)$. Lastly, as regards $\|c-d\|$, $\frac{1}{\sqrt{n}} \left( \mathbf{\Sigma}^{\chi} - \hat{\mathbf{\Sigma}}^{x} \right) = \sqrt{n} \left( \frac{1}{n} \left( \mathbf{\Sigma}^{\chi} - \hat{\mathbf{\Sigma}}^{x} \right) \right)$. The conclusion follows from Lemma 3, (iii). Q.E.D.

**Lemma 9** (i) For $n \to \infty$ and $T \to \infty$,

$$|p_{if}^{\chi} - \hat{p}_{if}^{x}| = O_P \left( \frac{1}{\sqrt{n}} \max \left( \frac{1}{\sqrt{n}}, \frac{1}{\sqrt{\rho_T}} \right) \right),$$

uniformly in $\theta$, $i \in \mathbb{N}$ and $f = 1, \ldots, q$.
(ii) For $n \to \infty$ and $T \to \infty$,

$$\frac{1}{n} \sum_{i=1}^{n} |p_{if}^{\chi} - \hat{p}_{if}^{x}| = O_P \left( \frac{1}{\sqrt{n}} \max \left( \frac{1}{n}, \frac{1}{\sqrt{\rho_T}} \right) \right),$$

uniformly in $\theta$ and $f = 1, \ldots, q$.

Proof. As regards (i), by (C.6)

$$p_{if}^{\chi} \sqrt{\lambda_f^{\chi}} - \hat{p}^{x} \sqrt{\hat{\lambda}_f^{x}} = O_P \left( \max \left( \frac{1}{\sqrt{n}}, \frac{1}{\sqrt{\rho_T}} \right) \right).$$

We have:

$$p_{if}^{\chi} \sqrt{\lambda_f^{\chi}} - \hat{p}^{x} \sqrt{\hat{\lambda}_f^{x}} = p_{if}^{\chi} \left( \sqrt{\lambda_f^{\chi}} - \sqrt{\hat{\lambda}_f^{x}} \right) + \sqrt{\hat{\lambda}_f^{x}} \left( p_{if}^{\chi} - \hat{p}_{if}^{x} \right). \tag{E.6}$$

For the first term on the right-hand side:

$$p_{if}^{\chi}\left(\sqrt{\lambda_f^{\chi}}-\sqrt{\hat{\lambda}_f^{x}}\right)=\sqrt{n}p_{if}^{\chi}\frac{\dfrac{\lambda_{if}^{\chi}-\hat{\lambda}_{if}^{x}}{n}}{\dfrac{\sqrt{\lambda_f^{\chi}}+\sqrt{\hat{\lambda}_f^{x}}}{\sqrt{n}}}=O_P\left(\max\left(\frac{1}{n},\frac{1}{\sqrt{\rho_T}}\right)\right),$$

by Lemma 4(i), Assumption A.8 and Lemma 9(i) above. Thus

$$\left(\sqrt{\frac{\hat{\lambda}_f^{x}}{n}}\right)\sqrt{n}\left(p_{if}^{\chi}-\hat{p}_{if}^{x}\right)=O_P\left(\max\left(\frac{1}{\sqrt{n}},\sqrt{\rho_T}\right)\right).$$

By Assumption A.8, $\sqrt{\hat{\lambda}_f^{x}}/\sqrt{n}$ is bounded away from zero. The conclusion follows. Regarding (ii), taking the modulus and summing across $i=1,...,n$ in (E.6) yields

$$\left(\sqrt{\frac{\hat{\lambda}_f^{x}}{n}}\right)\sum_{i=1}^{n}\left|p_{if}^{\chi}-\hat{p}_{if}^{x}\right|\leq\frac{1}{\sqrt{n}}\left(\sum_{i=1}^{n}\left|p_{if}^{\chi}\sqrt{\lambda_f^{\chi}}-\hat{p}^{x}\sqrt{\hat{\lambda}_f^{x}}\right|\right)+\left|\sqrt{\frac{\lambda_f^{\chi}}{n}}-\frac{\sqrt{\hat{\lambda}_f^{x}}}{n}\right|\sum_{i=1}^{n}|p_{if}^{\chi}|.$$

Regarding the first term on the right hand side above, by Jensen's inequality and Lemma 8

$$\sum_{i=1}^{n}\left|p_{if}^{\chi}\sqrt{\lambda_f^{\chi}}-\hat{p}^{x}\sqrt{\hat{\lambda}_f^{x}}\right|\leq\sqrt{n}\left(\sum_{i=1}^{n}\left|p_{if}^{\chi}\sqrt{\lambda_f^{\chi}}-\hat{p}^{x}\sqrt{\hat{\lambda}_f^{x}}\right|^2\right)^{\frac{1}{2}}=O_P\left(n\max[\frac{1}{n},\frac{1}{\sqrt{\rho_T}}]\right).$$

Lemma 4(i)-(ii) and Lemma 7(ii) permit to bound the second term. Q.E.D.

**Lemma 10** For any integer $d\in N$, for $n\to\infty$ and $T\to\infty$,

$$\frac{1}{n}\sum_{i=1}^{n}\mid p_{if}^{\chi}-\hat{p}_{if}^{x}\mid^{d}=O_P\left(\left(\frac{1}{n}\max\left(\frac{1}{n},\frac{1}{\rho_T}\right)\right)^{\frac{d}{2}}\right),$$

uniformly in $\theta$ and $f=1,\dots,q$.

Proof. By induction. Lemma 9(ii) implies

$$\frac{1}{n}\sum_{i=1}^{n}\mid \hat{p}_{if}^{\chi}-p_{ij}^{\chi}\mid=O_P\left(\left(\frac{1}{n}\max\left(\frac{1}{n},\frac{1}{\rho_T}\right)\right)^{\frac{1}{2}}\right).$$

In fact, to avoid un-necessary complications we prefer to use here a slightly looser bound than the one provided by Lemma 9. Assume now that $d\geq 2$ and that the result holds for $d-1$. Using summation by parts,

$$\frac{1}{n}\sum_{i=1}^{n} \mid p_{if}^{\chi} - \hat{p}_{if}^{x} \mid^{d} = \frac{1}{n}\sum_{i=1}^{n} \mid p_{if}^{\chi} - \hat{p}_{if}^{x} \mid^{d-1} \mid p_{if}^{\chi} - \hat{p}_{if}^{x} \mid$$
$$= \mid p_{nf}^{\chi} - \hat{p}_{nf}^{x} \mid \frac{1}{n}\sum_{i=1}^{n} \mid p_{if}^{\chi} - \hat{p}_{if}^{x} \mid^{d-1}$$
$$- \frac{1}{n}\sum_{i=1}^{n-1}\left(\sum_{k=1}^{i} \mid p_{kf}^{\chi} - \hat{p}_{kf}^{x} \mid^{d-1}\right)(\mid p_{i+1,f}^{\chi} - \hat{p}_{i+1,f}^{x} \mid - \mid p_{if}^{\chi} - \hat{p}_{if}^{x} \mid)$$
$$\leq \mid p_{nf}^{\chi} - \hat{p}_{nf}^{x} \mid \frac{1}{n}\sum_{i=1}^{n} \mid p_{if}^{\chi} - \hat{p}_{if}^{x} \mid^{d-1} = \mid p_{nf}^{\chi} - \hat{p}_{nf}^{x} \mid O_P\left(\left(\frac{1}{n}\max\left(\frac{1}{n}, \frac{1}{\rho_T}\right)\right)^{\frac{d-1}{2}}\right),$$

the inequality holding since without loss of generality we can assume $\mid p_{i+1,f}^{\chi} - \hat{p}_{i+1,f}^{x} \mid \geq \mid p_{if}^{\chi} - \hat{p}_{if}^{x} \mid$. Finally use Lemma 9(i). Q.E.D.

**Lemma 11** For $n \to \infty$ and $T \to \infty$,

$$\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} \mid \hat{\sigma}_{ij}^{\chi}(\theta) - \sigma_{ij}^{\chi}(\theta)|^{d} = \left(\left(\max\left(\frac{1}{n}, \frac{1}{\rho_T}\right)\right)^{\frac{d}{2}}\right)$$

uniformly in $\theta$.

$$\frac{1}{n}\sum_{i=1}^{n} \mid \hat{\sigma}_{ij}^{\chi}(\theta) - \sigma_{ij}^{\chi}(\theta) \mid^{d} = O_P\left(\left(\max\left(\frac{1}{n}, \frac{1}{\rho_T}\right)\right)^{\frac{d}{2}}\right),$$

uniformly in $\theta$ and any $1 \leq j \leq n$.

$$\frac{1}{n}\sum_{i=1}^{n} \mid \hat{\sigma}_{ii}^{\chi}(\theta) - \sigma_{ii}^{\chi}(\theta) \mid^{d} = O_P\left(\left(\max\left(\frac{1}{n}, \frac{1}{\rho_T}\right)\right)^{\frac{d}{2}}\right),$$

uniformly in $\theta$.

PROOF. We have

$$\begin{aligned}\hat{\sigma}_{ij}^{\chi} - \sigma_{ij}^{\chi} = \; & (\hat{\lambda}_1^{x} - \lambda_1^{\chi})\hat{p}_{i1}^{x}\bar{\hat{p}}_{j1}^{x} + \cdots + (\hat{\lambda}_q^{x} - \lambda_q^{\chi})\hat{p}_{iq}^{x}\bar{\hat{p}}_{jq}^{x} \\ & + \lambda_1^{\chi}\hat{p}_{i1}^{x}(\bar{\hat{p}}_{j1}^{x} - \bar{p}_{j1}^{\chi}) + \lambda_1^{\chi}\bar{p}_{j1}^{\chi}(\hat{p}_{i1}^{x} - p_{i1}^{\chi}) \\ & \vdots \\ & + \lambda_q^{\chi}\hat{p}_{iq}^{x}(\bar{\hat{p}}_{jq}^{x} - \bar{p}_{jq}^{\chi}) + \lambda_q^{\chi}\bar{p}_{jq}^{\chi}(\hat{p}_{iq}^{x} - p_{iq}^{\chi}).\end{aligned}$$

Using the triangular and $c_r$ inequalities, by Lemmas 4, 7 and 10,

$$\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} \mid \hat{\sigma}_{ij}^{\chi} - \sigma_{ij}^{\chi} \mid^d \leq$$
$$(3q)^{d-1}\left(\mid \lambda_1^{\chi} - \hat{\lambda}_1^{x} \mid^d \left(\frac{1}{n}\sum_{i=1}^{n} \mid \hat{p}_{i1}^{x} \mid^d\right)^2 + \cdots + \mid \lambda_q^{\chi} - \hat{\lambda}_q^{x} \mid^d \left(\frac{1}{n}\sum_{i=1}^{n} \mid \hat{p}_{iq}^{x} \mid^d\right)^2\right)$$
$$+(3q)^{d-1}\left(\lambda_1^{\chi}\right)^d\left(\frac{1}{n^2}\sum_{i=1}^{n} \mid \hat{p}_{i1}^{x} \mid^d \sum_{j=1}^{n} \mid p_{j1}^{\chi} - \hat{p}_{j1}^{x} \mid^d + \frac{1}{n^2}\sum_{j=1}^{n} \mid p_{j1}^{\chi} \mid^d \sum_{i=1}^{n} \mid p_{i1}^{\chi} - \hat{p}_{i1}^{x} \mid^d\right)$$
$$\vdots$$
$$+(3q)^{d-1}\left(\lambda_q^{\chi}\right)^d\left(\frac{1}{n^2}\sum_{i=1}^{n} \mid \hat{p}_{iq}^{x} \mid^d \sum_{j=1}^{n} \mid p_{jq}^{\chi} - \hat{p}_{jq}^{x} \mid^d + \frac{1}{n^2}\sum_{j=1}^{n} \mid p_{jq}^{\chi} \mid^d \sum_{i=1}^{n} \mid p_{iq}^{\chi} - \hat{p}_{iq}^{x} \mid^d\right)$$
$$= O_P\left(\left(\max\left(\frac{1}{n}, \frac{1}{\sqrt{\rho_T}}\right)\right)^d\right) + O_P\left(\left(\max\left(\frac{1}{n}, \frac{1}{\rho_T}\right)\right)^{\frac{d}{2}}\right) = O_P\left(\left(\max\left(\frac{1}{n}, \frac{1}{\rho_T}\right)\right)^{\frac{d}{2}}\right).$$

For the second statement

$$\frac{1}{n}\sum_{i=1}^{n} \mid \hat{\sigma}_{ij}^{\chi} - \sigma_{ij}^{\chi} \mid^d \leq$$
$$(3q)^{d-1}\left(\mid \lambda_1^{\chi} - \hat{\lambda}_1^{x}\mid^d |\hat{p}_{j1}^{x}|^d \frac{1}{n}\sum_{i=1}^{n} |\hat{p}_{i1}^{x}|^d + \cdots + \mid \lambda_q^{\chi} - \hat{\lambda}_q^{x} \mid^d |\hat{p}_{jq}^{x}|^d \frac{1}{n}\sum_{i=1}^{n} \mid \hat{p}_{iq}^{x} \mid^d\right)$$
$$+(3q)^{d-1}\left(\lambda_1^{\chi}\right)^d\left(\left|p_{j1}^{\chi} - \hat{p}_{j1}^{x}\right|^d \frac{1}{n}\sum_{i=1}^{n} |\hat{p}_{i1}^{x}|^d + \left|p_{j1}^{\chi}\right|^d \frac{1}{n}\sum_{i=1}^{n} \mid p_{i1}^{\chi} - \hat{p}_{i1}^{x} \mid^d\right)$$
$$\vdots$$
$$+(3q)^{d-1}\left(\lambda_q^{\chi}\right)^d\left(\left|p_{jq}^{\chi} - \hat{p}_{jq}^{x}\right|^d \frac{1}{n}\sum_{i=1}^{n} \left|\hat{p}_{iq}^{x}\right|^d + \left|p_{jq}^{\chi}\right|^d \frac{1}{n}\sum_{i=1}^{n} \mid p_{iq}^{\chi} - \hat{p}_{iq}^{x} \mid^d\right)$$
$$= O_P\left(\left(\max\left(\frac{1}{n}, \frac{1}{\sqrt{\rho_T}}\right)\right)^d\right) + O_P\left(\left(\max\left(\frac{1}{n}, \frac{1}{\rho_T}\right)\right)^{\frac{d}{2}}\right) = O_P\left(\left(\max\left(\frac{1}{n}, \frac{1}{\rho_T}\right)\right)^{\frac{d}{2}}\right).$$

Case $j = i$ easily follows.

Q.E.D.

The arguments used in Section 3.1 to obtain Proposition 3 from Proposition 2 can be used to prove the following statement. Let us simplify the notation by setting $\gamma_{ij}^{\chi} = \gamma_{ij,0}^{\chi}$ and $\hat{\gamma}_{ij}^{\chi} = \hat{\gamma}_{ij,0}^{\chi}$.

**Lemma 12** For $n \to \infty$ and $T \to \infty$,

$$\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} \mid \hat{\gamma}_{ij}^{\chi} - \gamma_{ij}^{\chi}|^d = O_P\left(\left(\max\left(\frac{1}{n}, \frac{1}{\rho_T}\right)\right)^{\frac{d}{2}}\right),$$

and

$$\frac{1}{n}\sum_{i=1}^{n} \mid \hat{\gamma}_{ij}^{\chi} - \gamma_{ij}^{\chi} \mid^d = O_P\left(\left(\max\left(\frac{1}{n}, \frac{1}{\rho_T}\right)\right)^{\frac{d}{2}}\right),$$

uniformly in $j$.

**Lemma 13** Assume that $n$ increases by blocks of size $q+1$, so that $n = m(q+1)$. Denote by $\hat{\mathbf{Z}}$ the $T \times n$ matrix with $\hat{z}_{it}$ in entry $(t,\ i)$ and define $\hat{\mathbf{\Gamma}}^z = (\hat{\mathbf{Z}}'\hat{\mathbf{Z}})/T$. We have:

$$\frac{1}{n}\|\hat{\mathbf{\Gamma}}^z - \mathbf{\Gamma}^z\| = O_P(\zeta_{Tn}),$$

and

$$\frac{1}{\sqrt{n}}\|\mathcal{S}'_{\mathbf{M}}(\hat{\mathbf{\Gamma}}^z - \mathbf{\Gamma}^z)\| = O_P(\zeta_{Tn}),$$

where $\mathbf{\Gamma}^z$ is the population covariance matrix of $\mathbf{z}_{nt}$.

Proof. Define $\check{\mathbf{\Gamma}}^z = (\mathbf{Z}'\mathbf{Z})/T$, the covariance matrix that we would estimate if $\mathbf{z}_{nt}$ were observed (the matrices $\mathbf{A}(L)$ known). We have

$$\|\hat{\mathbf{\Gamma}}^z - \mathbf{\Gamma}^z\| \leq \|\hat{\mathbf{\Gamma}}^z - \check{\mathbf{\Gamma}}^z\| + \|\check{\mathbf{\Gamma}}^z - \mathbf{\Gamma}^z\|, \tag{E.7}$$

so that the lemma can be proved separately for the two terms on the right hand side. Consider firstly $\|\check{\mathbf{\Gamma}}^z - \mathbf{\Gamma}^z\|$. Using

$$\mathbf{A}(L) = \mathbf{I}_n - \mathbf{A}_1 L - \cdots - \mathbf{A}_S L^S,$$

where

$$\mathbf{A}_s = \begin{pmatrix} \mathbf{A}_s^1 & 0 & \cdots & 0 \\ 0 & \mathbf{A}_s^2 & \cdots & 0 \\ \vdots & & \ddots & \\ 0 & 0 & \cdots & \mathbf{A}_s^m, \end{pmatrix}$$

and setting $\mathbf{A}_0 = \mathbf{I}_n$,

$$\|\check{\mathbf{\Gamma}}^z - \mathbf{\Gamma}^z\|^2 \leq \sum_{s=0}^{S}\sum_{r=0}^{S} \|\mathbf{A}_s\check{\mathbf{\Gamma}}^x_{s-r}\mathbf{A}_r{}' - \mathbf{A}_s\mathbf{\Gamma}^x_{s-r}\mathbf{A}_r{}'\|^2 = \sum_{s=0}^{S}\sum_{r=0}^{S} \|\mathbf{A}_s\left(\check{\mathbf{\Gamma}}^x_{s-r} - \mathbf{\Gamma}^x_{s-r}\right)\mathbf{A}_r{}'\|^2,$$

which is the sum of $(S+1)^2$ terms. Inspection of the right hand term shows that a proof of the lemma can be easily obtained for $\|\check{\mathbf{\Gamma}}^z - \mathbf{\Gamma}^z\|$ under our assumptions (in particular, A.4, A.5, A.7, A.9).

Turning to $\|\hat{\mathbf{\Gamma}}^z - \check{\mathbf{\Gamma}}^z\|$,

$$\|\hat{\mathbf{\Gamma}}^z - \check{\mathbf{\Gamma}}^z\|^2 \leq \sum_{s=0}^{S}\sum_{r=0}^{S} \|\hat{\mathbf{A}}_s\hat{\mathbf{\Gamma}}^x_{s-r}\hat{\mathbf{A}}_r{}' - \mathbf{A}_s\hat{\mathbf{\Gamma}}^x_{s-r}\mathbf{A}_r{}'\|^2,$$

it is sufficient to prove our statements for each couple $(s,\ r)$. Denoting by $\mathbf{a}^j_{s\alpha}$ the $\alpha$-th column of $\mathbf{A}^{j\prime}_s$, we have

$$\begin{aligned}\|\hat{\mathbf{A}}^s\hat{\boldsymbol{\Gamma}}^x_{s-r}\hat{\mathbf{A}}^{r\prime}-\mathbf{A}^s\hat{\boldsymbol{\Gamma}}^x_{s-r}\mathbf{A}^{r\prime}\|^2 &\leq \sum_{j=1}^m\sum_{k=1}^m\sum_{\alpha=1}^{q+1}\sum_{\beta=1}^{q+1}\left(\hat{\mathbf{a}}^j_{s\alpha}{}'\hat{\boldsymbol{\Gamma}}^x_{jk,s-r}\hat{\mathbf{a}}^k_{r\beta}-\mathbf{a}^j_{s\alpha}{}'\hat{\boldsymbol{\Gamma}}^x_{jk,s-r}\mathbf{a}^k_{r\beta}\right)^2\\ &\leq 2\sum_{j=1}^m\sum_{k=1}^m\sum_{\alpha=1}^{q+1}\sum_{\beta=1}^{q+1}\left((\hat{\mathbf{a}}^j_{s\alpha}-\mathbf{a}^j_{s\alpha})'\hat{\boldsymbol{\Gamma}}^x_{jk,s-r}\hat{\mathbf{a}}^k_{r\beta}\right)^2\\ &\quad+2\sum_{j=1}^m\sum_{k=1}^m\sum_{\alpha=1}^{q+1}\sum_{\beta=1}^{q+1}\left(\mathbf{a}^j_{s\alpha}{}'\hat{\boldsymbol{\Gamma}}^x_{jk,s-r}(\hat{\mathbf{a}}^k_{r\beta}-\mathbf{a}^k_{r\beta})\right)^2,\end{aligned}\tag{E.8}$$

where $\hat{\boldsymbol{\Gamma}}^x_{jk,s-r}$ is the $(j,k)$-th block of $\hat{\boldsymbol{\Gamma}}^x_{s-r}=T^{-1}\sum_{t=1}^T\mathbf{x}_{t-r}\mathbf{x}'_{t-s}$, the second inequality being obtained applying the $c_r$ inequality to each term

$$\left(\hat{\mathbf{a}}^j_{s\alpha}{}'\hat{\boldsymbol{\Gamma}}^x_{jk,s-r}\hat{\mathbf{a}}^k_{r\beta}-\mathbf{a}^j_{s\alpha}{}'\hat{\boldsymbol{\Gamma}}^x_{jk,s-r}\mathbf{a}^k_{r\beta}\right)^2=\left((\hat{\mathbf{a}}^j_{s\alpha}-\mathbf{a}^j_{s\alpha})'\hat{\boldsymbol{\Gamma}}^x_{jk,s-r}\hat{\mathbf{a}}^k_{r\beta}-\mathbf{a}^j_{s\alpha}{}'\hat{\boldsymbol{\Gamma}}^x_{jk,s-r}(\hat{\mathbf{a}}^k_{r\beta}-\mathbf{a}^k_{r\beta})\right)^2.$$

The two terms on the right-hand side of (E.8) can be dealt with in the same way. Let us focus on the first. We have;

$$\begin{aligned}&\sum_{j=1}^m\sum_{k=1}^m\sum_{\alpha=1}^{q+1}\sum_{\beta=1}^{q+1}\left((\hat{\mathbf{a}}^j_{s\alpha}-\mathbf{a}^j_{s\alpha})'\hat{\boldsymbol{\Gamma}}^x_{jk,s-r}\hat{\mathbf{a}}^k_{r\beta}\right)^2\\ [1]\qquad&\leq\sum_{j=1}^m\sum_{k=1}^m\sum_{\alpha=1}^{q+1}\sum_{\beta=1}^{q+1}\left((\hat{\mathbf{a}}^j_{s\alpha}-\mathbf{a}^j_{s\alpha})'(\hat{\mathbf{a}}^j_{s\alpha}-\mathbf{a}^j_{s\alpha})\right)\left(\hat{\mathbf{a}}^k_{r\beta}{}'\hat{\boldsymbol{\Gamma}}^x_{jk,s-r}{}'\hat{\boldsymbol{\Gamma}}^x_{jk,s-r}\hat{\mathbf{a}}^k_{r\beta}\right)\\ &=\sum_{k=1}^m\sum_{\beta=1}^{q+1}\left\{\sum_{j=1}^m\left[\sum_{\alpha=1}^{q+1}\left((\hat{\mathbf{a}}^j_{s\alpha}-\mathbf{a}^j_{s\alpha})'(\hat{\mathbf{a}}^j_{s\alpha}-\mathbf{a}^j_{s\alpha})\right)\right]\left(\hat{\mathbf{a}}^k_{r\beta}{}'\hat{\boldsymbol{\Gamma}}^x_{jk,s-r}{}'\hat{\boldsymbol{\Gamma}}^x_{jk,s-r}\hat{\mathbf{a}}^k_{r\beta}\right)\right\}\\ [2]\qquad&\leq\sum_{k=1}^m\sum_{\beta=1}^{q+1}\left\{\left[\sum_{j=1}^m\left[\sum_{\alpha=1}^{q+1}(\hat{\mathbf{a}}^j_{s\alpha}-\mathbf{a}^j_{s\alpha})'(\hat{\mathbf{a}}^j_{s\alpha}-\mathbf{a}^j_{s\alpha})\right]^2\right]^{\frac12}\left[\sum_{j=1}^m\left(\hat{\mathbf{a}}^k_{r\beta}{}'\hat{\boldsymbol{\Gamma}}^x_{jk,s-r}{}'\hat{\boldsymbol{\Gamma}}^x_{jk,s-r}\hat{\mathbf{a}}^k_{r\beta}\right)^2\right]^{\frac12}\right\}\\ &=m\left[\sum_{j=1}^m\left[\sum_{\alpha=1}^{q+1}(\hat{\mathbf{a}}^j_{s\alpha}-\mathbf{a}^j_{s\alpha})'(\hat{\mathbf{a}}^j_{s\alpha}-\mathbf{a}^j_{s\alpha})\right]^2\right]^{\frac12}\left\{\frac1m\sum_{k=1}^m\sum_{\beta=1}^{q+1}\left[\sum_{j=1}^m\left(\hat{\mathbf{a}}^k_{r\beta}{}'\hat{\boldsymbol{\Gamma}}^x_{jk,s-r}{}'\hat{\boldsymbol{\Gamma}}^x_{jk,s-r}\hat{\mathbf{a}}^k_{r\beta}\right)^2\right]^{\frac12}\right\}\\ [3]\qquad&\leq\mathcal{A}\mathcal{B},\end{aligned}$$

where

$$\begin{aligned}\mathcal{A}&=m(q+1)^{\frac12}\left[\sum_{j=1}^m\sum_{\alpha=1}^{q+1}\left((\hat{\mathbf{a}}^j_{s\alpha}-\mathbf{a}^j_{s\alpha})'(\hat{\mathbf{a}}^j_{s\alpha}-\mathbf{a}^j_{s\alpha})\right)^2\right]^{\frac12}\\ \mathcal{B}&=\left\{\frac1m\sum_{k=1}^m\sum_{\beta=1}^{q+1}\left[\sum_{j=1}^m\left(\hat{\mathbf{a}}^k_{r\beta}{}'\hat{\boldsymbol{\Gamma}}^x_{jk,s-r}{}'\hat{\boldsymbol{\Gamma}}^x_{jk,s-r}\hat{\mathbf{a}}^k_{r\beta}\right)^2\right]^{\frac12}\right\}\\ [4]\qquad&\leq\left\{\frac{q+1}{m}\sum_{k=1}^m\sum_{\beta=1}^{q+1}\sum_{j=1}^m\left(\hat{\mathbf{a}}^k_{r\beta}{}'\hat{\boldsymbol{\Gamma}}^x_{jk,s-r}{}'\hat{\boldsymbol{\Gamma}}^x_{jk,s-r}\hat{\mathbf{a}}^k_{r\beta}\right)^2\right\}^{\frac12},\end{aligned}$$

call it $\mathcal{C}$. Inequalities [1] and [2] are obtained using the Cauchy-Schwartz inequality, [3] using the $c_r$ and Jensen inequalities for the terms on the right and left respectively, [4] using the $c_r$ inequality. Let us focus on $\mathcal{A}$. Let

$$\mathbf{a}_{s\alpha}^{j\ \prime} = (a_{s\alpha,1}^j\ a_{s\alpha,2}^j\ \cdots\ a_{s\alpha,q+1}^j),$$

and note that $\mathbf{a}_{s\alpha,\delta}^j = \mathbf{e}_\alpha'\mathbf{A}^{[j]}\mathbf{g}_{s\delta}$, where $\mathbf{e}_\alpha$ is the $(q+1)\times 1$ vector with unity in entry $\alpha$ and zero elsewhere, while $\mathbf{g}_{s\delta}$ is $(q+1)S\times 1$ with unity in entry $(s-1)(q+1)+\delta$ and zero elsewhere. Setting, for sake of simplicity,

$$\mathbf{B}_j = \mathbf{B}_j^\chi,\ \ \mathbf{C}_j = \mathbf{C}_{jj}^\chi,$$

where $\mathbf{B}_j^\chi, \mathbf{C}_{jj}^\chi$ are defined in (2.15), we have:

$$\left[\sum_{j=1}^m\sum_{\alpha=1}^{q+1}\left((\hat{\mathbf{a}}_{s\alpha}^j - \mathbf{a}_{s\alpha}^j)'(\hat{\mathbf{a}}_{s\alpha}^j - \mathbf{a}_{s\alpha}^j)\right)^2\right]^{\frac{1}{2}}$$

$$[5] \qquad \leq (q+1)^{\frac{1}{2}}\left(\sum_{j=1}^m\sum_{\alpha=1}^{q+1}\sum_{\delta=1}^{q+1}(\hat{\mathbf{a}}_{s\alpha,\delta}^j - \mathbf{a}_{s\alpha,\delta}^j)^4\right)^{\frac{1}{2}}$$

$$= (q+1)^{\frac{1}{2}}\left(\sum_{j=1}^m\sum_{\alpha=1}^{q+1}\sum_{\delta=1}^{q+1}\left[\mathbf{e}_\alpha\left((\hat{\mathbf{B}}_j - \mathbf{B}_j)(\hat{\mathbf{C}}^j)^{-1} + \mathbf{B}_j(\hat{\mathbf{C}}_j)^{-1}(\hat{\mathbf{C}}_j - \mathbf{C}_j)(\mathbf{C}_j)^{-1}\right)\mathbf{g}_{s\delta}\right]^4\right)^{\frac{1}{2}}$$

$$[6] \qquad \leq 2^{\frac{3}{2}}(q+1)^{\frac{3}{2}}\left(\sum_{j=1}^m \|(\hat{\mathbf{B}}_j - \mathbf{B}_j)(\hat{\mathbf{C}}_j)^{-1}\|^4 + \|\mathbf{B}_j(\hat{\mathbf{C}}_j)^{-1}(\hat{\mathbf{C}}_j - \mathbf{C}_j)(\mathbf{C}_j)^{-1}\|^4\right)^{\frac{1}{2}}$$

$$[7] \qquad \leq 2^{\frac{3}{2}}(q+1)^{\frac{3}{2}}\left(\left[\sum_{j=1}^m\|\hat{\mathbf{B}}_j - \mathbf{B}_j\|^8\right]^{\frac{1}{2}}\left[\sum_{j=1}^m\|(\hat{\mathbf{C}}_j)^{-1}\|^8\right]^{\frac{1}{2}} + \left[\sum_{j=1}^m\|\hat{\mathbf{C}}_j - \mathbf{C}_j\|^8\right]^{\frac{1}{2}}\left[\sum_{j=1}^m\|\hat{\mathbf{B}}_j(\hat{\mathbf{C}}_j)^{-1}\|^8\|(\mathbf{C}_j)^{-1}\|^8\right]^{\frac{1}{2}}\right)^{\frac{1}{2}}$$

$$[8] \qquad \leq 2^{\frac{3}{2}}(q+1)^{\frac{3}{2}}\left(\left[\sum_{j=1}^m\|\hat{\mathbf{B}}_j - \mathbf{B}_j\|^8\right]^{\frac{1}{2}}\left[\sum_{j=1}^m\|(\hat{\mathbf{C}}_j)^{-1}\|^8\right]^{\frac{1}{2}} + \left[\sum_{j=1}^m\|\hat{\mathbf{C}}_j - \mathbf{C}_j\|^8\right]^{\frac{1}{2}}\left[\sum_{j=1}^m\|\hat{\mathbf{B}}_j\|^{16}\right]^{\frac{1}{4}}\left[\sum_{j=1}^m\|(\hat{\mathbf{C}}_j)^{-1}\|^{16}\|(\mathbf{C}_j)^{-1}\|^{16}\right]^{\frac{1}{4}}\right)^{\frac{1}{2}},$$

where [5] has been obtained using the $c_r$ inequality, [6] using (C.1), the triangular and the $c_r$ inequality, [7] and [8] using the Cauchy-Schwartz inequality. Now, denoting by $b_{il}^j$ the entries of $\mathbf{B}_j$, $i = 1,\ldots,q+1$, $l = 1,\ldots,S(q+1)$, using the

$c_r$ inequality and Lemma 12,

$$\begin{aligned}\sum_{j=1}^{m}\|\hat{\mathbf{B}}_j-\mathbf{B}_j\|^8 &\leq \sum_{j=1}^{m}\left(\sum_{i=1}^{q+1}\sum_{l=1}^{S(q+1)}(\hat{b}_{il}^j-b_{il}^j)^2\right)^4\\ &\leq (q+1)^6S^3\sum_{j=1}^{m}\sum_{i=1}^{q+1}\sum_{l=1}^{S(q+1)}(\hat{b}_{il}^j-b_{il}^j)^8=O_P\left(m\left(\max\left[\frac{1}{n},\frac{1}{\rho_T}\right]\right)^4\right).\end{aligned}$$

In the same way

$$\sum_{j=1}^{m}\|\hat{\mathbf{C}}_j-\mathbf{C}_j\|^8=O_P\left(m\left(\max\left[\frac{1}{n},\frac{1}{\rho_T}\right]\right)^4\right).$$

We are left to consider terms such as $m^{-1}\sum_{i=1}^{m}\|\mathbf{C}_j\|^d$ for $d>0$. By a Taylor expansion around $\lambda_{S(q+1)}(\mathbf{C}_j^2)$, the minimum eigenvalue of $\mathbf{C}_j^2$:

$$\begin{aligned}&\frac{1}{m}\sum_{i=1}^{m}\|\hat{\mathbf{C}}_i^{-1}\|^d=\frac{1}{m}\sum_{i=1}^{m}\frac{1}{(\lambda_{S(q+1)}(\hat{\mathbf{C}}_i^2))^{\frac{d}{2}}}\\ &=\frac{1}{m}\sum_{i=1}^{m}\frac{1}{(\lambda_{S(q+1)}(\mathbf{C}_i^2))^{\frac{d}{2}}}+\sum_{k=1}^{\infty}\frac{(-1)^k\Gamma(d+k)}{k!\Gamma(d)}\left(\frac{1}{m}\sum_{i=1}^{m}\frac{(\lambda_{S(q+1)}(\hat{\mathbf{C}}_i^2)-\lambda_{S(q+1)}(\mathbf{C}_i^2))^k}{(\lambda_{S(q+1)}(\mathbf{C}_i^2))^k}\right)\\ &=\frac{1}{m}\sum_{i=1}^{m}\frac{1}{(\lambda_{S(q+1)}(\mathbf{C}_i^2))^{\frac{d}{2}}}+\sum_{k=1}^{\infty}\frac{(-1)^k\Gamma(d+k)}{k!\Gamma(d)}O_P(\delta^k)=\frac{1}{m}\sum_{i=1}^{m}\|\mathbf{C}_i^{-1}\|^d+O_P(1)=O_P(1),\end{aligned}$$

for $\delta$ satisfying $0<\zeta_{Tn}\leq\delta<1$ when $n,T$ are sufficiently large. In fact, $\lambda_{S(q+1)}(\mathbf{C}_i)$ is analytic in the constant dimensional matrix $\mathbf{C}_i$, thus by Lemma 4,

$$|\lambda_{S(q+1)}(\hat{\mathbf{C}}_i^2)-\lambda_{S(q+1)}(\mathbf{C}_i^2)|^d=O_P(\zeta_{Tn}^d)$$

for any $d>0$.

Collecting terms yields

$$\left[\sum_{j=1}^{m}\sum_{\alpha=1}^{q+1}\left((\hat{\mathbf{a}}_{s\alpha}^j-\mathbf{a}_{s\alpha}^j)'(\hat{\mathbf{a}}_{s\alpha}^j-\mathbf{a}_{s\alpha}^j)\right)^2\right]^{\frac{1}{2}}\leq(2(q+1))^{\frac{3}{2}}\left(\sum_{i=1}^{m}\|\hat{\mathbf{A}}_s^i-\mathbf{A}_s^i\|^4\right)^{\frac{1}{2}}=O_P\left(m^{\frac{1}{2}}\max\left(\frac{1}{n},\frac{1}{\rho_T}\right)\right) \quad \text{(E.9)}$$

Consider now $\mathcal{C}$.

$$\mathcal{C} \leq \left(\frac{q+1}{m}\right)^{\frac{1}{2}} \left\{ \left[ \sum_{k=1}^{m} \left( \sum_{\beta=1}^{q+1} \left( \hat{\mathbf{a}}_{r\beta}^{k\prime} \hat{\mathbf{a}}_{r\beta}^{k} \right)^2 \right)^2 \right]^{\frac{1}{2}} \left[ \sum_{j=1}^{m} \left( \sum_{k=1}^{m} \left( \text{trace} \left[ \hat{\mathbf{\Gamma}}_{jk,s-r}^{x\prime} \hat{\mathbf{\Gamma}}_{jk,s-r}^{x} \right] \right)^4 \right)^{\frac{1}{2}} \right] \right\}^{\frac{1}{2}}$$

$$\leq \left(\frac{q+1}{m}\right)^{\frac{1}{2}} \left\{ \left[ (q+1) \sum_{k=1}^{m} \sum_{\beta=1}^{q+1} \left( \hat{\mathbf{a}}_{r\beta}^{k\prime} \hat{\mathbf{a}}_{r\beta}^{k} \right)^4 \right]^{\frac{1}{2}} \left[ \frac{m}{m} \sum_{j=1}^{m} \left( \sum_{k=1}^{m} \left( \text{trace} \left[ \hat{\mathbf{\Gamma}}_{jk,s-r}^{x\prime} \hat{\mathbf{\Gamma}}_{jk,s-r}^{x} \right] \right)^4 \right)^{\frac{1}{2}} \right] \right\}^{\frac{1}{2}}$$

$$\leq (q+1)^{\frac{1}{2}} \left[ (q+1)^4 \sum_{k=1}^{m} \sum_{\alpha=1}^{q+1} \sum_{\beta=1}^{q+1} (\hat{a}_{r,\alpha\beta}^{k})^8 \right]^{\frac{1}{4}} \left[ \frac{1}{m} \sum_{j=1}^{m} \sum_{k=1}^{m} \left( \text{trace} \left[ \hat{\mathbf{\Gamma}}_{jk,s-r}^{x\prime} \hat{\mathbf{\Gamma}}_{jk,s-r}^{x} \right] \right)^4 \right]^{\frac{1}{4}}$$

$$\leq (q+1)^{\frac{3}{2}} \left[ \sum_{k=1}^{m} \sum_{\alpha=1}^{q+1} \sum_{\beta=1}^{q+1} (\hat{a}_{r,\alpha\beta}^{k})^8 \right]^{\frac{1}{4}} \left[ \frac{(q+1)^6}{m} \sum_{j=1}^{m} \sum_{k=1}^{m} \sum_{\alpha=1}^{q+1} \sum_{\beta=1}^{q+1} (\hat{\gamma}_{jk,\alpha\beta}^{x}(s-r))^8 \right]^{\frac{1}{4}} = O_P(m^{\frac{1}{2}}),$$

where $\hat{\gamma}_{jk,\alpha\beta}^{x}(s-r)$ denotes the $(\alpha,\beta)$-th entry of $\hat{\mathbf{\Gamma}}_{jk,s-r}^{x}$. The first inequality uses $\mathbf{a}'\mathbf{A}\mathbf{a} \leq (\mathbf{a}'\mathbf{a})\text{trace}(\mathbf{A})$, the second uses the Cauchy-Schwarz inequality, the third uses the $c_r$ inequality, the fourth uses the $c_r$ and Jensen inequalities, the fifth and the sixth use the $c_r$ inequality. Collecting terms

$$\frac{1}{m} \| \hat{\mathbf{A}}_s \hat{\mathbf{\Gamma}}_{s-r}^{x} \hat{\mathbf{A}}_r' - \mathbf{A}_s \hat{\mathbf{\Gamma}}_{s-r}^{x} \mathbf{A}_r' \| = O_P\left(\zeta_{Tn}\right), \quad r, s = 0, ..., p.$$

Consider now the second statement of the lemma. The two terms on the right hand side of (E.8) must be dealt with separately since there is only one summation ranging from 1 to $n$. We can assume, without loss of generality, that $s$, the number of elements selected by $\mathcal{S}_{\mathbf{M}}$, $s = m^*(q+1)$ for a given $m^*$. Regarding the first term of (E.8), substitution of the summation $\sum_{k=1}^{m}$ with $\sum_{k=1}^{m^*}$ gives

$$\sum_{j=1}^{m} \sum_{k=1}^{m^*} \sum_{\alpha=1}^{q+1} \sum_{\beta=1}^{q+1} \left( (\hat{\mathbf{a}}_{s\alpha}^{j} - \mathbf{a}_{s\alpha}^{j})' \hat{\mathbf{\Gamma}}_{jk,s-r}^{x} \hat{\mathbf{a}}_{r\beta}^{k} \right)^2 = O_P\left( m \left( \max\left( \frac{1}{n}, \frac{1}{\rho_T} \right) \right) \right).$$

For, the term on the left hand side is bounded by $\mathcal{D}\mathcal{E}$ where:

$$\mathcal{D} = m^{\frac{1}{2}} (q+1)^{\frac{1}{2}} \left[ \sum_{j=1}^{m} \sum_{\alpha=1}^{q+1} \left( (\hat{\mathbf{a}}_{s\alpha}^{j} - \mathbf{a}_{s\alpha}^{j})' (\hat{\mathbf{a}}_{s\alpha}^{j} - \mathbf{a}_{s\alpha}^{j}) \right)^2 \right]^{\frac{1}{2}}$$

$$\mathcal{E} = \left\{ \sum_{k=1}^{m^*} \sum_{\beta=1}^{q+1} \left( \frac{1}{m} \sum_{j=1}^{m} \left( \hat{\mathbf{a}}_{r\beta}^{k\prime} \hat{\mathbf{\Gamma}}_{jk,s-r}^{x\prime} \hat{\mathbf{\Gamma}}_{jk,s-r}^{x} \hat{\mathbf{a}}_{r\beta}^{k} \right)^2 \right)^{\frac{1}{2}} \right\},$$

while, in turn, $\mathcal{D}$ and $\mathcal{E}$ can be bounded in the same way used to bound the terms $\mathcal{A}$ and $\mathcal{B}$ in the proof of the first first statement.

Consider now the second term on the right hand side of $(E.8)$. By arguments already used in first part of the proof,

$$\sum_{k=1}^{m^*}\sum_{j=1}^{m}\sum_{\alpha=1}^{q+1}\sum_{\beta=1}^{q+1}\left((\hat{\mathbf{a}}_{s\alpha}^{k}-\mathbf{a}_{s\alpha}^{k})'\hat{\mathbf{\Gamma}}_{jk,s-r}^{x\prime}\mathbf{a}_{r\beta}^{j}\right)^2$$
$$\leq m\left[\sum_{k=1}^{m^*}\left[\sum_{\alpha=1}^{q+1}(\hat{\mathbf{a}}_{s\alpha}^{k}-\mathbf{a}_{s\alpha}^{k})'(\hat{\mathbf{a}}_{s\alpha}^{k}-\mathbf{a}_{s\alpha}^{k})\right]^2\right]^{\frac{1}{2}}\left\{\frac{1}{m}\sum_{j=1}^{m}\sum_{\beta=1}^{q+1}\left[\sum_{k=1}^{m^*}\left(\mathbf{a}_{r\beta}^{j}{}'\hat{\mathbf{\Gamma}}_{jk,s-r}^{x}\hat{\mathbf{\Gamma}}_{jk,s-r}^{x\prime}\mathbf{a}_{r\beta}^{j}\right)^2\right]^{\frac{1}{2}}\right\}$$
$$\equiv \mathcal{F}\mathcal{G}.$$

Now, $\mathcal{F}=O_P(m\,\zeta_{TN}^2)$ easily follows from Proposition 4, while $\mathcal{G}=O_P(1)$ can be obtained using the arguments used to bound $\mathcal{C}$ in the proof of the first statement. Collecting terms

$$\frac{1}{\sqrt{m}}\|\mathcal{S}_{\mathbf{M}}'(\hat{\mathbf{A}}_s\hat{\mathbf{\Gamma}}_{s-r}^{x}\hat{\mathbf{A}}_r'-\mathbf{A}_s\hat{\mathbf{\Gamma}}_{s-r}^{x}\mathbf{A}_r')\|=O_p\left(\zeta_{Tn}\right),\quad r,s=0,...,p.$$

Q.E.D.

# F Proof of Proposition 6

Consider

$$\hat{\mathbf{v}}_t = ((\hat{\mathbf{\Lambda}}^z)^{\frac{1}{2}}\hat{\mathbf{P}}^{z\prime}\hat{\mathbf{P}}^z(\hat{\mathbf{\Lambda}}^z)^{\frac{1}{2}})^{-1}(\hat{\mathbf{\Lambda}}^z)^{\frac{1}{2}}\hat{\mathbf{P}}^{z\prime}\hat{\mathbf{z}}_t = (\hat{\mathbf{\Lambda}}^z)^{-\frac{1}{2}}\hat{\mathbf{P}}^{z\prime}\hat{\mathbf{z}}_t$$
$$= (\hat{\mathbf{\Lambda}}^z)^{-\frac{1}{2}}\hat{\mathbf{P}}^{z\prime}\left(\hat{\mathbf{A}}(L) - \mathbf{A}(L)\right)\mathbf{x}_t + \left((\hat{\mathbf{\Lambda}}^z)^{-\frac{1}{2}}\hat{\mathbf{P}}^{z\prime} - \hat{\mathbf{W}}^z(\mathbf{\Lambda}^\psi)^{-\frac{1}{2}}\mathbf{P}^{\psi\prime}\right)\mathbf{A}(L)\mathbf{x}_t$$
$$+\hat{\mathbf{W}}^z(\mathbf{\Lambda}^\psi)^{-\frac{1}{2}}\mathbf{P}^{\psi\prime}\mathbf{A}(L)\xi_t + \hat{\mathbf{W}}^z(\mathbf{\Lambda}^\psi)^{-\frac{1}{2}}\mathbf{P}^{\psi\prime}\mathbf{P}^\psi(\mathbf{\Lambda}^\psi)^{\frac{1}{2}}\mathbf{v}_t \qquad \text{(F.10)}$$

Regarding the first term on the right hand side of (F.10),

$$\|(\hat{\mathbf{\Lambda}}^z)^{-\frac{1}{2}}\hat{\mathbf{P}}^{z\prime}\left(\hat{\mathbf{A}}(L) - \mathbf{A}(L)\right)\mathbf{x}_t\| = \|(\frac{\hat{\mathbf{\Lambda}}^z}{n})^{-\frac{1}{2}}\hat{\mathbf{P}}^{z\prime}\frac{\left(\hat{\mathbf{A}}(L) - \mathbf{A}(L)\right)\mathbf{x}_t}{n^{\frac{1}{2}}}\|$$
$$\leq \|(\frac{\hat{\mathbf{\Lambda}}^z}{n})^{-\frac{1}{2}}\|\|\hat{\mathbf{P}}^{z\prime}\|\|\frac{(\hat{\mathbf{A}}(L) - \mathbf{A}(L))\mathbf{x}_t}{n^{\frac{1}{2}}}\|.$$

Since $\|(\hat{\mathbf{\Lambda}}^z/n)^{-\frac{1}{2}}\| = O_P(1)$ and $\|\hat{\mathbf{P}}^z\| = 1$, by (E.9) one gets

$$\|\frac{(\hat{\mathbf{A}}(L) - \mathbf{A}(L))\mathbf{x}_t}{n^{\frac{1}{2}}}\| \leq \sum_{r=0}^{p}\left(n^{-1}\sum_{i=1}^{m}\mathbf{x}^{i\prime}_{t-r}(\hat{\mathbf{A}}^i_r - \mathbf{A}^i_r)'(\hat{\mathbf{A}}^i_r - \mathbf{A}^i_r)\mathbf{x}^i_{t-r}\right)^{\frac{1}{2}}$$
$$\leq \sum_{r=0}^{p}\left(n^{-1}\sum_{i=1}^{m}(\mathbf{x}^{i\prime}_{t-r}\mathbf{x}^i_{t-r})^2\right)^{\frac{1}{4}}\left(n^{-1}\sum_{i=1}^{m}\left(\sum_{j=1}^{q+1}\sum_{h=1}^{q+1}(\hat{a}^i_{r,jh} - a^i_{r,jh})^2\right)^2\right)^{\frac{1}{4}}$$
$$\leq \sum_{r=0}^{p}\left(n^{-1}\sum_{i=1}^{m}(\mathbf{x}^{i\prime}_{t-r}\mathbf{x}^i_{t-r})^2\right)^{\frac{1}{4}}\left((q+1)^3 n^{-1}\sum_{i=1}^{m}\|\hat{\mathbf{A}}^i_r - \mathbf{A}^i_r\|^4\right)^{\frac{1}{4}}$$
$$= O_P\left(\zeta_{Tn}\right)$$

setting $\mathbf{x}_t = (\mathbf{x}^{1\prime}_t...\mathbf{x}^{i\prime}_t...\mathbf{x}^{m\prime}_t)'$ for sub-vectors $\mathbf{x}^i_t$ of size $q+1 \times 1$.

Regarding the second term on the right hand side of (F.10)

$$\|\left((\hat{\mathbf{\Lambda}}^z)^{-\frac{1}{2}}\hat{\mathbf{P}}^{z\prime} - \hat{\mathbf{W}}^z(\mathbf{\Lambda}^\psi)^{-\frac{1}{2}}\mathbf{P}^{\psi\prime}\right)\mathbf{A}(L)\mathbf{x}_t\|$$
$$= \|(\hat{\mathbf{\Lambda}}^z)^{-1}\hat{\mathbf{\Lambda}}^z\left((\hat{\mathbf{\Lambda}}^z)^{-\frac{1}{2}}\hat{\mathbf{P}}^{z\prime} - \hat{\mathbf{W}}^z(\mathbf{\Lambda}^\psi)^{-\frac{1}{2}}\mathbf{P}^{\psi\prime}\right)\mathbf{A}(L)\mathbf{x}_t\|$$
$$\|(\frac{\hat{\mathbf{\Lambda}}^z}{n})^{-1}\hat{\mathbf{\Lambda}}^z\left((\hat{\mathbf{\Lambda}}^z)^{-\frac{1}{2}}\hat{\mathbf{P}}^{z\prime} - \hat{\mathbf{W}}^z(\mathbf{\Lambda}^\psi)^{-\frac{1}{2}}\mathbf{P}^{\psi\prime}\right)\frac{\mathbf{A}(L)\mathbf{x}_t}{n}\|$$
$$= \|(\frac{\hat{\mathbf{\Lambda}}^z}{n})^{-1}\left((\hat{\mathbf{\Lambda}}^z)^{\frac{1}{2}}\hat{\mathbf{P}}^{z\prime} - \hat{\mathbf{W}}^z\hat{\mathbf{\Lambda}}^z(\mathbf{\Lambda}^\psi)^{-\frac{1}{2}}\mathbf{P}^{\psi\prime}\right)\frac{\mathbf{A}(L)\mathbf{x}_t}{n}\|$$
$$= \|(\frac{\hat{\mathbf{\Lambda}}^z}{n})^{-1}\left((\hat{\mathbf{\Lambda}}^z)^{\frac{1}{2}}\hat{\mathbf{P}}^{z\prime} - \hat{\mathbf{W}}^z[\hat{\mathbf{\Lambda}}^z - \mathbf{\Lambda}^\psi + \mathbf{\Lambda}^\psi](\mathbf{\Lambda}^\psi)^{-\frac{1}{2}}\mathbf{P}^{\psi\prime}\right)\frac{\mathbf{A}(L)\mathbf{x}_t}{n}\|$$
$$\leq \|(\frac{\hat{\mathbf{\Lambda}}^z}{n})^{-1}\|\|\left((\hat{\mathbf{\Lambda}}^z)^{\frac{1}{2}}\hat{\mathbf{P}}^{z\prime} - \hat{\mathbf{W}}^z(\mathbf{\Lambda}^\psi)^{\frac{1}{2}}\mathbf{P}^{\psi\prime}\right)\|\|\frac{\mathbf{A}(L)\mathbf{x}_t}{n}\|$$
$$+\|(\frac{\hat{\mathbf{\Lambda}}^z}{n})^{-1}\|\|\hat{\mathbf{W}}^z(\hat{\mathbf{\Lambda}}^z - \mathbf{\Lambda}^\psi)(\mathbf{\Lambda}^\psi)^{-\frac{1}{2}}\mathbf{P}^{\psi\prime}\|\|\frac{\mathbf{A}(L)\mathbf{x}_t}{n}\|$$
$$= O_P\left(\zeta_{Tn}\right)$$

since by (E.3)

$$\|(\hat{\mathbf{P}}^z(\hat{\mathbf{\Lambda}}^z)^{\frac{1}{2}} - \mathbf{P}^\psi(\mathbf{\Lambda}^\psi)^{\frac{1}{2}}\hat{\mathbf{W}}^z)\| = O_P\left(n^{\frac{1}{2}}\zeta_{Tn}\right),$$

and

$$\begin{aligned}
&\|\frac{\hat{\mathbf{A}}(L)\mathbf{x}_t}{n}\| = n^{-\frac{1}{2}}\left(\frac{\mathbf{x}_t'\hat{\mathbf{A}}'(L)\hat{\mathbf{A}}(L)\mathbf{x}_t}{n}\right)^{\frac{1}{2}} \\
&\leq n^{-\frac{1}{2}}\sum_{r=0}^{p}\left(\frac{\mathbf{x}_{t-r}'\hat{\mathbf{A}}_r'\hat{\mathbf{A}}_r\mathbf{x}_{t-r}}{n}\right)^{\frac{1}{2}} \leq n^{-\frac{1}{2}}\sum_{r=0}^{p}\left(\frac{\mathbf{x}_{t-r}'\hat{\mathbf{A}}_r'\hat{\mathbf{A}}_r\mathbf{x}_{t-r}}{n}\right)^{\frac{1}{2}} \\
&\leq n^{-\frac{1}{2}}\sum_{r=0}^{p}(\mathbf{x}_{t-r}'\mathbf{x}_{t-r}/n)^{\frac{1}{2}}(\lambda_1(\hat{\mathbf{A}}_r'\hat{\mathbf{A}}_r))^{\frac{1}{2}} = O_P(n^{-\frac{1}{2}}).
\end{aligned}$$

For the third term on the right hand side of (F.10)

$$(\mathbf{\Lambda}^\psi)^{-\frac{1}{2}}\mathbf{P}^{\psi\prime}\mathbf{A}(L)\xi_t = O_P(n^{-\frac{1}{2}}).$$

Finally, note that $\hat{\mathbf{W}}^z(\mathbf{\Lambda}^\psi)^{-\frac{1}{2}}\mathbf{P}^{\psi\prime}\mathbf{P}^\psi(\mathbf{\Lambda}^\psi)^{\frac{1}{2}}\mathbf{v}_t = \hat{\mathbf{W}}^z\mathbf{v}_t$. Q.E.D.

# G Data description

**Quarterly data**. Most series are taken from the FRED data base. A few stock market and leading indicators are taken from Data Stream. Some series have been constructed as transformations of the original FRED series. Monthly data have been temporally aggregated to get quarterly figures. Outliers are treated as in Stock and Watson (2002b). Transformations: 1 = levels, 2 = first differences of the original series, 5 = first differences of logs of the original series, 6 = second differences of logs of the original series.

| no.series | Transf. | Mnemonic | Long Label |
|---|---|---|---|
| 1 | 5 | GDPC1 | Real Gross Domestic Product, 1 Decimal |
| 2 | 5 | GNPC96 | Real Gross National Product |
| 3 | 5 | NICUR/GDPDEF | National Income/GDPDEF |
| 4 | 5 | DPIC96 | Real Disposable Personal Income |
| 5 | 5 | OUTNFB | Nonfarm Business Sector: Output |
| 6 | 5 | FINSLC1 | Real Final Sales of Domestic Product, 1 Decimal |
| 7 | 5 | FPIC1 | Real Private Fixed Investment, 1 Decimal |
| 8 | 5 | PRFIC1 | Real Private Residential Fixed Investment, 1 Decimal |
| 9 | 5 | PNFIC1 | Real Private Nonresidential Fixed Investment, 1 Decimal |
| 10 | 5 | GPDIC1 | Real Gross Private Domestic Investment, 1 Decimal |
| 11 | 5 | PCECC96 | Real Personal Consumption Expenditures |
| 12 | 5 | PCNDGC96 | Real Personal Consumption Expenditures: Nondurable Goods |
| 13 | 5 | PCDGCC96 | Real Personal Consumption Expenditures: Durable Goods |
| 14 | 5 | PCESVC96 | Real Personal Consumption Expenditures: Services |
| 15 | 5 | GPSAVE/GDPDEF | Gross Private Saving/GDP Deflator |
| 16 | 5 | FGCEC1 | Real Federal Consumption Expenditures & Gross Investment, 1 Decimal |
| 17 | 5 | FGEXPND/GDPDEF | Federal Government: Current Expenditures/ GDP deflator |
| 18 | 5 | FGRECPT/GDPDEF | Federal Government Current Receipts/ GDP deflator |
| 19 | 2 | FGDEF | Federal Real Expend-Real Receipts |
| 20 | 1 | CBIC1 | Real Change in Private Inventories, 1 Decimal |
| 21 | 5 | EXPGSC1 | Real Exports of Goods & Services, 1 Decimal |
| 22 | 5 | IMPGSC1 | Real Imports of Goods & Services, 1 Decimal |
| 23 | 5 | CP/GDPDEF | Corporate Profits After Tax/GDP deflator |
| 24 | 5 | NFCPATAX/GDPDEF | Nonfinancial Corporate Business: Profits After Tax/GDP deflator |
| 25 | 5 | CNCF/GDPDEF | Corporate Net Cash Flow/GDP deflator |
| 26 | 5 | DIVIDEND/GDPDEF | Net Corporate Dividends/GDP deflator |
| 27 | 5 | HOANBS | Nonfarm Business Sector: Hours of All Persons |
| 28 | 5 | OPHNFB | Nonfarm Business Sector: Output Per Hour of All Persons |
| 29 | 5 | UNLPNBS | Nonfarm Business Sector: Unit Nonlabor Payments |
| 30 | 5 | ULCNFB | Nonfarm Business Sector: Unit Labor Cost |
| 31 | 5 | WASCUR/CPI | Compensation of Employees: Wages & Salary Accruals/CPI |
| 32 | 6 | COMPNFB | Nonfarm Business Sector: Compensation Per Hour |
| 33 | 5 | COMPRNFB | Nonfarm Business Sector: Real Compensation Per Hour |
| 34 | 6 | GDPCTPI | Gross Domestic Product: Chain-type Price Index |
| 35 | 6 | GNPCTPI | Gross National Product: Chain-type Price Index |
| 36 | 6 | GDPDEF | Gross Domestic Product: Implicit Price Deflator |
| 37 | 6 | GNPDEF | Gross National Product: Implicit Price Deflator |
| 38 | 5 | INDPRO | Industrial Production Index |
| 39 | 5 | IPBUSEQ | Industrial Production: Business Equipment |
| 40 | 5 | IPCONGD | Industrial Production: Consumer Goods |
| 41 | 5 | IPDCONGD | Industrial Production: Durable Consumer Goods |
| 42 | 5 | IPFINAL | Industrial Production: Final Products (Market Group) |
| 43 | 5 | IPMAT | Industrial Production: Materials |
| 44 | 5 | IPNCONGD | Industrial Production: Nondurable Consumer Goods |
| 45 | 2 | AWHMAN | Average Weekly Hours: Manufacturing |
| 46 | 2 | AWOTMAN | Average Weekly Hours: Overtime: Manufacturing |
| 47 | 2 | CIVPART | Civilian Participation Rate |
| 48 | 5 | CLF16OV | Civilian Labor Force |
| 49 | 5 | CE16OV | Civilian Employment |
| 50 | 5 | USPRIV | All Employees: Total Private Industries |
| 51 | 5 | USGOOD | All Employees: Goods-Producing Industries |
| 52 | 5 | SRVPRD | All Employees: Service-Providing Industries |
| 53 | 5 | UNEMPLOY | Unemployed |
| 54 | 5 | UEMPMEAN | Average (Mean) Duration of Unemployment |
| 55 | 2 | UNRATE | Civilian Unemployment Rate |
| 56 | 5 | HOUST | Housing Starts: Total: New Privately Owned Housing Units Started |

| | | | |
|---|---|---|---|
| 57 | 2 | FEDFUNDS | Effective Federal Funds Rate |
| 58 | 2 | TB3MS | 3-Month Treasury Bill: Secondary Market Rate |
| 59 | 2 | GS1 | 1-Year Treasury Constant Maturity Rate |
| 60 | 2 | GS10 | 10-Year Treasury Constant Maturity Rate |
| 61 | 2 | AAA | Moody's Seasoned Aaa Corporate Bond Yield |
| 62 | 2 | BAA | Moody's Seasoned Baa Corporate Bond Yield |
| 63 | 2 | MPRIME | Bank Prime Loan Rate |
| 64 | 6 | BOGNONBR | Non-Borrowed Reserves of Depository Institutions |
| 65 | 6 | TRARR | Board of Governors Total Reserves, Adjusted for Changes in Reserve |
| 66 | 6 | BOGAMBSL | Board of Governors Monetary Base, Adjusted for Changes in Reserve |
| 67 | 6 | M1SL | M1 Money Stock |
| 68 | 6 | M2MSL | M2 Minus |
| 69 | 6 | M2SL | M2 Money Stock |
| 70 | 6 | BUSLOANS | Commercial and Industrial Loans at All Commercial Banks |
| 71 | 6 | CONSUMER | Consumer (Individual) Loans at All Commercial Banks |
| 72 | 6 | LOANINV | Total Loans and Investments at All Commercial Banks |
| 73 | 6 | REALLN | Real Estate Loans at All Commercial Banks |
| 74 | 6 | TOTALSL | Total Consumer Credit Outstanding |
| 75 | 6 | CPIAUCSL | Consumer Price Index For All Urban Consumers: All Items |
| 76 | 6 | CPIULFSL | Consumer Price Index for All Urban Consumers: All Items Less Food |
| 77 | 6 | CPILEGSL | Consumer Price Index for All Urban Consumers: All Items Less Energy |
| 78 | 6 | CPILFESL | Consumer Price Index for All Urban Consumers: All Items Less Food & Energy |
| 79 | 6 | CPIENGSL | Consumer Price Index for All Urban Consumers: Energy |
| 80 | 6 | CPIUFDSL | Consumer Price Index for All Urban Consumers: Food |
| 81 | 6 | PPICPE | Producer Price Index Finished Goods: Capital Equipment |
| 82 | 6 | PPICRM | Producer Price Index: Crude Materials for Further Processing |
| 83 | 6 | PPIFCG | Producer Price Index: Finished Consumer Goods |
| 84 | 6 | PPIFGS | Producer Price Index: Finished Goods |
| 85 | 6 | OILPRICE | Spot Oil Price: West Texas Intermediate |
| 86 | 5 | USSHRPRCF | US Dow Jones Industrials Share Price Index (EP) NADJ |
| 87 | 5 | US500STK | US Standard & Poor's Index of 500 Common Stocks |
| 88 | 5 | USI62...F | US Share Price Index NADJ |
| 89 | 5 | USNOIDN.D | US Manufacturers New Orders for Non Defense Capital Goods (BCI 27) |
| 90 | 5 | USCNORCGD | US New Orders of Consumer Goods & Materials (BCI 8) CONA |
| 91 | 1 | USNAPMNO | US ISM Manufacturers Survey: New Orders Index SADJ |
| 92 | 5 | USVACTOTO | US Index of Help Wanted Advertising VOLA |
| 93 | 5 | USCYLEAD | US The Conference Board Leading Economic Indicators Index SADJ |
| 94 | 5 | USECRIWLH | US Economic Cycle Research Institute Weekly Leading Index |
| 95 | 2 | GS10-FEDFUNDS | |
| 96 | 2 | GS1-FEDFUNDS | |
| 97 | 2 | BAA-FEDFUNDS | |
| 98 | 5 | GEXPND/GDPDEF | Government Current Expenditures/ GDP deflator |
| 99 | 5 | GRECPT/GDPDEF | Government Current Receipts/ GDP deflator |
| 100 | 2 | GDEF | Governnent Real Expend-Real Receipts |
| 101 | 5 | GCEC1 | Real Government Consumption Expenditures & Gross Investment, 1 Decimal |

**Monthly data**. Most series are those of the Stock-Watson data set used in Bernanke *et al.* (2005). A few real exchange rates and short-term interest rate spreads between US and some foreign countries are added, and some discontinued series are eliminated. The basic source is the FRED data base; some series have been constructed as transformations of the original series. Outliers are treated as in Stock and Watson (2002b). Transformations: 1 = levels, 4 = logs, 5 = first differences of logs of the original series.

| no.series | Mnemonic | Long Label | Transformation |
|---|---|---|---|
| 1 | DSPIC96 | Real Disposable Personal Income | 5 |
| 2 | A0M051 | Personal Income Less Transfer Payments | 5 |
| 3 | PCEC96 | Real Personal Consumption Expenditures | 5 |
| 4 | A0M059 | Sales, Orders, And Deliveries, Sales, Retail Stores | 5 |
| 5 | IPS10 | Industrial Production Index - Total Index | 5 |
| 6 | IPS11 | Industrial Production Index - Products, Total | 5 |
| 7 | IPS12 | Industrial Production Index - Consumer Goods | 5 |
| 8 | IPS13 | Industrial Production Index - Durable Consumer Goods | 5 |
| 9 | IPS18 | Industrial Production Index - Nondurable Consumer Goods | 5 |
| 10 | IPS25 | Industrial Production Index - Business Equipment | 5 |
| 11 | IPS299 | Industrial Production Index - Final Products | 5 |
| 12 | IPS306 | Industrial Production Index - Fuels | 5 |
| 13 | IPS307 | Industrial Production Index - Residential Utilities | 5 |
| 14 | IPS32 | Industrial Production Index - Materials | 5 |
| 15 | IPS34 | Industrial Production Index - Durable Goods Materials | 5 |
| 16 | IPS38 | Industrial Production Index - Nondurable Goods Materials | 5 |
| 17 | IPS43 | Industrial Production Index - Manufacturing (SIC) | 5 |
| 18 | PMP | NAPM Production Index (Percent) | 1 |
| 19 | MCUMFN | Capacity Utilization: Manufacturing (NAICS) | 1 |
| 20 | LHEL | Index Of Help-Wanted Advertising In Newspapers | 5 |
| 21 | LHELX | Employment: Ratio; Help-Wanted | 4 |
| 22 | LHEM | Civilian Labor Force: Employed, Total | 5 |
| 23 | LHNAG | Civilian Labor Force: Employed, Nonagric.Industries | 5 |
| 24 | LHU14 | Unemploy.By Duration: Persons Unempl.5 To 14 Wks | 1 |
| 25 | LHU15 | Unemploy.By Duration: Persons Unempl.15 Wks + | 1 |
| 26 | LHU26 | Unemploy.By Duration: Persons Unempl.15 To 26 Wks | 1 |
| 27 | LHU27 | Unemploy.By Duration: Persons Unempl.27 Wks + | 1 |
| 28 | LHU5 | Unemploy.By Duration: Persons Unempl.Less Than 5 Wks | 1 |
| 29 | LHU680 | Unemploy.By Duration: Average(Mean)Duration In Weeks | 1 |
| 30 | LHUR | Unemployment Rate: All Workers, 16 Years & Over (%,SA) | 1 |
| 31 | CES002 | Employees On Nonfarm Payrolls - Total Private | 5 |
| 32 | CES003 | Employees On Nonfarm Payrolls - Goods-Producing | 5 |
| 33 | CES006 | Employees On Nonfarm Payrolls - Mining, Thousands | 5 |
| 34 | CES011 | Employees On Nonfarm Payrolls - Construction | 5 |
| 35 | CES015 | Employees On Nonfarm Payrolls - Manufacturing | 5 |
| 36 | CES017 | Employees On Nonfarm Payrolls - Durable Goods | 5 |
| 37 | CES033 | Employees On Nonfarm Payrolls - Nondurable Goods | 5 |
| 38 | CES046 | Employees On Nonfarm Payrolls - Service-Providing | 5 |
| 39 | CES048 | Employees On Nonfarm Payrolls - Trade, Transp., Utilities | 5 |
| 40 | CES049 | Employees On Nonfarm Payrolls - Wholesale Trade | 5 |
| 41 | CES053 | Employees On Nonfarm Payrolls - Retail Trade | 5 |
| 42 | CES088 | Employees On Nonfarm Payrolls - Financial Activities | 5 |
| 43 | CES140 | Employees On Nonfarm Payrolls - Government | 5 |
| 44 | AWHI | Aggregate Weekly Hours Index: Total Private Industries | 5 |
| 45 | CES151 | Average Weekly Hours Goods-Producing | 1 |
| 46 | CES155 | Average Weekly Hours Manufacturing Overtime Hours | 1 |
| 47 | AWHMAN | Average Weekly Hours: Manufacturing | 1 |
| 48 | PMEMP | Napm Employment Index (Percent) | 1 |
| 49 | HSBMW | Houses Authorized By Build. Permits:Midwest | 4 |
| 50 | HSBNE | Houses Authorized By Build. Permits:Northeast | 4 |
| 51 | HSBR | Housing Authorized: Total New Priv Housing Units | 4 |
| 52 | HSBSOU | Houses Authorized By Build. Permits:South | 4 |
| 53 | HSBWST | Houses Authorized By Build. Permits:West | 4 |
| 54 | HSFR | Housing Starts:Nonfarm (1947-58);Total Farm&Nonfarm(1959-) | 4 |
| 55 | HSMW | Housing Starts:Midwest | 4 |

| no.series | Mnemonic | Long Label | Transformation |
|---|---|---|---|
| 56 | HSNE | Housing Starts:Northeast | 4 |
| 57 | HSSOU | Housing Starts:South | 4 |
| 58 | HSWST | Housing Starts:West | 4 |
| 59 | PMDEL | Napm Vendor Deliveries Index | 1 |
| 60 | PMI | Purchasing Managers' Index | 1 |
| 61 | PMNO | Napm New Orders Index | 1 |
| 62 | PMNV | Napm Inventories Index | 1 |
| 63 | A0M007 | New Orders, Durable Goods Industries | 5 |
| 64 | A0M027 | New Orders, Capital Goods Industries, Nondefense | 5 |
| 65 | A1M092 | Manufacturers' Unfilled Orders, Durable Goods Industries | 5 |
| 66 | FM1 | Money Stock: M1 | 5 |
| 67 | FM2 | Money Stock:M2 | 5 |
| 68 | FMFBA | Monetary Base, Adj For Reserve Requirement Changes | 5 |
| 69 | FMRNBA | Depository Inst Reserves:Nonborrowed,Adj Res Req Chgs | 5 |
| 70 | FMRRA | Depository Inst Reserves:Total,Adj For Reserve Req Chgs | 5 |
| 71 | FCLBMC | Wkly Rp Lg Com'L Banks:Net Change Com'L & Indus Loans | 1 |
| 72 | CCINRV | Consumer Credit Outstanding - Nonrevolving(G19) | 5 |
| 73 | FSPCOM | S&P'S Common Stock Price Index: Composite | 5 |
| 74 | FSPIN | S&P'S Common Stock Price Index: Industrials | 5 |
| 75 | FYFF | Interest Rate: Federal Funds (Effective) | 1 |
| 76 | FYGM3 | Interest Rate: U.S.Treasury Bills,Sec Mkt,3-Mo. | 1 |
| 77 | FYGM6 | Interest Rate: U.S.Treasury Bills,Sec Mkt,6-Mo.0 | 1 |
| 78 | FYGT1 | Interest Rate: U.S.Treasury Const Maturities,1-Yr. | 1 |
| 79 | FYGT10 | Interest Rate: U.S.Treasury Const Maturities,10-Yr. | 1 |
| 80 | FYGT5 | Interest Rate: U.S.Treasury Const Maturities,5-Yr. | 1 |
| 81 | FYAAAC | Bond Yield: Moody'S Aaa Corporate | 1 |
| 82 | FYBAAC | Bond Yield: Moody'S Baa Corporate | 1 |
| 83 | EXRUS | United States;Effective Exchange Rate (MERM) | 5 |
| 84 | EXRCAN | Foreign Exchange Rate: Canada (Canadian $ Per U.S.$) | 5 |
| 85 | EXRJAN | Foreign Exchange Rate: Japan (Yen Per U.S.$) | 5 |
| 86 | EXRSW | Foreign Exchange Rate: Switzerland (Swiss Franc Per U.S.$) | 5 |
| 87 | EXRUK | Foreign Exchange Rate: United Kingdom (Cents Per Pound) | 5 |
| 88 | PWFCSA | Producer Price Index:Finished Consumer Goods | 5 |
| 89 | PWFSA | Producer Price Index: Finished Goods | 5 |
| 90 | PWCMSA | Producer Price Index:Crude Materials | 5 |
| 91 | PWIMSA | Producer Price Index:Intermed Mat.Supplies & Components | 5 |
| 92 | PMCP | Napm Commodity Prices Index | 1 |
| 93 | PU83 | CPI-U: Apparel & Upkeep | 5 |
| 94 | PU84 | CPI-U: Transportation | 5 |
| 95 | PU85 | CPI-U: Medical Care | 5 |
| 96 | PUNEW | CPI-U: All Items | 5 |
| 97 | PUC | CPI-U: Commodities | 5 |
| 98 | PUCD | CPI-U: Durables | 5 |
| 99 | PUS | CPI-U: Services | 5 |
| 100 | PUXF | CPI-U: All Items Less Food | 5 |
| 101 | PUXHS | CPI-U: All Items Less Shelter | 5 |
| 102 | PUXM | CPI-U: All Items Less Medical Care | 5 |
| 103 | CES277 | Average Hourly Earnings - Construction | 5 |
| 104 | CES278 | Average Hourly Earnings - Manufacturing | 5 |
| 105 | CES275 | Average Hourly Earnings Goods-Producing | 5 |
| 106 | | Real Foreign Exchange Rate: Swiss | 4 |
| 107 | | Real Foreign Exchange Rate: Japan | 4 |
| 108 | | Real Foreign Exchange Rate: Uk | 4 |
| 109 | | Real Foreign Exchange Rate: Canada | 4 |
| 110 | | Us - Canada Interest Rates Spread | 1 |
| 111 | | Us - Japan Interest Rates Spread | 1 |
| 112 | | Us - Uk Interest Rates Spread | 1 |