ECARES

Dynamic Factors in the Presence of Block Structure

Marc Hallin
ECARES, Institut de Recherche en Statistiques, Département de Mathématique,
Académie Royale de belgique

Roman Liska
IPSC, European Commission, Ispra, Italy

ECARES working paper 2008_012

ECARES
ULB - CP 114
50, F.D. Roosevelt Ave., B-1050 Brussels BELGIUM
www.ecares.org

Dynamic Factors in the Presence of Block Structure

Marc Hallin[*†‡] and Roman Liška[*§]
Université Libre de Bruxelles
Brussels, Belgium

May 25, 2008

Abstract

Macroeconometric data often come under the form of large panels of time series, themselves decomposing into smaller but still quite large subpanels or blocks. We show how the dynamic factor analysis method proposed in Forni *et al* (2000), combined with the identification method of Hallin and Liška (2007), allows for identifying and estimating joint and block-specific common factors. This leads to a more sophisticated analysis of the structures of dynamic interrelations within and between the blocks in such datasets, along with an informative decomposition of explained variances. The method is illustrated with an analysis of the Industrial Production Index data for France, Germany, and Italy.

Key Words: Panel data; Time series; High dimensional data; Dynamic factor model; Business cycle; Block specific factors; Dynamic principal components; Information criterion.

1 Introduction

1.1 Panel data and dynamic factor models

In many fields—macroeconometrics, finance, environmental sciences, chemometrics, ...—information comes under the form of a large number of observed time series or *panel data*. Panel data consist of series of observations (length T) made on n individuals or "cross-sectional items" that have been put together on purpose, because, mainly, they carry some information about some

[*]E.C.A.R.E.S., ECORE, Institut de Recherche en Statistique, and Département de Mathématique.

[†]Académie Royale de Belgique.

[‡]Part of this work was completed while Marc Hallin was visiting the Economics Department at the European University Institute in Florence under a Fernand Braudel Fellowship; the hospitality and financial support of EUI is gratefully acknowledged.

[§]Current affiliation: Institute for the Protection and Security of the Citizen (IPSC), European Commission Joint Research Centre, Ispra, Italy.

common feature or unobservable process of interest, or are expected to do so. This "commonness" is a distinctive feature of panel data : mutually independent cross-sectional items, in that respect, do not constitute a panel (or then, a degenerate one).

On the other hand, the cross-sectional items of a panel, although carrying some common information, also are distinct from each other. Cross-sectional heterogeneity is another distinctive feature of panel data: n (possibly non independent) replications of the same time series would be another form of degenarecy of a panel. Moreover, the impact of item-specific or *idiosyncratic* effects, which have the role of a nuisance, very often dominate, quantitatively, that of the common features one is interested in.

Finally, all individuals in a panel are exposed to the influence of unobservable or unrecorded covariates, which create complex interdependencies, both in the cross-sectional as in the time dimension, which cannot be modelled , as this would require criticable modelling assumptions and a prohibitive number of nuisance parameters. These interdependencies may affect all (or almost all) items in the panel, in which case they are "common"; they also may be specific to a small number of items, hence "idiosyncratic".

The idea of separating "common" and "idiosyncratic" effects is thus at the core of panel data analysis. The same idea is the cornerstone of another statistical domain : factor analysis. There is little surprise, thus, to see a time series version of factor analysis emerging as a powerful tool in the analysis of panel data. This time series version of factor models, however, requires an adequate definition of "commonness" and "idiosyncrasy". This definition should not simply allow for identifying the decomposition of the observation into a "common" component and an "idiosyncratic" one, but also should provide an adequate translation of the inuitive meanings of "common" and "idiosyncratic".

Denote by X_{it} the observation of item i ($i = 1, \ldots, n$) at time t ($t = 1, \ldots, T$); this observation is usually decomposed into a sum

$$X_{it} = \chi_{it} + \xi_{it}, \quad i = 1, \ldots, n, \quad t = 1, \ldots, T$$

of two mutually orthogonal (at all leads and lags) unobservable components : a *common component* χ_{it} and an idiosyncratic one ξ_{it}. Some authors identify this decomposition by requiring the idiosyncratic components to be "small" or "negligible", as in dimension reduction techniques. Some others require that the n idiosyncratic processes be mutually orthogonal white noises. Such characterizations are not reflecting the fundamental nature of factor models: idiosyncratic components indeed can be "large" and strongly autocorrelated, while white noise can be common. For instance, in a model of the form $X_{it} = \chi_t + \xi_{it}$, where χ_t is white noise and orthogonal to $\xi_{it} = \varepsilon_{it} + a_i\varepsilon_{i,t-1}$, with i.i.d. ε_{it}'s, the white noise component χ_t, which is present in all cross-sectional items, very much qualifies as being "common", while the cross-sectionally independent autocorrelated ξ_{it}'s, being item-specific, exhibit all the attributes one would like to see in an "idiosyncratic" component.

A possible characterization of commonness/idiosyncrasy is obtained by requiring the common component to account for all cross-sectional correlations, leading to possibly autocorrelated but cross-sectionally orthogonal idiosyncratic components. This yields the so-called "exact factor models" considered, for instance, by Sargent and Sims (1997) and Geweke (1997). These exact models, however, are too restrictive in most real life applications, where it often happens that two (or a small number of) cross-sectional items, being neighbours in some broad sense, exhibit cross-sectional correlation also in variables that are orthogonal, at all leads and lags, to all other observations throughout the panel. A "weak" or "approximate factor model", allowing for mildly cross-sectionally correlated idiosyncratic components, therefore also has been proposed (Chamberlain 1983; Chamberlain and Rothschild 1983), in which, however, the common and idiosyncratic components are only asymptotically (as $n \to \infty$) identified. Under its most general form, the characterization of idiosyncrasy, in this weak factor model, can be based on the behavior, as $n \to \infty$, of the eigenvalues of the spectral density matrices of the unobservable idiosyncratic components, but also (Forni and Lippi 2001) on the asympotic behavior of the eigenvalues of the spectral density matrices of the observations themselves : see Section 2 for details. This general characterization is the one we are adopting here.

Finally, once the common and idiosyncratic components are identified, two types of factor models can be found in the literature, depending on the way *factors* are driving the common components. In *static* factor models, it is assumed that common components are of the form

$$\chi_{it} = \sum_{l=1}^{q} b_{il} f_{lt}, \quad i = 1, \ldots, n, \quad t = 1, \ldots, T, \tag{1.1}$$

that is, the χ_{it}'s are driven by q factors $f_{1t}, \ldots, f_{qt}$ which are loaded instantaneously. This static approach is the one adopted by Chamberlain (1983), Chamberlain and Rothschild (1983), Stock and Watson (1989, 2002a and 2002b), Bai and Ng (2002 and 2007), and a large number of applied studies. The so-called *general dynamic model* decomposes common components into

$$\chi_{it} = \sum_{l=1}^{q} b_{il}(L) u_{lt}, \quad i = 1, \ldots, n, \quad t = 1, \ldots, T, \tag{1.2}$$

where $u_{1t}, \ldots, u_{qt}$, the *common shocks,* are loaded via one-sided linear filters $b_{il}(L)$. That "truly dynamic" approach (the terminology is not unified and the adjective "dynamic" is often used in an ambiguous way) goes back, under exact factor form, to Chamberlain (1983) and Chamberlain and Rothschild (1983), but was developed, mainly, by Forni *et al* (2000, 2003, 2004, 2005), Forni and Lippi (2001), Hallin and Liška (2007).

The static model (1.1) clearly is a particular case of the general dynamic one (1.2). Its main advantage is simplicity. On the other hand, both models share the same assumption on the asympotic behavior of spectral eigenvalues—a behavior which is confirmed by empirical evidence. But the static model (1.1) places an additional and rather severe restriction on the data generating process, while the dynamic one (1.2), as shown by Lippi and Forni (2001), does not—we refer to Section 2 for details. Moreover, the synchronization of clocks and calendars across the panel is often quite approximative, so that the concept of "instantaneous loading" itself may be questionable.

Both the static and the general dynamic models are receiving increasing attention in finance and macroeconometric applications where information usually is scattered through a (very) large number n of interrelated time series (n values of the order of several hundreds, or even one thousand, are not uncommon). Classical multivariate time series techniques are totally helpless in the presence of such values of n, and factor model methods, to the best of our knowledge, are the only ones that can handle such datasets. In macroeconomics, factor models are used in business cycle analysis (Forni and Reichlin 1998; Giannone, Reichlin, and Sala 2006), in the identification of economy-wide and global shocks (Forni, Giannone *et al* 2005), in the construction of indexes and forecasts exploiting the information scattered in a huge number of interrelated series (Altissimo *et al* 2001), in the monitoring of economic policy (Giannone, Reichlin, and Sala 2004), and in monetary policy applications (Bernanke and Boivin 2003; Favero *et al* 2005). In finance, factor models are at the heart of the extensions proposed by Chamberlain and Rothschild (1983) and Ingersol (1984) of the classical arbitrage pricing theory; they also have been considered in performance evaluation and risk measurement (Chapters 5 and 6 of Campbell *et al* 1997), and in the statistic analysis of the structure of stock returns (Yao 2008).

Factor models in the recent years also generated a huge amount of applied work: see Artis *et al* (2002), Bruneau *et al* (2003), den Reijer (2005), Dreger and Schumacher (2004), Nieuwenhuyzen (2004), Schneider and Spitzer (2004), Giannone and Matheson (2007), and Stock and Watson (2002b) for applications to data from UK, France, the Netherlands, Germany, Belgium, Austria, New Zealand, and the US, respectively; Altissimo *et al* (2001), Angelini *et al* (2001), Forni *et al* (2003), and Marcellino *et al* (2003) for the Euro area and Aiolfi *et al* (2006) for South American data—to quote only a few. Dynamic factor models also have entered the practice of a number of economic and financial institutions, including several central banks and national statistical offices, who are using them in their current analysis and prediction of economic activity. A real time coincident indicator of the EURO area business cycle (EuroCOIN), based on Forni *et al* (2000), is published monthly by the London-based Center for Economic Policy Research and the Banca dItalia: see [http://www.cepr.org/data/EuroCOIN/]. A similar index, based on the methods, is established for the US economy by the Federal Reserve Bank of Chicago.

1.2 Dynamic factor models in the presence of blocks: outline of the paper

Although heterogeneous, panel data very often are obtained by pooling together several "blocks" which themselves can be considered as "large" subpanels. In macroeconometrics, for instance, data typically are organized either by country or sectoral origin: the database which is used in the construction of EuroCOIN, the monthly indicator of the euro area business cycle published by CEPR, includes almost 1000 time series that cover six European countries and are organized into eleven blocks including industrial production, producer prices, monetary aggregates, etc. When these blocks are large enough, several dynamic factor models can be considered and analyzed, allowing for a refined analysis of interblock relations. In the simple two-block case, "marginal common factors" can be defined for each block, and need not coincide with the "joint common factors" resulting from pooling the two blocks.

The objective of this paper is to provide a theoretical basis for that type of analysis. For simplicity, we start with the simple case of two blocks. We show (Section 2) how the Hilbert space spanned by the n observed series decomposes into four mutually orthogonal subspaces: the space of *strongly common variables*, which are common to both subpanesls, the space of *strongly idiosyncratic variables*, which are idiosyncratic to both subpanels, and two spaces of *weakly common/weakly idiosyncratic variables*, which are common to one subpanel but idiosyncratic to the other. In Sections 3 and 4, we show how the projection of each observation onto those various subspaces is asymptotically identified and how it can be consistently reconstructed from the observations. Section 5 is devoted to the general case of $K \geq 2$ blocks, allowing for a decomposition of each observation into 2^K mutually orthogonal components. The tools we are using throughout are Brillinger's theory of dynamic principal components and the identification method developed by Hallin and Liška (2007). Proofs are concentrated in an appendix (Section 7).

The potential of the method is briefly illustrated, in Section 6, with a panel of Industrial Production Index data for France and Germany ($K = 2$, four distinct components), then France, Germany, and Italy ($K = 3$, hence eight distinct components). Simple as it is, the analysis of that dataset reveals some striking facts. For instance, both Germany and Italy exhibit a "national common factor" which is idiosyncratic to the other two countries, while France's common factors are included in the space spanned by Germany's. The (estimated) percentages of explained variation associated with the various cases also are quite illuminating : Germany, with 25.9% of common variation, is the "most common" out of the three countries. But it also is, with only 4.9% of its total variation, the "least strongly common" one. France has the highest proportion (79.6%) of marginal idiosyncratic variation but also the highest proportions of strongly and weakly idiosyncratic variations (72.7% and 6.9%, respectively).

We do not attempt here to provide an economic interpretation for such facts. Nor do we apply the method to a more sophisticated dataset. But we feel that the simple application we are proposing provides sufficient evidence of the potential power of the method, both from a structural as from a quantitative point of view.

2 The dynamic factor model in the presence of blocks

We throughout assume that all stochastic variables considered in this paper belong to the Hilbert space $L_2(\Omega, \mathcal{F}, \mathrm{P})$, where $(\Omega, \mathcal{F}, \mathrm{P})$ is some given probability space. We will study two double-indexed sequences of observed random variables

$$\mathbf{Y} := \{Y_{it}, i \in \mathbb{N}, t \in \mathbb{Z}\} \quad \text{and} \quad \mathbf{Z} := \{Z_{jt}, j \in \mathbb{N}, t \in \mathbb{Z}\},$$

where t stands for time and i, j are cross-sectional indices. Let $\mathbf{Y}_{n_y} := \{\mathbf{Y}_{n_y,t}, t \in \mathbb{Z}\}$ and $\mathbf{Z}_{n_z} := \{\mathbf{Z}_{n_z,t}, t \in \mathbb{Z}\}$ be the n_y- and n_z-dimensional subprocesses of $\mathbf{Y}$ and $\mathbf{Z}$, respectively, where

$$\mathbf{Y}_{n_y,t} := (Y_{1t} \ldots, Y_{n_y t})' \quad \text{and} \quad \mathbf{Z}_{n_z,t} := (Z_{1t} \ldots, Z_{n_z t})',$$

and write $\mathbf{X}_{\mathbf{n},t} := (Y_{1t} \ldots, Y_{n_y t}, Z_{1t} \ldots, Z_{n_z t})' := (\mathbf{Y}'_{n_y,t} \mathbf{Z}'_{n_z,t})'$ with $\mathbf{n} := (n_y, n_z)$ and $n := n_y + n_z$. The Hilbert subspaces spanned by the processes $\mathbf{Y}$, $\mathbf{Z}$ and $\mathbf{X}$ are denoted by $\mathcal{H}_y$, $\mathcal{H}_z$ and $\mathcal{H}$, respectively.

The following assumption is made throughout the paper.

Assumption A1. *For all* $\mathbf{n}$*, the vector process* $\{\mathbf{X}_{\mathbf{n},t}; t \in \mathbb{Z}\}$ *is a zero mean second order stationary process.*

Denoting by $\mathbf{\Sigma}_{y;n_y}(\theta)$ and $\mathbf{\Sigma}_{z;n_z}(\theta)$ the $(n_y \times n_y)$ and $(n_z \times n_z)$ spectral density matrices of $\mathbf{Y}_{n_y,t}$, $\mathbf{Z}_{n_z,t}$, respectively, and by $\mathbf{\Sigma}_{yz;\mathbf{n}}(\theta) = \mathbf{\Sigma}'_{zy;\mathbf{n}}(\theta)$ their $(n_y \times n_z)$ cross-spectrum matrix, write

$$\mathbf{\Sigma}_{\mathbf{n}}(\theta) =: \begin{pmatrix} \mathbf{\Sigma}_{y;n_y}(\theta) & \mathbf{\Sigma}_{yz;\mathbf{n}}(\theta) \\ \mathbf{\Sigma}_{zy;\mathbf{n}}(\theta) & \mathbf{\Sigma}_{z;n_z}(\theta) \end{pmatrix}$$

for the $(n \times n)$ spectral density matrix of $\mathbf{X}_{\mathbf{n},t}$, with elements $\sigma_{i_1 i_2}(\theta)$, $\sigma_{j_1 j_2}(\theta)$ or $\sigma_{ii}(\theta)$, $i, i_1, i_2 = 1, \ldots, n_y$, $j, j_1, j_2 = 1, \ldots, n_z$. On these matrices, we make the following assumption.

Assumption A2. *For any* $k \in \mathbb{N}$*, there exists a real* $c_k > 0$ *such that* $\sigma_{kk}(\theta) \leq c_k$ *for any* $\theta \in [-\pi, \pi]$*.*

For any $\theta \in [-\pi, \pi]$, let $\lambda_{y;n_y,i}(\theta)$ be $\mathbf{\Sigma}_{y;n_y}(\theta)$'s i-th eigenvalue (in decreasing order of magnitude). The function $\theta \mapsto \lambda_{y;n_y,i}(\theta)$ is called $\mathbf{\Sigma}_{y;n_y}(\theta)$' *i-th dynamic eigenvalue*. The notation $\theta \mapsto \lambda_{z;n_z,j}(\theta)$ and $\theta \mapsto \lambda_{\mathbf{n},k}(\theta)$ is used in an obvious way for the dynamic eigenvalues of $\mathbf{\Sigma}_{z;n_z}(\theta)$ and $\mathbf{\Sigma}_{\mathbf{n}}(\theta)$, respectively.

The corresponding *dynamic* eigenvectors, of dimensions $(n_y \times 1)$, $(n_z \times 1)$, and $(n \times 1)$, are denoted by $\mathbf{p}_{y;n_y,i}(\theta)$, $\mathbf{p}_{z;n_z,j}(\theta)$, and $\mathbf{p}_{\mathbf{n},k}(\theta)$, respectively. These dynamic eigenvectors can be expanded in Fourier series, e.g.

$$\mathbf{p}_{\mathbf{n},k}(\theta) = \frac{1}{2\pi} \sum_{s=-\infty}^{\infty} \left[\int_{-\pi}^{\pi} \mathbf{p}_{\mathbf{n},k}(\theta) e^{\mathrm{i}s\theta} d\theta \right] e^{-\mathrm{i}s\theta}$$

where the series on the right hand side converge in quadratic mean, which in turn defines square summable filters of the form

$$\underline{\mathbf{p}}_{\mathbf{n},k}(L) = \frac{1}{2\pi} \sum_{s=-\infty}^{\infty} \left[\int_{-\pi}^{\pi} \mathbf{p}_{\mathbf{n},k}(\theta) e^{\mathrm{i}s\theta} d\theta \right] L^s.$$

Similarly define $\underline{\mathbf{p}}_{y;n_y,i}(L)$ and $\underline{\mathbf{p}}_{z;n_z,j}(L)$ from $\mathbf{p}_{y;n_y,i}(\theta)$ and $\mathbf{p}_{z;n_z,j}(\theta)$, respectively.

On those dynamic eigenvalues, we make the following assumptions.

Assumption A3. *For some* $q_y, q_z \in \mathbb{N}$,

(i) the q_y*-th dynamic eigenvalue of* $\mathbf{\Sigma}_{y;n_y}(\theta)$, $\lambda_{y;n_y,q_y}(\theta)$*, diverges as* $n_y \to \infty$*, a.e. in* $[-\pi, \pi]$*, while the* $(q_y + 1)$*-th one ,* $\lambda_{y;n_y,q_y+1}(\theta)$*, is* θ*-a.e. bounded;*

(ii) the q_z-th dynamic eigenvalue of $\mathbf{\Sigma}_{z;n_z}(\theta)$, $\lambda_{z;n_z,q_z}(\theta)$, diverges as $n_z \to \infty$, a.e. in $[-\pi, \pi]$, while the $(q_z + 1)$-th one, $\lambda_{z;n_z,q_z+1}(\theta)$, is θ-a.e. bounded.

The following lemma shows that this behavior of the dynamic eigenvalues of the subpanel spectral matrices $\mathbf{\Sigma}_{y;n_y}(\theta)$ and $\mathbf{\Sigma}_{z;n_z}(\theta)$ entails a similar behavior for the dynamic eigenvalues $\lambda_{\mathbf{n},k}(\theta)$ of $\mathbf{\Sigma}_{\mathbf{n}}(\theta)$.

Lemma 1. *Let Assumptions A1-A3 hold. Then, there exists $q \in \mathbb{N}$, with $\max(q_y, q_z) \leq q \leq q_y + q_z$, such that $\mathbf{\Sigma}_{\mathbf{n}}(\theta)$'s q-th dynamic eigenvalue $\lambda_{\mathbf{n},q}(\theta)$ diverges as $\min(n_y, n_z) \to \infty$, a.e. in $[-\pi, \pi]$, while the $(q + 1)$-th one, $\lambda_{\mathbf{n},q+1}(\theta)$, is θ-a.e. bounded.*

Proof. See the appendix (Section 8.1). □

Theorem 2 in Forni and Lippi (2001) establishes that the behavior of dynamic eigenvalues described in Assumption A3 and Lemma 1 characterizes the existence of a *dynamic factor representation*. We say that a process $\mathbf{X} := \{X_{kt}, k \in \mathbb{N},\ t \in \mathbb{Z}\}$ admits a *dynamic factor representation* with q factors if X_{kt} decomposes into a sum

$$X_{kt} = \chi_{kt} + \xi_{kt}, \quad \text{with} \quad \chi_{kt} := \sum_{l=1}^{q} b_{kl}(L)u_{lt} \quad \text{and} \quad b_{kl}(L) := \sum_{m=1}^{\infty} b_{klm}L^m,\ k \in \mathbb{N},\ t \in \mathbb{Z}\ ,$$

such that

(i) the q-dimensional vector process $\{\mathbf{u}_t := (u_{1t}\ u_{2t}\ ...\ u_{qt})'; t \in \mathbb{Z}\}$ is orthonormal white noise;

(ii) the (unobservable) n-dimensional processes $\{\boldsymbol{\xi}_n := (\xi_{1t}\ \xi_{2t}\ \cdots\ \xi_{nt})'; t \in \mathbb{Z}\}$ are zero-mean stationary for any n, with *(idiosyncrasy)* θ-a.e. bounded (as $n \to \infty$) dynamic eigenvalues;

(iii) ξ_{k,t_1} and u_{l,t_2} are mutually orthogonal for any k, l, t_1 and t_2;

(iv) the filters $b_{kl}(L)$ are square summable: $\sum_{m=1}^{\infty} b_{klm}^2 < \infty$ for all $k \in \mathbb{N}$ and $l = 1, \ldots, q$, and

(v) q is minimal with respect to (i)-(iv).

The processes $\{u_{lt}, t \in \mathbb{Z}\}$, $l = 1, \ldots, q$, are called the *common shocks* or *factors*, the random variables ξ_{kt} and χ_{kt} the *idiosyncratic* and *common components* of X_{kt}, respectively. Actually, Forni and Lippi define *idiosyncrasy* via the behavior of dynamic aggregates, then show (their Theorem 1) that this definition is equivalent to the condition on dynamic eigenvalues we are giving here.

This result of Forni and Lippi (2001), along with Lemma 1, leads to the following proposition.

Proposition 1. *Let Assumption A1 and A2 hold. Then,*

(a) Assumption A3(i) is satisfied iff the process $\mathbf{Y}$ has a dynamic factor representation (q_y factors; call them the (common) $y-$factors)

$$Y_{it} = \chi_{y;it} + \xi_{y;it} = \sum_{l=1}^{q_y} b_{y;il}(L)u_{y;lt} + \xi_{y;it}\ , \quad i \in \mathbb{N},\ t \in \mathbb{Z}; \tag{2.3}$$

(b) *Assumption A3(ii) is satisfied iff the process* $\mathbf{Z}$ *has a dynamic factor representation* (q_z *factors; call them the (common)* $z-$*factors)*

$$Z_{jt} = \chi_{z;jt} + \xi_{z;jt} = \sum_{l=1}^{q_z} b_{z;jl}(L) u_{z;lt} + \xi_{z;jt} \ , \quad j \in \mathbb{N}, \ t \in \mathbb{Z}; \tag{2.4}$$

(c) *Assumption A3 is satisfied iff the process* $\mathbf{X}$ *has a dynamic factor representation (*q *factors, with* q *characterized in Lemma 1; call them the* joint *common factors)*

$$X_{kt} = Y_{it} = \chi_{xy;it} + \xi_{xy;it} = \sum_{l=1}^{q} b_{xy;il}(L) u_{lt} + \xi_{xy;it} \ , \quad k \in \mathbb{N}, \ t \in \mathbb{Z} \tag{2.5}$$

in case $X_{kt} = Y_{it}$ *and*

$$X_{kt} = Z_{jt} = \chi_{xz;jt} + \xi_{xz;jt} = \sum_{l=1}^{q} b_{xz;jl}(L) u_{lt} + \xi_{xz;jt} \ , \quad k \in \mathbb{N}, \ t \in \mathbb{Z} \tag{2.6}$$

in case $X_{kt} = Z_{jt}$.

All filters involved have square-summable coefficients.

Proof. The proof follows directly from the characterization theorem of Forni and Lippi(2001), along, for part (c), with Lemma 1. □

It follows that, under Assumption A3, the processes $\mathbf{Y}$ and $\mathbf{Z}$ admit two distinct decompositions each: the *marginal factor models* (a) and (b), with *marginal common shocks* $u_{y;lt}$ ($l = 1, \dots q_y$) and $u_{z;lt}$ ($l = 1, \dots q_z$), respectively, and the *joint factor model* (c), with *joint common shocks* u_{lt} ($l = 1, \dots q$). This double representation allows for refining the factor decomposition. Call $x-$, $y-$, or $z-$*idiosyncratic* a process which is orthogonal (at all leads and lags) to the $x-$, $y-$, or $z-$factors, respectively. Similarly, call $x-$, $y-$, or $z-$*common* any process belonging to the Hilbert space generated by the $x-$, $y-$, or $z-$factors. The joint common components $\chi_{xy;it}$ and $\chi_{xz;jt}$ then further decompose into

$$\chi_{xy;it} = \phi_{y;it} + \psi_{y;it} + \nu_{y;it} \quad \text{and} \quad \chi_{xz;jt} = \phi_{z;ij} + \psi_{z;jt} + \nu_{z;jt},$$

where $\phi_{y;it}$ and $\phi_{z;jt}$ are $y-$ and $z-$common, $\psi_{y;it}$ and $\nu_{z;jt}$ are $y-$common but $z-$idiosyncratic, and $\nu_{y;it}$ and $\psi_{z;jt}$ are $z-$common but $y-$idiosyncratic. We thus have

$$Y_{it} = \overbrace{\underbrace{\phi_{y;it} + \psi_{y;it}}_{\chi_{y;it}} + \underbrace{\nu_{y;it}}_{} }^{\chi_{xy;it}} \underbrace{+ \xi_{xy;it}}_{\xi_{y;it}} \quad \text{and} \quad Z_{jt} = \overbrace{\underbrace{\phi_{z;jt} + \psi_{z;jt}}_{\chi_{z;jt}} + \underbrace{\nu_{z;jt}}_{}}^{\chi_{xz;jt}} \underbrace{+ \xi_{xz;jt}}_{\xi_{z;jt}} \ , \quad i, j \in \mathbb{N}, \ t \in \mathbb{Z}. \tag{2.7}$$

More precisely, consider the Hilbert subspaces $\mathcal{H}_y^\chi$, $\mathcal{H}_z^\chi$, and $\mathcal{H}^\chi$ of $\mathcal{H}$ spanned by the common components $\{\chi_{y;it}, \ i \in \mathbb{N}, t \in \mathbb{Z}\}$, $\{\chi_{z;jt}, \ j \in \mathbb{N}, t \in \mathbb{Z}\}$, and $\{\chi_{y;it}, \chi_{z;jt}, \ i, j \in \mathbb{N}, t \in \mathbb{Z}\}$, respectively. Similarly define, for idiosyncratic components, $\mathcal{H}_y^\xi$, $\mathcal{H}_z^\xi$, and $\mathcal{H}^\xi$. These subspaces induce a partition of $\mathcal{H}$ into four mutually orthogonal subspaces: $\mathcal{H}^\phi := \mathcal{H}_y^\chi \bigcap \mathcal{H}_z^\chi$ (containing

$\phi_{y;it}$ and $\phi_{z;jt}$), $\mathcal{H}^{\psi}_y := \mathcal{H}^{\chi}_y \bigcap \mathcal{H}^{\xi}_z$ (containing $\psi_{y;it}$ and $\nu_{z;jt}$), $\mathcal{H}^{\psi}_z := \mathcal{H}^{\xi}_y \bigcap \mathcal{H}^{\chi}_z$ (containing $\nu_{y;it}$ and $\psi_{z;jt}$), and $\mathcal{H}^{\xi} := \mathcal{H}^{\xi}_y \bigcap \mathcal{H}^{\xi}_z$ (containing $\xi_{xy;it}$ and $\xi_{xz;jt}$).

Clearly, $\mathcal{H}^{\chi}_y$ and $\mathcal{H}^{\chi}_z$ are subspaces of $\mathcal{H}^{\chi}$. Since $\mathcal{H}^{\chi}_y$ is spanned by the q_y-tuple of mutually orthogonal white noises $\{u_{y;lt},\ 1 \leq l \leq q_y, t \in \mathbb{Z}\}$, it has *dynamic dimension* q_y. Similarly, $\mathcal{H}^{\chi}_z$ has dynamic dimension q_z, and $\mathcal{H}^{\chi}$ dynamic dimension q. Denote by q_{yz} the dynamic dimension of the intersection $\mathcal{H}^{\phi}$ of $\mathcal{H}^{\chi}_y$ and $\mathcal{H}^{\chi}_z$. This intersection may reduce to the origin in $\mathcal{H}$, in which case $q_{yz} = 0$; it may coincide with $\mathcal{H}^{\chi}_y$ (resp., with $\mathcal{H}^{\chi}_z$) when $\mathcal{H}^{\chi}_y \subseteq \mathcal{H}^{\chi}_z$ (resp., $\mathcal{H}^{\chi}_z \subseteq \mathcal{H}^{\chi}_y$), with $q_{yz} = \min(q_y, q_z)$. Whenever $q_{yz} \geq 1$, let $\{v_{lt},\ 1 \leq l \leq q_{yz},\ t \in \mathbb{Z}\}$ denote a q_{yz}-tuple of of mutually orthogonal white noises spanning this intersection. This q_{yz}-tuple can be extended into a q_y-tuple $\{v_{lt},\ v_{y,mt},\ 1 \leq l \leq q_{yz},\ 1 \leq m \leq q_y - q_{yz},\ t \in \mathbb{Z}\}$ spanning $\mathcal{H}^{\chi}_y$, or into a q_z-tuple $\{v_{lt},\ v_{z,mt},\ 1 \leq l \leq q_{yz},\ 1 \leq m \leq q_z - q_{yz},\ t \in \mathbb{Z}\}$ spanning $\mathcal{H}^{\chi}_z$. We then have

$$\phi_{y;it} = \sum_{l=1}^{q_{yz}} d_{il}(L) v_{lt}\ , \quad \psi_{y;it} = \sum_{l=1}^{q_y - q_{yz}} d_{y;il}(L) v_{y;lt}\ , \quad \nu_{y;it} = \sum_{l=1}^{q_z - q_{yz}} d_{yz;il}(L) v_{z;lt}\ ,$$

and

$$\phi_{z,jt} = \sum_{l=1}^{q_{yz}} d_{jl}(L) v_{lt}\ , \quad \psi_{z,jt} = \sum_{l=1}^{q_z - q_{yz}} d_{z,jl}(L) v_{z;lt}\ , \quad \nu_{z,jt} = \sum_{l=1}^{q_y - q_{yz}} d_{zy,jl}(L) v_{y;lt}.$$

Note that $\psi_{z,jt}$ and $\psi_{y,it}$ are common in the joint model (2.5)-(2.6), but that $\psi_{z,jt}$ is idiosyncratic in the marginal models (2.3), $\psi_{y,it}$ in the marginal model (2.4)—therefore call them *weakly common*. Similarly, $\nu_{y;it}$ and $\nu_{z;jt}$ are are common in the joint model (2.5)-(2.6), but idiosyncratic in in the marginal models (2.3) and (2.4), respectively—call them *weakly idiosyncratic*. We say that $\phi_{y;it}$ and $\phi_{z,jt}$, which are both y- and z-common, are *strongly common*; similarly, $\xi_{xy;it}$ and $\xi_{xz;jt}$, being y- and z-idiosyncratic, are called *strongly idiosyncratic*.

In the following sections, we propose a procedure that provides consistent estimates of $\phi_{y;it}$, $\psi_{y;it}$, $\nu_{y;it}$, $\xi_{xy;it}$ and $\phi_{z,jt}$, $\psi_{z,jt}$, $\nu_{z,jt}$, $\xi_{xz;jt}$, hence $\xi_{y;it}$, $\xi_{xy;it}$, $\xi_{z;jt}$, and $\xi_{xz;jt}$.

3 Identifying the factor structure; population results

Based on the n-dimensional vector process $\mathbf{X}_{\mathbf{n},t} = \left(\mathbf{Y}'_{n_y,t}, \mathbf{Z}'_{n_z,t}\right)'$, we first asymptotically identify $\phi_{y;it}$, $\psi_{y;it}$, $\nu_{y;it}$, $\phi_{z;jt}$, $\psi_{z;jt}$ and $\nu_{z;jt}$ as $\min(n_y, n_z) \to \infty$. More precisely, we show that, under specified spectral structure, all those quantities can be consistently recovered from the observations $\mathbf{X}_{\mathbf{n},t}$.

3.1 Recovering the joint common and strongly idiosyncratic components

Under the joint factor model, Proposition 2 in Forni *et al* (2000) provides $\mathbf{X}_{\mathbf{n},t}$-measurable reconstructions—denoted by $\chi^{\mathbf{n}}_{xy;it}$ and $\chi^{\mathbf{n}}_{xz;jt}$, respectively—of the joint common components $\chi_{xy;it}$ and $\chi_{xz;jt}$, which converge in quadratic mean for any i, j and t, as $\min(n_y, n_z) \to \infty$; we are using the terminology "reconstruction" rather than "estimation" to emphasize that spectral densities here, unlike in Section 4, are assumed to be known.

Write $\mathbf{M}^*$ for the adjoint (transposed, complex conjugate) of a matrix $\mathbf{M}$. The scalar process $\{V_{\mathbf{n},kt} := \underline{\mathbf{p}}^*_{\mathbf{n},k}(L)\mathbf{X}_{\mathbf{n},t}, t \in \mathbb{Z}\}$, $k = 1, \ldots, n$, the spectral density of which is $\lambda_{\mathbf{n},j}(\theta)$, will be called $\mathbf{X}_{\mathbf{n},t}$'s *k-th dynamic principal component.* The basic properties of dynamic principal components imply that $\{V_{\mathbf{n},k_1t}\}$ and $\{V_{\mathbf{n},k_2t}\}$, for $k_1 \neq k_2$, are mutually orthogonal at all leads and lags. Forni *et al* (2000) show that the projections of Y_{it} and Z_{jt} onto the closed space spanned by the present, past and future values of $V_{\mathbf{n},kt}$, $k = 1, \ldots, q$ yield the desired reconstructions of of $\chi_{xy;it}$ and $\chi_{xz;jt}$. They also provide (up to a minor change due to the fact that they are considering row rather than column-eigenvectors, as we do here) the explicit forms

$$\chi^{\mathbf{n}}_{xy;it} = \underline{\mathbf{K}}^*_{y;\mathbf{n},i}(L)\mathbf{X}_{\mathbf{n},t} \quad \text{and} \quad \chi^{\mathbf{n}}_{xz;jt} = \underline{\mathbf{K}}^*_{z;\mathbf{n},j}(L)\mathbf{X}_{\mathbf{n},t} \quad i = 1, \ldots, n_y, \quad j = 1, \ldots, n_z, \tag{3.8}$$

with

$$\underline{\mathbf{K}}_{y;\mathbf{n},i}(L) := \sum_{k=1}^{q} \underline{p}^*_{\mathbf{n},k,i}(L)\underline{\mathbf{p}}_{\mathbf{n},k}(L) \quad \text{and} \quad \underline{\mathbf{K}}_{z,\mathbf{n},j}(L) = \sum_{k=1}^{q} \underline{p}^*_{\mathbf{n},k,j}(L)\underline{\mathbf{p}}_{\mathbf{n},k}(L)$$

where $\underline{p}_{\mathbf{n},k,i}(L)$ denotes the i-th component of $\underline{\mathbf{p}}_{\mathbf{n},k}(L)$ such that $X_{kt} = Y_{it}$ and $\underline{p}_{\mathbf{n},k,j}(L)$ the j-th component of $\underline{\mathbf{p}}_{\mathbf{n},k}(L)$ such that $X_{kt} = Z_{jt}$.

We then can state a first consistency result.

Proposition 2. *Let Assumptions A1-A3 hold. Then,*

$$\lim_{\min(n_y,n_z)\to\infty} \chi^{\mathbf{n}}_{xy;it} = \chi_{xy;it} \quad \textit{and} \quad \lim_{\min(n_y,n_z)\to\infty} \chi^{\mathbf{n}}_{xz;jt} = \chi_{xz;jt}$$

in quadratic mean, for any i, j, and t.

Proof. The proof consists in applying Proposition 2 in Forni *et al* (2000) to the joint panel.□

It follows from (3.8) that $\chi^{\mathbf{n}}_{xy;it}$ has variance

$$\text{Var}(\chi^{\mathbf{n}}_{xy;it}) = \sum_{k=1}^{q} \int_{-\pi}^{\pi} |p_{\mathbf{n},k,i}(\theta)|^2 \lambda_{\mathbf{n},k}(\theta) d\theta.$$

Averaging this variance over the subpanel produces a measure

$$\frac{1}{n_y} \sum_{i=1}^{n_y} \text{Var}(\chi^{\mathbf{n}}_{xy;it}) = \frac{1}{n_y} \sum_{i=1}^{n_y} \sum_{k=1}^{q} \int_{-\pi}^{\pi} |p_{\mathbf{n},k,i}(\theta)|^2 \lambda_{\mathbf{n},k}(\theta) d\theta \tag{3.9}$$

of the contribution of joint common factors in the variability of the y-subpanel. Dividing it by the averaged variance

$$\frac{1}{n_y} \sum_{i=1}^{n_y} \text{Var}(Y_{it}) = \frac{1}{n_y} \sum_{i=1}^{n_y} \int_{-\pi}^{\pi} \lambda_{y;n_y,i}(\theta) d\theta$$

of the y-subpanel yields an evaluation

$$\sum_{i=1}^{n_y} \sum_{k=1}^{q} \int_{-\pi}^{\pi} |p_{\mathbf{n},k,i}(\theta)|^2 \lambda_{\mathbf{n},k}(\theta) d\theta \Big/ \sum_{i=1}^{n_y} \int_{-\pi}^{\pi} \lambda_{y;n_y,i}(\theta) d\theta \tag{3.10}$$

of its "degree of commonness" within the joint panel. For the z-subpanel, this measure takes the form

$$\sum_{j=1}^{n_z}\sum_{k=1}^{q}\int_{-\pi}^{\pi}|p_{\mathbf{n},k,j}(\theta)|^2\lambda_{\mathbf{n},k}(\theta)d\theta/\sum_{j=1}^{n_z}\int_{-\pi}^{\pi}\lambda_{z;n_z,j}(\theta)d\theta. \tag{3.11}$$

As for the strongly idiosyncratic components $\xi_{xy;it}$, and $\xi_{xz;jt}$, they are consistently recovered, as $\min(n_y, n_z) \to \infty$, by

$$\xi^{\mathbf{n}}_{xy;it} := Y_{it} - \chi^{\mathbf{n}}_{xy;it} \quad \text{and} \quad \xi^{\mathbf{n}}_{xz;jt} := Z_{jt} - \chi^{\mathbf{n}}_{xz;jt},$$

respectively. In view of the mutual orthogonality of common and idiosyncatic components, the variance of $\xi^{\mathbf{n}}_{xy;it}$ writes

$$\mathrm{Var}(\xi^{\mathbf{n}}_{xy;it}) = \mathrm{Var}(Y_{it}) - \sum_{k=1}^{q}\int_{-\pi}^{\pi}|p_{\mathbf{n},k,i}(\theta)|^2\lambda_{\mathbf{n},k}(\theta)d\theta;$$

the complement to one of (3.10) therefore constitutes a measure of the "degree of idiosyncrasy" of the y-subpanel within the joint panel. Similar formulas hold for the strongly idiosyncratic component $\xi^{\mathbf{n}}_{xz;jt}$.

3.2 Recovering the marginal common, marginal idiosyncratic, and weakly idiosyncratic components

If $q_y = q$, then the marginal common and idiosyncratic components $\chi_{y;it}$ and $\xi_{y;it}$ coincide with their joint counterparts $\chi_{xy;it}$ and $\xi_{xy;it}$, which were taken care of in the previous section. Assume therefore that $q > q_y$; the marginal and joint y-common spaces then do not coincide anymore.

Applying to the y- and z-subpanels separately the same type of technique as we used in Section 3.1, consider the spectral density matrix $\mathbf{\Sigma}_{y;n_y}(\theta)$, with eigenvectors $\mathbf{p}_{y;n_y,i}(\theta)$ and the corresponding filters $\underline{\mathbf{p}}_{y;n_y,i}(L)$, $i = 1,\dots,q_y$. A consistent reconstruction of $\chi_{y;it}$ is obtained by projecting Y_{it} onto the closed subspace spanned by the first q_y dynamic principal components $V^{n_y}_{y;1t},\dots,V^{n_y}_{y;q_yt}$ of $\mathbf{\Sigma}_{y;n_y}(\theta)$, where $V^{n_y}_{y;kt} := \underline{\mathbf{p}}^*_{y;n_y,k}(L)\mathbf{Y}_{n_y,t}$. This projection takes the form

$$\chi^{n_y}_{y;it} = \underline{\mathbf{G}}^*_{y;n_y,i}(L)\mathbf{Y}_{n_y,t} \tag{3.12}$$

with

$$\underline{\mathbf{G}}_{y;n_y,i}(L) := \sum_{k=1}^{q_y}\underline{p}^*_{y;n_y,k,i}(L)\underline{\mathbf{p}}_{y;n_y,k}(L). \tag{3.13}$$

Similarly, the reconstruction $\chi^{n_z}_{z;jt}$ of $\chi_{z;jt}$ is

$$\chi^{n_z}_{z;jt} = \underline{\mathbf{G}}^*_{z;n_z,j}(L)\mathbf{Z}_{n_z,t} = \sum_{k=1}^{q_z}\underline{p}_{z;n_z,k,j}(L)\underline{\mathbf{p}}^*_{z;n_z,k}(L)\mathbf{Z}_{n_z,t}. \tag{3.14}$$

We then have a second consistency result.

Proposition 3. *Let Assumptions A1-A3 hold. Then*

$$\lim_{n_y \to \infty} \chi^{n_y}_{y,it} = \chi_{y;it} \quad \textit{and} \quad \lim_{n_z \to \infty} \chi^{n_z}_{z;jt} = \chi_{z;jt}$$

in quadratic mean for any i, j, and t.

Proof. The proof again is a direct application of Proposition 2 in Forni *et al* (2000) to the y- and z-subpanels, respectively. □

The variance of the reconstructed marginal y-common component $\chi^{n_y}_{y;it}$ writes

$$\mathrm{Var}(\chi^{n_y}_{y;it}) = \sum_{k=1}^{q_y} \int_{-\pi}^{\pi} |p_{y;n_y,k,i}(\theta)|^2 \lambda_{y;n_y,k}(\theta) d\theta$$

The averaged variance explained by the y-common factors in the y-subpanel is thus

$$\frac{1}{n_y} \sum_{i=1}^{n_y} \mathrm{Var}(\chi^{n_y}_{y;it}) = \frac{1}{n_y} \sum_{i=1}^{n_y} \sum_{k=1}^{q_y} \int_{-\pi}^{\pi} |p_{y;n_y,k,i}(\theta)|^2 \lambda_{y;n_y,k}(\theta) d\theta = \frac{1}{n_y} \sum_{k=1}^{q_y} \int_{-\pi}^{\pi} \lambda_{y;n_y,k}(\theta) d\theta. \quad (3.15)$$

Similarly, the averaged variance explained by the z-common factors in the z-subpanel is

$$\frac{1}{n_z} \int_{-\pi}^{\pi} \left[\sum_{k=1}^{q_z} \lambda_{z;n_z,k}(\theta) \right] d\theta. \quad (3.16)$$

Consistent reconstructions of the marginal idiosyncratic components $\xi_{y;it}$ and $\xi_{z;jt}$ are straightforwardly obtained as

$$\xi^{n_y}_{y;it} := Y_{it} - \chi^{n_y}_{y;it} \quad \text{and} \quad \xi^{n_z}_{z;jt} := Z_{jt} - \chi^{n_z}_{z;jt}, \quad (3.17)$$

whereas the weakly idiosyncratic components $\nu_{y;it}$ and $\nu_{z;it}$ can be recovered as

$$\nu^{\mathbf{n}}_{y;it} := \chi^{\mathbf{n}}_{xy;it} - \chi^{n_y}_{y;it} = \xi^{n_y}_{y;it} - \xi^{\mathbf{n}}_{xy;it} \quad \text{and} \quad \nu^{\mathbf{n}}_{z;jt} := \chi^{\mathbf{n}}_{xz;jt} - \chi^{n_z}_{z;jt} = \xi^{n_z}_{z;jt} - \xi^{\mathbf{n}}_{xz;jt}, \quad (3.18)$$

respectively. The averaged variance of weakly idiosyncratic components (or its ratio to $\sum_{i=1}^{n_y} \mathrm{Var}(Y_{it})$), which measures extent to which the z-common factors contribute to y-idiosyncratic variation, is also a quantity of interest. Clearly, since $\xi^{n_y}_{y;it} = \xi^{\mathbf{n}}_{xy;it} + \nu^{\mathbf{n}}_{y;it}$, where $\xi^{\mathbf{n}}_{xy;it}$ (which is joint idiosyncratic) and $\nu^{\mathbf{n}}_{y;it}$ (which is joint comon) are mutually orthogonal,

$$\mathrm{Var}(\nu^{\mathbf{n}}_{y;it}) = \mathrm{Var}(\xi^{n_y}_{y;it}) - \mathrm{Var}(\xi^{\mathbf{n}}_{xy;it}), \quad (3.19)$$

so that

$$\frac{1}{n_y} \sum_{i=1}^{n_y} \mathrm{Var}(\nu^{\mathbf{n}}_{y;it}) = \frac{1}{n_y} \Big[\sum_{i=1}^{n_y} \sum_{k=1}^{q} \int_{-\pi}^{\pi} |p_{\mathbf{n},k,i}(\theta)|^2 \lambda_{\mathbf{n},k}(\theta) d\theta - \sum_{k=1}^{q_y} \int_{-\pi}^{\pi} \lambda_{y;n_y,k}(\theta) d\theta \Big].$$

Similar formulas hold for $\nu^{\mathbf{n}}_{z;jt}$.

3.3 Disentangling the strongly and weakly common components

As explained in Section 2, each element of the Hilbert space spanned by the observed variables decomposes into a sum of four mutually orthogonal components—the strongly common (both y- and z-common), the weakly common/weakly idiosyncratic (either y-common and z-idiosyncratic or y-idiosyncratic and z-common), and the strongly idiosyncratic one (both y- and z-idiosyncratic). So far, we have been able to reconstruct some of these components by implementing the Forni *et al* (2000) filtering, which asymptotically separates common and idiosyncratic components. In order to separate the strongly common component $\phi_{y,it}$ of Y_{it} from the weakly common one $\psi_{y,it}$, however, we need another procedure. Intuitively, three equivalent projections are possible, all on the z-common space or, more precisely, on the approximation ot the z-common space based on the n_z-dimensional z-subpanel :

(a) either Y_{it} is projected, yielding a consistent reconstruction $\chi^{n_z}_{yz,it}$ (see (3.20)) of the z-common component $\phi_{y,it} + \nu_{y,it}$ of Y_{it}, from which $\nu^{\mathbf{n}}_{y,it}$ (obtained in Section 3.2) is easily subtracted, yielding the desired $\phi^{\mathbf{n}}_{y,it}$;

(b) or $\chi^{\mathbf{n}}_{xy,it}$ (obtained in Section 3.1) is projected, leading, up to quadratic mean negligible quantities, to the same result, as the difference $Y_{it} - \chi^{\mathbf{n}}_{xy,it}$ is $\xi^{\mathbf{n}}_{xy,it}$, which consistently reconstructs the strongly idiosyncratic $\xi_{xy,it}$;

(c) or $\chi^{n_y}_{y,it}$ (obtained in Section 3.2) is projected, immediately providing the result $\phi^{\mathbf{n}}_{y,it}$, since $\chi^{n_y}_{y,it} = \phi_{y,it} + \psi_{y,it}$, where $\psi_{y,it}$ is z-idiosyncratic.

For the sake of simplicity, as all these projections eventually coincide, we concentrate on projection (a).

The following result is adapted from Theorem 8.3.1 in Brillinger (1981), and provides the explicit form of such projections.

Proposition 4. *Assume that the $(r + s)$ vector valued second-order mean zero stationary process $\{(\boldsymbol{\zeta}'_t, \boldsymbol{\eta}'_t)',\ t \in \mathbb{Z}\}$ is such that the spectral density matrix $\mathbf{f}_{\boldsymbol{\eta}\boldsymbol{\eta}}(\theta)$ of $\boldsymbol{\eta}_t$, is nonsingular. Then, the projection of $\boldsymbol{\zeta}_t$ onto the closed space $\mathcal{H}_\eta$ spanned by $\{\boldsymbol{\eta}_t\ , t \in \mathbb{Z}\}$ —that is, the r-tuple $\mathbf{A}(L)\boldsymbol{\eta}_t$ of square summable linear combinations of the present, past and future of $\boldsymbol{\eta}_t$ minimizing $\mathrm{E}[(\boldsymbol{\zeta}_t - \mathbf{A}(L)\boldsymbol{\eta}_t)(\boldsymbol{\zeta}_t - \mathbf{A}(L)\boldsymbol{\eta}_t)']$ is $\underline{\mathbf{f}}_{\boldsymbol{\zeta}\boldsymbol{\eta}}(L)\underline{\mathbf{f}}^{-1}_{\boldsymbol{\eta}\boldsymbol{\eta}}(L)\boldsymbol{\eta}_t$, where*

$$\underline{\mathbf{f}}_{\boldsymbol{\zeta}\boldsymbol{\eta}}(L) := \frac{1}{2\pi}\sum_{s=-\infty}^{\infty}\left[\int_{-\pi}^{\pi}\mathbf{f}_{\boldsymbol{\zeta}\boldsymbol{\eta}}(\theta)e^{\mathrm{i}s\theta}\, d\theta\right]L^s \quad \text{and} \quad \underline{\mathbf{f}}^{-1}_{\boldsymbol{\eta}\boldsymbol{\eta}}(L) := \frac{1}{2\pi}\sum_{s=-\infty}^{\infty}\left[\int_{-\pi}^{\pi}[\mathbf{f}_{\boldsymbol{\eta}\boldsymbol{\eta}}(\theta)]^{-1}e^{\mathrm{i}s\theta}\, d\theta\right]L^s,$$

and $\mathbf{f}_{\boldsymbol{\zeta}\boldsymbol{\eta}}(\theta)$ denotes the cross-spectrum of $\boldsymbol{\zeta}_t$ and $\boldsymbol{\eta}_t$. vspace2mm

Actually, Brillinger also requires $(\boldsymbol{\zeta}'_t, \boldsymbol{\eta}'_t)'$ to have absolutely summable autocovariances, so that the filter $\underline{\mathbf{f}}_{\boldsymbol{\zeta}\boldsymbol{\eta}}(L)\underline{\mathbf{f}}^{-1}_{\boldsymbol{\eta}\boldsymbol{\eta}}(L)$ also is absolutely summable. We, however, do not need this here.

Now, the z-common space $\mathcal{H}^{\chi}_z$ on which we have to project Y_{it} has reduced dimension $q_z < n_z$, and Proposition 4 thus does not directly solve our problem. Nor does it apply it to the finite-sample reconstruction of $\mathcal{H}^{\chi}_z$ based on $\chi_{z;n_z,t}$, as the spectral density $\mathbf{f}_{\boldsymbol{\eta}\boldsymbol{\eta}}(\theta)$, for $\boldsymbol{\eta}_t = \chi_{z;n_z,t}$, is singular. Fortunately, a full rank q_z-dimensional random vector spanning the same space as $\chi_{z;n_z,t}$ is available : the q_z-tuple

$$\mathbf{V}'_{z;n_z,t} := \left(V_{z;n_z,1t}, \ldots, V_{z;n_z,q_y t}\right)' \quad \text{where} \quad V_{z;n_z,kt} := \underline{\mathbf{p}}^*_{z;n_z,k}(L)\mathbf{Z}_{n_z,t},$$

of $\mathbf{\Sigma}_{z;n_z}(\theta)$'s first dynamic principal components, which are mutually orthogonal. Proposition 4 thus applies to the (n_y+q_z) random vector $(\mathbf{Y}'_{n_y,t}, \mathbf{V}'_{z;n_z,t})'$. The spectral matrix for that vector is

$$\begin{pmatrix} \mathbf{\Sigma}_{\mathbf{YY}}(\theta) & \mathbf{\Sigma}_{\mathbf{YV}}(\theta) \\ \mathbf{\Sigma}_{\mathbf{VY}}(\theta) & \mathbf{\Sigma}_{\mathbf{VV}}(\theta) \end{pmatrix}$$

with $\mathbf{\Sigma}_{\mathbf{YY}}(\theta) = \mathbf{\Sigma}_{y;n_y}(\theta)$ $(n_y \times n_y)$, $\mathbf{\Sigma}_{\mathbf{YV}}(\theta) = \mathbf{\Sigma}_{yz;\mathbf{n}}(\theta)(\mathbf{p}_{z;n_z,1}(\theta), \ldots, \mathbf{p}_{z;n_z,q_z}(\theta))$ $(n_y \times q_z)$, and (since the principal components $V_{z;n_z,kt}$'s, with spectral densities $\lambda_{z;n_z,k}(\theta)$, are mutually orthogonal) $\mathbf{\Sigma}_{\mathbf{VV}}(\theta) = \text{diag}(\lambda_{z;n_z,1}(\theta), \ldots, \lambda_{z;n_z,q_z}(\theta))$ $(q_z \times q_z)$.

This yields, for Y_{it}, a projection (which we propose as a reconstruction of the z-common component $\phi_{y,it} + \nu_{y,it}$ of Y_{it})

$$\begin{aligned} \chi^{n_z}_{yz,it} &:= \left(\underline{\sigma}_{i1}(L), \ldots, \underline{\sigma}_{in_z}(L)\right)\left(\underline{\mathbf{p}}_{z;n_z,1}(L), \ldots, \underline{\mathbf{p}}_{z;n_z,q_z}(L)\right)\text{diag}(\underline{\lambda}^{-1}_{z;n_z,1}(L), \ldots, \underline{\lambda}^{-1}_{z;n_z,q_z}(L))\mathbf{V}_{z;n_z t} \\ &= \left(\underline{\sigma}_{i1}(L), \ldots, \underline{\sigma}_{in_z}(L)\right)\sum_{k=1}^{q_z} \underline{\lambda}^{-1}_{z;n_z,k}(L)\underline{\mathbf{p}}_{z;n_z,k}(L)\underline{\mathbf{p}}^*_{z;n_z,k}(L)\mathbf{Z}_{n_z,t} \\ &=: \underline{\mathbf{H}}^*_{y;n_z,i}(L)\mathbf{Z}_{n_z,t} \qquad (3.20) \end{aligned}$$

where (denoting by $\sigma_{ij}(\theta)$ the element in $\mathbf{\Sigma}_{\mathbf{n}}(\theta)$ corresponding to the cross-spectrum of Y_{it} and Z_{jt})

$$\underline{\sigma}_{ij}(L) := \frac{1}{2\pi}\sum_{s=-\infty}^{\infty}\left[\int_{-\pi}^{\pi}\sigma_{ij}(\theta)e^{\text{i}s\theta}\, d\theta\right]L^s \quad \text{and} \quad \underline{\lambda}^{-1}_{z;n_z,k}(L) := \frac{1}{2\pi}\sum_{s=-\infty}^{\infty}\left[\int_{-\pi}^{\pi}[\lambda_{z;n_z,k}(\theta)]^{-1}e^{\text{i}s\theta}\, d\theta\right]L^s$$

($|\lambda_{z;n_z,k}(\theta)|$, $k = 1, \ldots, q_z$ safely can be assumed to be θ-a.e. larger than one: see p. 551 of Forni *et al* (2001), Assumption (A) and the comments thereafter; the filters associated with their inverses then are well defined, and square summable).

Our reconstruction of Y_{it}'s strongly common component then is

$$\phi^{\mathbf{n}}_{y,it} := \underline{\mathbf{H}}^*_{y;n_z,i}(L)\mathbf{Z}_{n_z,t} - \nu^{\mathbf{n}}_{y,it}.$$

Similar definitions, with obvious changes, are made for $\phi^{\mathbf{n}}_{z,jt}$. Parallel to Propositions 2 and 3, we then have the following consistency result for $\phi^{\mathbf{n}}_{y,it}$ and $\phi^{\mathbf{n}}_{z,jt}$.

PROPOSITION 5. *Let Assumptions A1, A2, and A3 hold. Then*

$$\lim_{\min(n_y,n_z)\to\infty} \phi^{\mathbf{n}}_{y,it} = \phi_{y;it} \quad \textit{and} \quad \lim_{\min(n_y,n_z)\to\infty} \phi^{\mathbf{n}}_{z,jt} = \phi_{z,jt}$$

in quadratic mean for any i, j, and t.

PROOF. See the appendix. □

It follows from (3.20) that the spectral density of $\chi^{n_z}_{yz,it}$ is of the form

$$\sum_{k=1}^{q_z} |\boldsymbol{\Sigma}_{YV;ik}(\theta)|^2 \, \lambda^{-1}_{z;n_z,k}(\theta);$$

the variance of $\chi^{n_z}_{yz,it}$ therefore writes

$$\mathrm{Var}(\chi^{n_z}_{yz,it}) = \int_{-\pi}^{\pi} \sum_{k=1}^{q_z} \left| \sum_{j=1}^{n_z} \sigma_{ij}(\theta) p_{z;j,k}(\theta) \right|^2 \lambda^{-1}_{z;n_z,k}(\theta) d\theta \tag{3.21}$$

as $\boldsymbol{\Sigma}_{YV;ik}(\theta) = \sum_{j=1}^{n_z} \sigma_{ij}(\theta) p_{z;j,k}(\theta)$. Since $\chi^{n_z}_{yz,it}$ decomposes into the sum of $\phi^{\mathbf{n}}_{y,it}$ and $\nu^{\mathbf{n}}_{y;it}$, which are mutually orthogonal, the reconstructed strongly common component $\phi^{\mathbf{n}}_{y,it}$, in view of (3.19), has variance

$$\begin{aligned} \mathrm{Var}(\phi^{\mathbf{n}}_{y,it}) \;\; = \;\; & \int_{-\pi}^{\pi} \sum_{k=1}^{q_z} \left| \sum_{j=1}^{n_z} \sigma_{ij}(\theta) p_{z;j,k}(\theta) \right|^2 \lambda^{-1}_{z;n_z,k}(\theta) d\theta - \sum_{k=1}^{q} \int_{-\pi}^{\pi} |p_{\mathbf{n},k,i}(\theta)|^2 \lambda_{\mathbf{n},k}(\theta) d\theta \\ & + \sum_{k=1}^{q_y} \int_{-\pi}^{\pi} |p_{y;n_y,k,i}(\theta)|^2 \lambda_{y;n_y,k}(\theta) d\theta. \end{aligned}$$

Averaged over the subpanel, this yields

$$\begin{aligned} \frac{1}{n_y} \sum_{i=1}^{n_y} \mathrm{Var}(\phi^{\mathbf{n}}_{y,it}) \;\; = \;\; & \frac{1}{n_y} \sum_{i=1}^{n_y} \int_{-\pi}^{\pi} \sum_{k=1}^{q_z} \left| \sum_{j=1}^{n_z} \sigma_{ij}(\theta) p_{z;j,k}(\theta) \right|^2 \lambda^{-1}_{z;n_z,k}(\theta) d\theta \\ & - \frac{1}{n_y} \sum_{i=1}^{n_y} \sum_{k=1}^{q} \int_{-\pi}^{\pi} |p_{\mathbf{n},k,i}(\theta)|^2 \lambda_{\mathbf{n},k}(\theta) d\theta + \frac{1}{n_y} \sum_{k=1}^{q_y} \int_{-\pi}^{\pi} \lambda_{y;n_y,k}(\theta) d\theta, \end{aligned}$$

which measures the contribution of the strongly common factors in the total variation of the y-subpanel. Similar quantities are easily computed for the z-subpanel.

Consistent reconstructions of the weakly common components now readily follow by taking differences :

$$\psi^{\mathbf{n}}_{y,it} := Y_{it} - \chi^{n_y}_{yz,it} - \xi^{\mathbf{n}}_{xy;it} \qquad \text{and} \qquad \psi^{\mathbf{n}}_{z,jt} := Z_{jt} - \chi^{n_z}_{zy,jt} - \xi^{\mathbf{n}}_{xz;jt}.$$

The contributions of those weakly common components to the total variation are be obtained along the same lines as above, i.e. for $\psi^{\mathbf{n}}_{y,it}$ we get

$$\begin{aligned} \frac{1}{n_y} \sum_{i=1}^{n_y} \mathrm{Var}(\psi^{\mathbf{n}}_{y,it}) \;\; = \;\; & \frac{1}{n_y} \sum_{i=1}^{n_y} \mathrm{Var}(Y_{it}) - \frac{1}{n_y} \sum_{i=1}^{n_y} \int_{-\pi}^{\pi} \sum_{k=1}^{q_z} \left| \sum_{j=1}^{n_z} \sigma_{ij}(\theta) p_{z;j,k}(\theta) \right|^2 \lambda^{-1}_{z;n_z,k}(\theta) d\theta \\ & - \frac{1}{n_y} \sum_{i=1}^{n_y} \sum_{k=1}^{q} \int_{-\pi}^{\pi} |p_{\mathbf{n},k,i}(\theta)|^2 \lambda_{\mathbf{n},k}(\theta) d\theta. \end{aligned}$$

Dividing by $\frac{1}{n_y} \sum_{i=1}^{n_y} \mathrm{Var}(Y_{it}) = \frac{1}{n_y} \sum_{i=1}^{n_y} \int_{-\pi}^{\pi} \lambda_{y;n_y,k}(\theta) d\theta$ yields the correponding relative quantities. Up to obvious changes, the formulas for the z-subpanel are identical.

4 Recovering the factor structure; estimation results

The previous section shows how all components of Y_{it} and Z_{jt} can be recovered asymptotically as $\min(n_y, n_z) \to \infty$, provided that the spectral density $\mathbf{\Sigma_n}$ and the numbers q, q_y, and q_z of factors are known. The estimates $\phi^{\mathbf{n}}_{y;it}$, $\psi^{\mathbf{n}}_{y;it}$ and $\nu^{\mathbf{n}}_{y;it}$ all take the form of a filtered series of the observed process $\mathbf{X}_{\mathbf{n},t}$. We have indeed

$$
\begin{aligned}
\phi^{\mathbf{n}}_{y,it} &= \underline{\mathbf{H}}^*_{y;n_z,i}(L)\mathbf{Z}_{n_z,t} - \nu^{\mathbf{n}}_{y,it} \\
&= \underline{\mathbf{H}}^*_{y;n_z,i}(L)\mathbf{Z}_{n_z,t} + \underline{\mathbf{G}}^*_{y;n_y,i}(L)\mathbf{Y}_{n_y,t} - \underline{\mathbf{K}}^*_{y;\mathbf{n},i}(L)\mathbf{X}_{\mathbf{n},t} \\
&= \left[(\underline{\mathbf{G}}^*_{y;n_y,i}(L),\ \underline{\mathbf{H}}^*_{y;n_z,i}(L)) - \underline{\mathbf{K}}^*_{y;\mathbf{n},i}(L)\right]\mathbf{X}_{\mathbf{n},t} \\
&=: \underline{\mathbf{K}}^*_{\phi_y;\mathbf{n},i}(L)\mathbf{X}_{\mathbf{n},t}, \\
\psi^{\mathbf{n}}_{y,it} &= \chi^{n_y}_{y,it} - \phi^{\mathbf{n}}_{y,it} \\
&= \underline{\mathbf{G}}^*_{y;n_y,i}(L)\mathbf{Y}_{n_y,t} - \underline{\mathbf{K}}^*_{\phi_y;\mathbf{n},i}(L)\mathbf{X}_{\mathbf{n},t} \\
&= \left[\underline{\mathbf{K}}^*_{y;\mathbf{n},i}(L) - (\mathbf{0},\ \underline{\mathbf{H}}^*_{y;n_z,i}(L))\right]\mathbf{X}_{\mathbf{n},t} \\
&=: \underline{\mathbf{K}}^*_{\psi_y;\mathbf{n},i}(L)\mathbf{X}_{\mathbf{n},t}, \quad \text{and} \\
\nu^{\mathbf{n}}_{y;it} &= \chi^{\mathbf{n}}_{xy;it} - \chi^{n_y}_{y;it} \\
&= \underline{\mathbf{K}}^*_{y;\mathbf{n},i}(L)\mathbf{X}_{\mathbf{n},t} - \underline{\mathbf{G}}^*_{y;n_y,i}(L)\mathbf{Y}_{n_y,t} \\
&= \left[\underline{\mathbf{K}}^*_{y;\mathbf{n},i}(L) - (\underline{\mathbf{G}}^*_{y;n_y,i}(L),\ \mathbf{0})\right]\mathbf{X}_{\mathbf{n},t} \\
&=: \underline{\mathbf{K}}^*_{\nu_y;\mathbf{n},i}(L)\mathbf{X}_{\mathbf{n},t},
\end{aligned}
$$

with

$$
\begin{aligned}
\underline{\mathbf{K}}^*_{\phi_y;\mathbf{n},i}(L) &:= \left[(\underline{\mathbf{G}}^*_{y;n_y,i}(L),\ \underline{\mathbf{H}}^*_{y;n_z,i}(L)) - \underline{\mathbf{K}}^*_{y;\mathbf{n},i}(L)\right], \\
\underline{\mathbf{K}}^*_{\psi_y;\mathbf{n},i}(L) &:= \left[\underline{\mathbf{K}}^*_{y;\mathbf{n},i}(L) - (\mathbf{0},\ \underline{\mathbf{H}}^*_{y;n_z,i}(L))\right], \quad \text{and} \\
\underline{\mathbf{K}}^*_{\nu_y;\mathbf{n},i}(L) &:= \left[\underline{\mathbf{K}}^*_{y;\mathbf{n},i}(L) - (\underline{\mathbf{G}}^*_{y;n_y,i}(L),\ \mathbf{0})\right].
\end{aligned}
$$

These three filters all are functions of the spectral density matrix $\mathbf{\Sigma_n}(\theta)$ which of course in practice is unknown, as we only observe a finite realization $\mathbf{X}^T_{\mathbf{n}} := (\mathbf{X}_{\mathbf{n}1}, \mathbf{X}_{\mathbf{n}2}, \ldots, \mathbf{X}_{\mathbf{n}T})$ of $\mathbf{X_n}$.

Since its actual value is unknown, we need an estimator $\mathbf{\Sigma}^T_{\mathbf{n}}(\theta)$ of $\mathbf{\Sigma_n}(\theta)$. Consistent estimation of the spectral density $\mathbf{\Sigma_n}(\theta)$ requires strengthening slightly Assumption A1 into the following Assumption A1′:

ASSUMPTION A1′. *For all* $\mathbf{n}$, *the vector process* $\{\mathbf{X}_{\mathbf{n},t}; t \in \mathbb{Z}\}$ *admits a Wold representation of the form* $\mathbf{X}_{\mathbf{n},t} = \sum_{k=-\infty}^{\infty} \mathbf{C}_k \boldsymbol{\zeta}_{t-k}$, *where* $\boldsymbol{\zeta}_t$ *is full-rank* n*-dimensional white noise with finite fourth order moments, and the* $n \times n$ *matrices* $\mathbf{C}_k = (C_{ij,k})$ *are such that* $\sum_{k=-\infty}^{\infty} |k||C_{ij,k}|^{1/2} < \infty$ *for all* i, j.

Under Assumption A1′, if $\mathbf{\Sigma}^T_{\mathbf{n}}(\theta)$, with elements $\sigma^T_{\mathbf{n},ij}(\theta)$, denotes any periodogram-smoothing or lag-window estimator of $\mathbf{\Sigma_n}(\theta)$, we have, for all $\mathbf{n}$, i, j, and $\varepsilon > 0$,

$$
\lim_{T\to\infty} \mathrm{P}\left[\sup_{\theta\in[-\pi,\pi]} \left|\sigma^T_{\mathbf{n},ij}(\theta) - \sigma_{ij}(\theta)\right| > \varepsilon\right] = 0
$$

(see e.g. Brockwell and Davis 1987, p. 433). In Section 6, we consider lag-window estimators of the form

$$\boldsymbol{\Sigma}^T_{\mathbf{n}}(\theta) := \sum_{k=-M_T}^{M_T} \boldsymbol{\Gamma}^T_{\mathbf{n}k}\omega_k e^{-ik\theta} \tag{4.22}$$

where $\boldsymbol{\Gamma}^T_{\mathbf{n}k}$ is the sample covariance matrix of $\mathbf{X}_{\mathbf{n},t}$ and $\mathbf{X}_{\mathbf{n},t-k}$ and $\omega_k := 1 - |k|/(M_T + 1)$ are the weights corresponding to the Bartlett lag window of size M_T. Consistency then is achieved provided that the following assumption holds:

Assumption B. *$M_T \to \infty$, and $M_T T^{-1} \to 0$, as $T \to \infty$.*

A consistent estimator $\boldsymbol{\Sigma}^T_{\mathbf{n}}(\theta)$ of $\boldsymbol{\Sigma}_{\mathbf{n}}(\theta)$ however is not sufficient here. Deriving, from this estimator$\boldsymbol{\Sigma}^T_{\mathbf{n}}(\theta)$, estimated versions $\underline{\mathbf{K}}^T_{\phi_y;\mathbf{n},i}(L)$, $\underline{\mathbf{K}}^T_{\psi_y;\mathbf{n},i}(L)$ and $\underline{\mathbf{K}}^T_{\nu_y;\mathbf{n},i}(L)$, of the filters $\underline{\mathbf{K}}_{\phi_y;\mathbf{n},i}(L)$, $\underline{\mathbf{K}}_{\psi_y;\mathbf{n},i}(L)$ and $\underline{\mathbf{K}}_{\nu_y;\mathbf{n},i}(L)$ indeed also requires an estimation of the numbers of factors q, q_y and q_z involved. The only method allowing for such estimation is the idendification method developed in Hallin and and Liška (2007), which we now briefly describe, with a few adjustments taking into account the particular notation of this paper. For a detailed description of the procedure, we refer to the section entitled "A practical guide to the selection of q" in Hallin and Liška (2007).

The lag window method described in (4.22) provides estimations $\boldsymbol{\Sigma}^T_{\mathbf{n}}(\theta_l)$ of the spectral density at frequencies $\theta_l := \pi l/(M_T + 1/2)$ for $l = -M_T, \ldots, M_T$. Based on these estimations, consider the information criterion

$$IC^T_{\mathbf{n};c}(k) := \log\left[\frac{1}{n}\sum_{i=k+1}^{n}\frac{1}{2M_T+1}\sum_{l=-M_T}^{M_T}\lambda^T_{\mathbf{n}i}(\theta_l)\right] + kcp(n,T), \quad 0 \le k \le q_{\max}, \quad c \in \mathbb{R}^+_0, \tag{4.23}$$

where the penalty function $p(n,T)$ is $o(1)$ while $p^{-1}(n,T) = o\left(\min(n, M_T^2, M_T^{-1/2}T^{1/2}\right)$ as both n and T tend to infinity, and $q_{\max}$ is some predetermined upper bound; the eigenvalues $\lambda^T_{\mathbf{n}i}(\theta_l)$ are those of $\boldsymbol{\Sigma}^T_{\mathbf{n}}(\theta_l)$. Depending on $c > 0$, the estimated number of factors, for given $\mathbf{n}$ and T, is

$$q^T_{\mathbf{n};c} := \mathrm{argmin}_{0\le k\le q_{\max}} IC^T_{\mathbf{n};c}(k). \tag{4.24}$$

Hallin and Liška (2007) prove that this $q^T_{\mathbf{n};c}$ is consistent for any $c > 0$. An "optimal" value c^* of c is then selected as follows. Consider a $J-$tuple of the form $q^{T_j}_{c,\mathbf{n}_j}$, $j = 1, \ldots, J$, where $\mathbf{n}_j = (n_{y;j}, n_{z;j})$ with $0 < n_{y;1} < \ldots < n_{y;J} = n_y$, $0 < n_{z;1} < \ldots < n_{z;J} = n_z$, and $0 < T_1 \le \ldots \le T_J = T$. This $J-$tuple can be interpreted as a "history" of the identification procedure, and characterizes, for each $c > 0$, a sequence $q^{T_j}_{c,\mathbf{n}_j}$, $j = 1, \ldots, J$ of estimated factor numbers. In order to keep a balanced representation of the two blocks, we only consider $J-$tuples along which $n_{y;j}/n_{z;j}$ is as close as possible to n_y/n_z.

The selection of c^* is based on the inspection of two mappings: $c \to q^T_{\mathbf{n};c}$, and $c \to S_c$, where $S^2_c := J^{-1}\sum_{j=1}^J(q^{T_j}_{\mathbf{n}_j;c} - J^{-1}\sum_{j=1}^J q^{T_j}_{\mathbf{n}_j}; c)^2$ measures the variability of $q^{T_j}_{\mathbf{n}_j;c}$ over the "history". For n and T large enough, S_c exhibits "stability intervals", that is, intervals of c values over which $S_c = 0$. The definition of S_c implies that $c \mapsto q^T_{\mathbf{n};c}$ is constant over such intervals. Starting in the neighborhood of $c = 0$, a first stability interval $(0, c_1^+)$ corresponds to $q_{\mathbf{n};c}^T = q_{\max}$; choose c^* as any point in the next one, (c_2^-, c_2^+). The selected number of factors is then $q_{\mathbf{n}}^T = q_{\mathbf{n};c^*}^T$. The same method, applied to the Y- and Z-subpanels, yields estimators $q_{n_y}^T$ and $q_{n_z}^T$ of q_y and q_z; $q_{\mathbf{n};yz}^T := q_{n_y}^T + q_{n_z}^T - q_{\mathbf{n}}^T$ provides a consistent estimator of q_{yz}.

The success of this identification method however also requires strengthening somewhat the assumptions; from now on, we reinforce Assumption A1$'$ into Assumption A1$''$ and Assumptions A2 and A3 into Assumptions A2$'$ and A3$'$:

Assumption A1$''$. *Same as Assumption A1$'$, but (i) the convergence condition on the $C_{ij,k}$'s is uniform,* $\sup_{i,j\in\mathbb{N}} \sum_{k=-\infty}^{\infty} |C_{ij,k}||k|^{1/2} < \infty$, *and, for all* $1 \leq \ell \leq 4$ *and* $1 \leq j < \ell$, $\sup_{i_1,\ldots,i_\ell} \left[\sum_{k_1=-\infty}^{\infty} \cdots \sum_{k_{\ell-1}=-\infty}^{\infty} |c_{i_1,\ldots,i_\ell}(k_1,\ldots,k_{\ell-1})|\right] < \infty$.

Assumption A2$'$. *The entries $\sigma_{ij}(\theta)$ of $\mathbf{\Sigma_n}(\theta)$ (i) are bounded, uniformly in $\mathbf{n}$ and θ—that is, there exists a real $c > 0$ such that $\sigma_{ij}(\theta) \leq c$ for any $i, j \in \mathbb{N}$ and $\theta \in [-\pi, \pi]$—and (ii) they have bounded, uniformly in $\mathbf{n}$ and θ, derivatives up to the order two—namely, there exists $Q < \infty$ such that* $\sup_{i,j\in\mathbb{N}} \sup_\theta \left|\frac{d^k}{d\theta^k}\sigma_{ij}(\theta)\right| \leq Q$, $k = 0, 1, 2$.

Assumption A3$'$. *Same as Assumption A3, but moreover*

(i) $\lambda_{y;n_y,q_y}(\theta)$ and $\lambda_{z;n_z,q_z}(\theta$ diverge at least linearly in n_y and n_z, respectively, that is, $\liminf_{n_y\to\infty} \inf_\theta n_y^{-1}\lambda_{y;n_y,q_y}(\theta) > 0$, *and* $\liminf_{n_z\to\infty} \inf_\theta n_z^{-1}\lambda_{z;n_z,q_z}(\theta) > 0$, *and*

(ii) both n_y/n_z and n_z/n_y are $O(1)$ as $\min(n_y, n_z) \to \infty$.

This "at least linear" divergence assumption is also made in Hallin and Liška (2007), and can be considered as a form of cross-sectional stability of the two panels under study.

Once estimated values of the numbers q, q_y and q_z of factors are available, the estimated counterparts of of $\underline{\mathbf{K}}_{\phi_y;\mathbf{n},i}(L)$, $\underline{\mathbf{K}}_{\psi_y;\mathbf{n},i}(L)$ and $\underline{\mathbf{K}}_{\nu_y;\mathbf{n},i}(L)$ are obtained by substituting $\mathbf{\Sigma}_{\mathbf{n}}^T(\theta)$, $q_{\mathbf{n}}^T$, $q_{n_y}^T$ and $q_{n_z}^T$ for $\mathbf{\Sigma_n}(\theta)$, q, q_y and q_z in all definitions of Section 3, then truncating infinite sums as explained in Section B of Forni *et al* (2000) (a truncation which depends on t, which explains the notation), yielding $\underline{\mathbf{K}}_{\phi_y;\mathbf{n},i}^{Tt}(L)$, $\underline{\mathbf{K}}_{\psi_y;\mathbf{n},i}^{Tt}(L)$ and $\underline{\mathbf{K}}_{\nu_y;\mathbf{n},i}^{Tt}(L)$. Parallel to Proposition 3 in Forni *et al* (2000), we then have the following result.

Proposition 6. *Let Assumption A1$''$, A2$'$, A3$'$, and B hold. Then, for all $\epsilon_k > 0$ and $\eta_k > 0$, $k = 1, 2, 3$, there exists $N_0(\epsilon_1, \epsilon_2, \epsilon_3, \eta_1, \eta_2, \eta_3)$ such that*

$$\mathrm{P}\left[\left|\underline{\mathbf{K}}_{\phi_y;\mathbf{n},i}^{Tt\ *}(L)\mathbf{X}_{\mathbf{n},t} - \phi_{y;it}\right| > \epsilon_1\right] \leq \eta_1, \quad \mathrm{P}\left[\left|\underline{\mathbf{K}}_{\psi_y;\mathbf{n},i}^{Tt\ *}(L)\mathbf{X}_{\mathbf{n},t} - \psi_{y;it}\right| > \epsilon_2\right] \leq \eta_2,$$

and

$$\mathrm{P}\left[\left|\underline{\mathbf{K}}_{\nu_y;\mathbf{n},i}^{Tt\ *}(L)\mathbf{X}_{\mathbf{n},t} - \nu_{y;it}\right| > \epsilon_3\right] \leq \eta_3,$$

for all $t = \check{t}(T)$ satisfying, for some a, b such that $0 < a < b < 1$,

$$a \leq \liminf_{T\to\infty} \frac{\check{t}(T)}{T} \leq \limsup_{T\to\infty} \frac{\check{t}(T)}{T} \leq b,$$

all $n \geq N_0$ and all T larger than some $T_0(\mathbf{n}, \epsilon_1, \epsilon_2, \epsilon_3, \eta_1, \eta_2, \eta_3)$.

Proof. The proof consists in reproducing, for each projection involved in the reconstruction of $\phi_{y;it}$, $\psi_{y;it}$ and $\nu_{y;it}$, the proof of Proposition 3 in Forni *et al* (2000). Lengthy but obvious details are left to the reader. $\square$

Consistent estimations of the various contributions to the total variance of each subpanel can be obtained either by substituting estimated spectral eigenvalues and eigenvectors for the exact ones in the formulas of Section 3, and replacing integrals with the corresponding finite sums over Fourier frequencies, or by computing the empirical variances of the estimated strongly and weakly common, strongly and weakly idiosyncratic components.

5 Dynamic factors in the presence of K blocks ($K > 2$).

The ideas developed in the previous sections readily extend to the more general case of $K \geq 2$ blocks. Instead of Y_{it} for the first block and Z_{jt} for the second one, denote all observations as X_{it} ($i = 1, \ldots, n$), with an additional label $[k]$ indicating, when needed, that X_{it} belongs to block k, $k = 1, \ldots, K$: the notation $X_{[1];1t}$ for instance means that the first series in the panel belongs to the first block. Marginal *k-common* and *k-idiosyncratic* spaces are defined in an obvious manner by considering the kth block as an individual subpanel. The number of mutually orthogonal components in the decomposition (2.7) of each observation $X_{[k];it}$ however increases exponentially with K, and the general case requires 2^K distinct components, with somewhat heavy notation: for each i and t, $X_{it} = X_{[k];it}$ decomposes into

(a) one *strongly common component* $\phi_{[k];it}$, denoting the projection of $X_{[k];it}$ on the intersection of the K marginal common spaces,

(b) $2^{K-1}-1$ *weakly common components*, of the form $\psi_{[k]\ (k,k_1,\ldots,k_\ell).(k_{\ell+1},\ldots,k_{K-1});it}$, denoting the projection of $X_{[k];it}$ on the intersection of the k-, k_1-, ..., k_ℓ-common and $k_{\ell+1}$-, ..., k_{K-1}-idiosyncratic spaces, where $\Big(\{k_1, \ldots, k_\ell\},\ \{k_{\ell+1}, \ldots, k_{K-1}\}\Big)$ ranges over all partitions of $\{1, \ldots, k-1, k+1, \ldots, K\}$ into two nonoverlapping nonempty subsets, $\ell = 0, 1, \ldots, K-2$;

(c) $2^{K-1}-1$ *weakly idiosyncratic components*, of the form $\nu_{[k]\ (k_1,\ldots,k_\ell).(k,k_{\ell+1},\ldots,k_{K-1});it}$, denoting the projection of $X_{[k];it}$ on the intersection of the k_1-, ..., k_ℓ-common and k-, $k_{\ell+1}$-, ..., k_{K-1}-idiosyncratic spaces, where $\Big(\{k_1, \ldots, k_\ell\},\ \{k_{\ell+1}, \ldots, k_{K-1}\}\Big)$ similarly ranges over all partitions of $\{1, \ldots, k-1, k+1, \ldots, K\}$ into two nonoverlapping nonempty subsets, $\ell = 1, \ldots, K-1$;

(d) one *strongly idiosyncratic component* $\xi_{[k];it}$, denoting the projection of $X_{[k];it}$ on the intersection of the K marginal idiosyncratic spaces.

In view of the notational burden, we will not pursue any further with formal developments, since it is clear that the methods previously described, with a well-designed sequence of projections, allow for a consistent reconstruction of all those components.

An application for $K = 3$ is considered in Section 6.2.

6 Real Data Applications

We applied our method to a dataset of monthly Industrial Production Indexes for France, Germany, and Italy, observed from January 1995 through December 2006. All data were preadjusted by taking a log-difference transformation ($T = 143$ throughout—one observation is lost due to differencing), then centered and normalized using their sample means and standard errors. A full description of the panels is given in Table 7.2.

6.1 A two-block analysis

First consider the data for France and Germany. Using Y_{it} or the Fench data and Z_{jt} for the German, we have $n_y = n_F = 96$, $n_z = n_G = 114$, hence $n = n_{FG} = 210$. Spectral densities were estimated from the pooled panel using a lag-window estimators of the form (4.22), with truncation parameter $M_T = 0.5\sqrt{T} = 5$. Based on this estimation, we ran the Hallin and Liška (2007) identification method on the French and German subpanels, with sequences $n_{F,j} = 96 - 2j$, $j = 1, \ldots,\ 5$ and $n_{G,j} = 96 - 2j$, $j = 1, \ldots,\ 5$, respectively, then on the pooled panel, with sequence $n_{FG,j} = 210-2j$, $j = 1, \ldots, 8$ and an "almost constant " proportion 96/210, 114/210 of French and German observations (namely, $\lceil 96 n_{FG,j}/210 \rceil$ French observations, and $\lfloor 114 n_{FG,j}/210 \rfloor$ German ones. In all cases, we put $T_j = T = 143$, $j = 1, \ldots,\ 5$. The range for c values, after some preliminary exploration, was taken as $[0, 0.0002, 0.0004, \ldots, 0.5]$, and $q_{\max}$ was set to 10. In all cases, the panels were randomly ordered prior to the analysis. The penalty function was $p(n,T) = \left(\min\left[n, M_T^2, M_T^{-1/2} T^{1/2}\right]\right)^{-1/2}$.

The results are shown in Figure 6.1, and very clearly conclude for $q^T_{(n_F,n_G)} = 3$ (for $c \in [0.1798, 0.1894]$), $q^T_{n_F,F} = 2$ (for $c \in [0.2222, 0.2344]$), and $q^T_{n_G,G} = 3$ (for $c \in [0.2032, 0.2138]$). This identification of 3 joint common factors, 3 German-common and 2 French-common factors also provides an estimation of 2 strongly common factors (as $q_{yz} = q_y + q_z - q$). The French-common factors thus are strongly common (no weakly common space), whereas one German-common factor is French-idiosyncratic.

Table 6.1 is summarizing these findings. For each of the mutually orthogonal subspaces appearing in the decomposition, we provide the percentage of total variation explained in each country. The two strongly common factors jointly account for 9.2 % of German total variability and 20.4 % of French total variability. Germany has an "all-German", French-idiosyncratic, common factor explaining 16.7 % of its total variance. Although French-idiosyncratic, that German factor nevertheless still accounts for 2.6 % of the French total variability. Estimated percentages of explained variation were obtained via estimated eigenvectors and eigenvalues.

6.2 A three-block analysis

Next consider the three-block case resulting from adding the corresponding Italian Industrial Production index, with $n_I = 91$ into the previous panel, yielding $K = 3$. The series length is

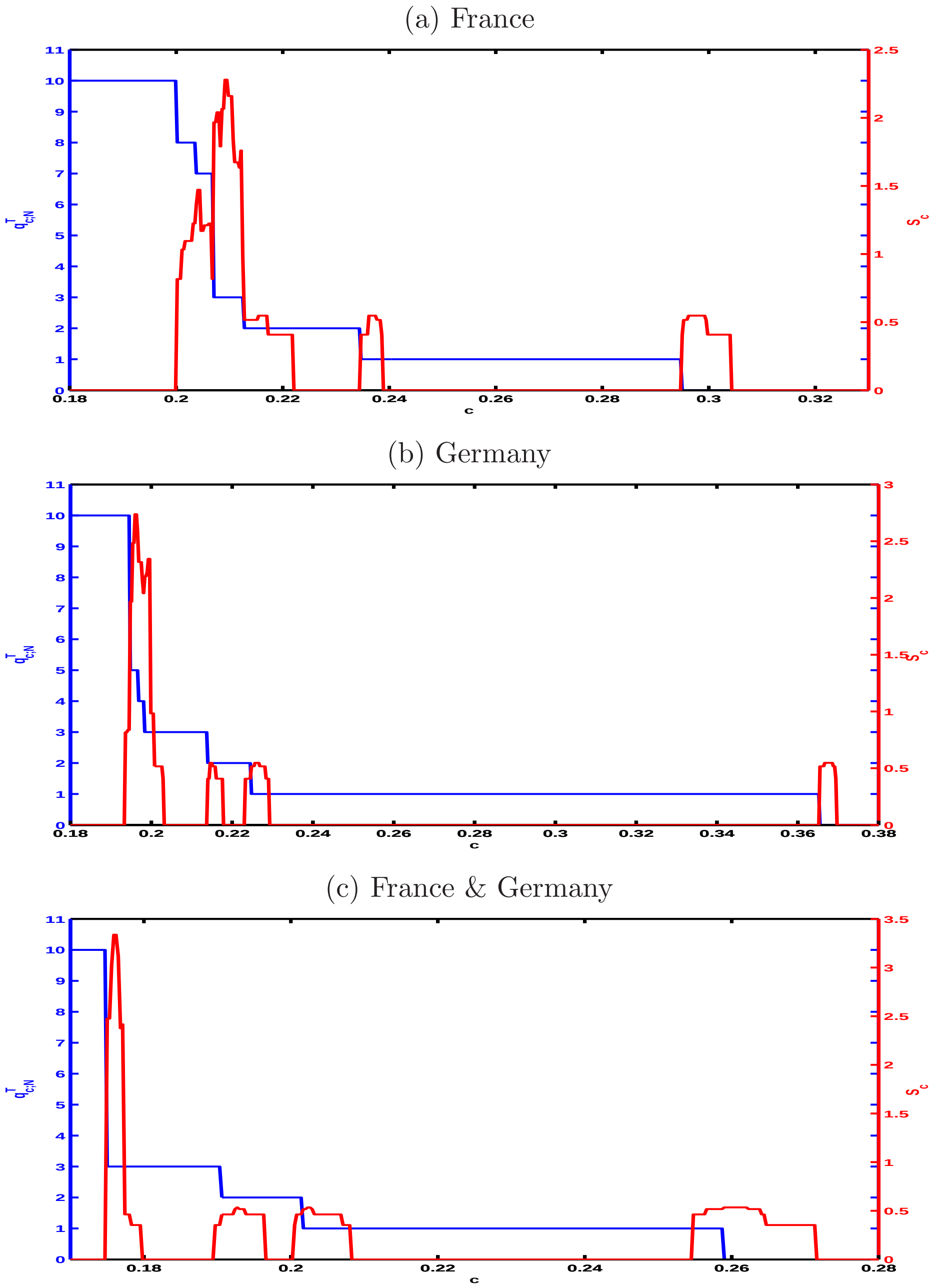

Figure 1: Identification of the numbers of factors for the France-Germany Industrial Production dataset. The three figures show he simultaneous plots of $c \mapsto S_c$ and $c \mapsto q^T_{c,n}$ needed for this identification, ((a) and (b)) in the marginal French and German subpanels, and (c) in the complete panel, respectively.

3 joint common factors	0 factor weakly F-common weakly G-idiosyncratic	2 factors strongly common	1 factor weakly F-idiosyncratic weakly G-common	strongly idiosyncratic
France	(ψ_F) 0 %	(ϕ_F) 20.4 %	(ν_F) 2.6 %	(ξ_F) 77.0 %
Germany	(ν_G) 0 %	(ϕ_G) 9.2 %	(ψ_G) 16.7 %	(ξ_G) 74.1 %

Table 1: Decomposition of the France-Germany panel data into four mutually orthogonal components, with the corresponding percentages of explained variation.

still $T = 143$. Adapting the notation of Section 5, let $X_{[F];it}$ correspond to the French, $X_{[G];it}$ to the German, and $X_{[I];it}$ to the Italian subpanel, respectively.

From the resulting panel ($n = n_{FGI} = 301$), we can extract seven subpanels—the three panels we already analysed in Section 6.1, one new one-block subpanel (the marginal Italian one, with $n_I = 91$) and two new two-block subpanels (the France-Italy one, with $n_{FI} = 187$ and the Germany-Italy one, with $n_{GI} = 205$, respectively). Analyzing these new subpanels along the same lines as in the previous section (with, using obvious notation, $n_{I,j} = 91 - 2j$, $j = 1, \dots, 5$, $n_{GI,j} = 191 - 2j$, $j = 1, \dots, 8$, $n_{FI,j} = 187 - 2j$, $j = 1, \dots, 8$, and $n_{FGI,j} = 301 - 2j$, $j = 1, \dots,$ 15), still with $M_T = 0.5\sqrt{T} = 5$, the same penalty function and the same $q_{\max} = 10$ as before, we obtain the results shown in the four graphs of Figure 6.2.

These graphs again very clearly allow for identifying a total umber of $q^T_{\mathbf{n},FGI} = 4$ joint common factors (for $c \in [0.1710, 0.1718]$), $q^T_{\mathbf{n},FI} = 3$ (for $c \in [0.1838, 0.1886]$), $q^T_{\mathbf{n},GI} = 4$ (for $c \in [0.1786, 0.1800]$), and $q^T_{n_I,I} = 2$ marginal Italian factor (for $c \in [0.2118, 0.22218]$). Along with the figures obtained in Section 6.1 for France and Germany, this leads to the results summarized in Figure 6.2. The space spanned by he three blocks now decomposes into eight mutually orthogonal subspaces: seven (jointly) common ones, namely, the strongly common (F,G,I-common), the F,G-common/I-idiosyncratic, the G,I-common/F-idiosyncratic, the F,I-common/G-idiosyncratic, the F-common/G,I-idiosyncratic, the G-common/F,I-idiosyncratic, the I-common/F,G-idiosyncratic one, and the strongly idiosyncratic (F,G,I-idiosyncratic) one. Since the total number of factors is 4, three at least of the common subspaces must have dimension zero (they only contain the origin). The relations between the various (dynamic) dimensions of the seven common spaces are very easily obtained; for instance

$$q_{(n_F,n_G),FG} = q_{n_F,F} + q_{n_G,G} - q_{(n_F,n_G)},$$

a relation which we already used in Section 6.1, or

$$q_{(n_F,n_G),FG} = q_{n_F,F} + q_{n_G,G} + q_{n_I,I} - q_{(n_F,n_G),FG} - q_{(n_F,n_I),FI} - q_{(n_G,n_I),GI} + q_{(n_F,n_G,n_I),FGI}.$$

A two-dimensional table however cannot display the various interrelations between the seven common subspaces, which we rather provide in the diagram shown in Figure 6.2, along with

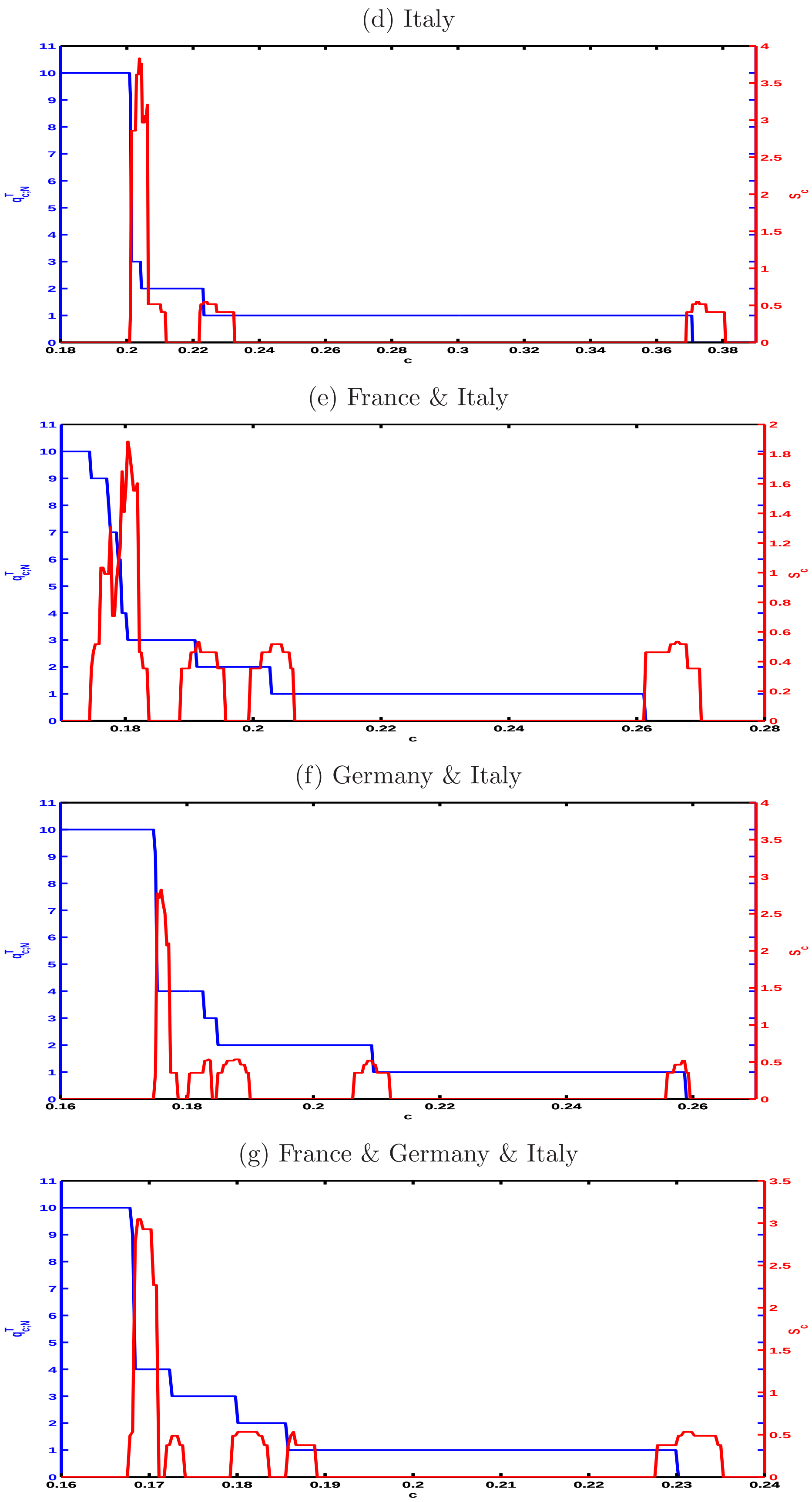

Figure 2: Identification of the numbers of factors for the France-Germany-Italy Industrial Production dataset. The four figures show he simultaneous plots of $c \mapsto S_c$ and $c \mapsto q^T_{c,n}$ needed for this identification: (d) for the marginal Italian subpanel, ((e) and (f)) for the France-Italy and Germany-Italy subpanels, and (g) for the complete three-country panel, respectively.

the various percentages of explained variances. Inspection of that diagram reveals that the three countries all exhibit a high percentage of about 60 % of strongly idiosyncratic variation. As already noted, France has no common components but the two shared with Germany and Italy (one), and with Germany alone (one). Both Italy and Germany have a "national common component. Italy's only "non national" common factor is the strongly common one, which is common to the three countries under study.

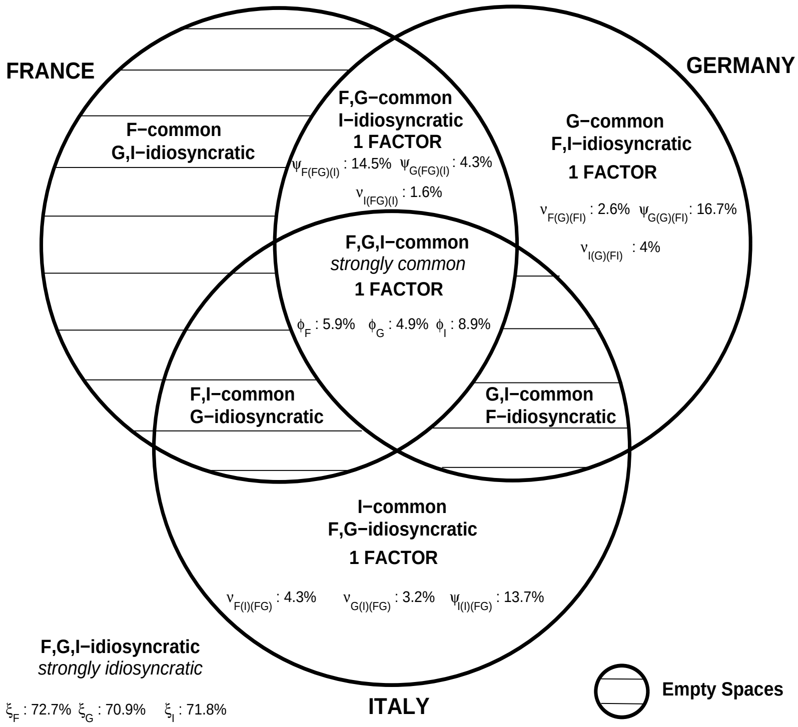

Figure 3: Decomposition of the France-Germany-Italy panel data into eight mutually orthogonal components, with the corresponding percentages of explained variation.

7 Appendix.

7.1 Proof of Lemma 1.

PROOF. Denote by $\bar{\Theta}_y$ the set (with Lebesgue measure zero) of θ values for which divergence in Assumption A2(i) does not hold. Similarly define $\bar{\Theta}_z$, and let $\bar{\Theta} := \bar{\Theta}_y \cup \bar{\Theta}_z$: $\bar{\Theta}$ also has Lebesgue measure zero. Since $\boldsymbol{\Sigma}_{y;n_y}(\theta)$ is a principal submatrix of $\boldsymbol{\Sigma}_{\mathbf{n}}(\theta)$, a classical result (see

Corollary 1, page 293, in Lancaster and Tismenetsky 1985) implies that, for any $\mathbf{n} = (n_y, n_z)$ and θ, $\lambda_{y;n_y,i}(\theta) \leq \lambda_{\mathbf{n},i}(\theta)$, $i = 1, \ldots, n_y$. Since $\lambda_{y;n_y,q_y}(\theta)$ diverges for all $\theta \in \Theta$ as $n_y \to \infty$, so does $\lambda_{\mathbf{n},q_y}(\theta)$. The same result of course also holds for the $\lambda_{z;n_z,j}$'s. It follows that, for all $\theta \in \Theta$, $\lambda_{\mathbf{n},\max(q_y,q_z)}(\theta)$ diverges as $\min(n_y, n_z)$ tends to infinity.

Note that the same result by Lancaster and Tismenetsky (1985) also implies that, for all θ and k, $\lambda_{\mathbf{n},k}(\theta)$ is a monotone nondecreasing function of both n_y and n_z and, therefore, either is bounded or goes to infinity as either n_y or $n_z \to \infty$.

Next, let us show that $\lambda_{\mathbf{n},q_y+q_z+1}(\theta)$ is bounded as $\min(n_y, n_z) \to \infty$, for all $\theta \in \Theta$. For all $\theta \in \Theta$, consider the sequences of n-dimensional vectors $\boldsymbol{\zeta_n}(\theta) := (\boldsymbol{\zeta}'_{y;n_y}(\theta), \boldsymbol{\xi}'_{z;n_z}(\theta))'$ which are orthogonal to the $q_y + q_z$ vectors $(\mathbf{p}'_{y;n_y,1}(\theta), 0, \ldots, 0)', \ldots, (\mathbf{p}'_{y;n_y,q_y}(\theta), 0, \ldots, 0)'$ and $(0, \ldots, 0, \mathbf{p}'_{z;n_z,1}(\theta))', \ldots, (0, \ldots, 0, \mathbf{p}'_{z;n_z,q_z}(\theta))'$. The collection of all such $\boldsymbol{\xi_n}$'s is a linear subspace $\boldsymbol{\Xi_n}(\theta)$ of dimension at least $n - q_y - q_z$. For any such $\boldsymbol{\xi_n}(\theta)$, in view of the orthogonality of $\boldsymbol{\xi}_{y;n_y}(\theta)$ and $\mathbf{p}_{y;n_y,1}(\theta), \ldots, \mathbf{p}_{y;n_y,q_y}(\theta)$ (resp., of $\boldsymbol{\xi}_{z;n_z}(\theta)$ and $\mathbf{p}_{z;n_z,1}(\theta), \ldots, \mathbf{p}_{z;n_z,q_z}(\theta)$),

$$
\begin{aligned}
&\|\boldsymbol{\xi_n}(\theta)\|^{-2}\boldsymbol{\xi}^*_{\boldsymbol{n}}(\theta)\boldsymbol{\Sigma}_{\mathbf{n}}(\theta)\boldsymbol{\xi_n}(\theta)\\
&\quad = \|\boldsymbol{\xi_n}\|^{-2}\boldsymbol{\xi}^*_{y;n_y}(\theta)\boldsymbol{\Sigma}_{y;n_y}(\theta)\boldsymbol{\xi}_{y;n_y}(\theta) + \|\boldsymbol{\xi_n}(\theta)\|^{-2}\boldsymbol{\xi}^*_{z;n_z}(\theta)\boldsymbol{\Sigma}_{z;n_z}(\theta)\boldsymbol{\xi}_{z;n_z}(\theta)\\
&\qquad + \|\boldsymbol{\xi_n}(\theta)\|^{-2}\boldsymbol{\xi}^*_{y;n_y}(\theta)\boldsymbol{\Sigma}_{yz;\mathbf{n}}(\theta)\boldsymbol{\xi}_{z;n_z}(\theta) + \|\boldsymbol{\xi_n}(\theta)\|^{-2}\boldsymbol{\xi}^*_{z;n_z}(\theta)\boldsymbol{\Sigma}_{zy,\mathbf{n}}(\theta)\boldsymbol{\xi}_{y;n_y}(\theta)\\
&\quad \leq 2\left(\|\boldsymbol{\xi}_{y;n_y}(\theta)\|^{-2}\boldsymbol{\xi}^*_{y;n_y}(\theta)\boldsymbol{\Sigma}_{y;n_y}\boldsymbol{\xi}_{y;n_y}(\theta) + \|\boldsymbol{\xi}_{z;n_z}(\theta)\|^{-2}\boldsymbol{\xi}^*_{z;n_z}(\theta)\boldsymbol{\Sigma}_{z;n_z}\boldsymbol{\xi}_{z;n_z}(\theta)\right)\\
&\quad \leq 2(\lambda^2_{y;n_y,q_y+1}(\theta) + \lambda^2_{z;n_z,q_z+1}(\theta))
\end{aligned}
$$

for all $\theta \in \Theta$ and $\mathbf{n} = (n_y, n_z)$. Since $\lambda^2_{y;n_y,q_y+1}(\theta)$ and $\lambda^2_{z,n_z,q_z+1}(\theta)$ are bounded, for any $\theta \in \Theta$, as $\min(n_y, n_z) \to \infty$, so is $\boldsymbol{\xi}^*_{\boldsymbol{n}}(\theta)\boldsymbol{\Sigma}_{\mathbf{n}}(\theta)\boldsymbol{\xi_n}(\theta)$. Hence, for all $\theta \in \Theta$ and $\mathbf{n} = (n_y, n_z)$, $\boldsymbol{\Xi_n}$ (with dimension at least $n - q_y - q_z$) is orthogonal to any eigenvector associated with a diverging sequence of eigenvalues of $\boldsymbol{\Sigma}_{\mathbf{n}}(\theta)$. It follows that the number of such eigenvalues cannot exceed$q_y + q_z$.

Summing up, for all $\theta \in \Theta$, the number of diverging eigenvalues of $\boldsymbol{\Sigma}_{\mathbf{n}}(\theta)$ is finite—denote it by q—and comprised between $\max(q_y, q_z)$ and $q_y + q_z$, as was to be shown. □

7.2 Proof of Proposition 5.

The proof of Proposition 5 is an extension of the proof of Proposition 2 in Forni et *et al* (2000).

We systematically denote by $\boldsymbol{\phi}_{y;n_y,t}$, $\boldsymbol{\chi}_{y;n_y,t}$, … column n_y-vectors of the form $(\phi_{y;1t}, \ldots, \phi_{y;n_yt})'$, $(\chi_{y;1t}, \ldots, \chi_{y;n_yt})'$, … ; these vectors thus belong to the "exact" strongly common, the "exact" y-weakly common, … spaces. The notation $\boldsymbol{\phi}^{\mathbf{n}}_{y;t}$, $\boldsymbol{\chi}^{n_y}_{y;t}$, … on the contrary is used for the corresponding "reconstructions" $(\phi^{\mathbf{n}}_{y;1t}, \ldots, \phi^{\mathbf{n}}_{y;n_yt})'$, $(\chi^{n_y}_{y;1t}, \ldots, \chi^{n_y}_{y;n_yt})'$, … ; these vectors which belong to the finite-(n_y, n_z) approximations of the same "exact" strongly common, "exact" y-weakly common, … spaces. Similar notation is used for $\mathbf{Z}_{n_z,t}$.

With this notation, each observation $\mathbf{Y}_{n_y,t}$, for given $\mathbf{n} = (n_y, n_z)$, decomposes into

$$\mathbf{Y}_{n_y,t} = \boldsymbol{\phi}_{y;n_y,t} + \boldsymbol{\psi}_{y;n_y,t} + \boldsymbol{\nu}_{y;n_y,t} + \boldsymbol{\xi}_{xy;n_y,t} = \boldsymbol{\phi}^{\mathbf{n}}_{y;t} + \boldsymbol{\psi}^{\mathbf{n}}_{y;t} + \boldsymbol{\nu}^{\mathbf{n}}_{y;t} + \boldsymbol{\xi}^{\mathbf{n}}_{xy;t},$$

where

$$\phi^{\mathbf{n}}_{y;it} + \nu^{\mathbf{n}}_{y;it} = \underline{\mathbf{H}}^*_{y;n_z,i}(L)\mathbf{Z}_{n_z,t} = \underline{\mathbf{H}}^*_{y;n_z,i}(L)(\boldsymbol{\phi}_{z;n_z,t} + \boldsymbol{\psi}_{z;n_z,t} + \boldsymbol{\nu}_{z;n_z,t} + \boldsymbol{\xi}_{xz;n_z,t}).$$

Hence, letting $\chi_{yz;it} := (\phi_{y;it} + \nu_{y;it})$ and $\chi_{yz;n_z,it} := \underline{\mathbf{H}}^*_{y;n_z,i}(L)(\boldsymbol{\phi}_{z;n_z,t} + \boldsymbol{\psi}_{z;n_z,t})$, we have

$$\Big[\chi_{yz;it} - \chi_{yz;n_z,it}\Big] + \Big[\psi_{y;it} - \psi^{\mathbf{n}}_{y;it}\Big] + \Big[\xi_{xy;it} - \xi^{\mathbf{n}}_{xy;it}\Big] = \underline{\mathbf{H}}^*_{y;n_z,i}(L)\Big[\boldsymbol{\nu}_{z;n_z,t} + \boldsymbol{\xi}_{xz;n_z,t}\Big]. \qquad (7.25)$$

The outline of the proof is as follows. We first show (Lemma 2) that the spectral density of $\underline{\mathbf{H}}^*_{y;n_z,i}(L)\Big[\boldsymbol{\nu}_{z;n_z,t}+\boldsymbol{\xi}_{xz;n_z,t}\Big] = \underline{\mathbf{H}}^*_{y;n_z,i}(L)\boldsymbol{\xi}_{z;n_z,t}$ tends to zero, θ-a.e. uniformly, as $n_z \to \infty$, which implies that the corresponding process tends to zero in quadratic mean. The same therefore also holds for the left-hand side of (7.25). Denote by $\mathcal{A}_{n_z,i}(\theta)$ the spectral density of that right-hand side, by $\mathcal{B}_{n_z,i}(\theta)$, $\mathcal{C}_{\mathbf{n},i}(\theta)$, and $\mathcal{D}_{\mathbf{n},i}(\theta)$ the spectral densities of $\chi_{yz;it} - \chi_{yz;n_z,it}$, $\psi_{y;it} - \psi^{\mathbf{n}}_{y;it}$, and $\xi_{xy;it} - \xi^{\mathbf{n}}_{xy;it}$, respectively (all these spectral densities are scalar). Noting that $\chi_{yz;it} - \chi_{yz;n_z,it}$ is z-common, whereas $\psi_{y;it}$ and $\xi_{xy;it}$ are z-idiosyncratic, and that $\psi_{y;it}$ is y-common whereas $\xi_{xy;it}$ is y-idiosyncratic, we have that

$$\begin{aligned}\mathcal{A}_{n_z,i}(\theta) \;&=\; \mathcal{B}_{n_z,i}(\theta) + \mathcal{C}_{\mathbf{n},i}(\theta) + \mathcal{D}_{\mathbf{n},i}(\theta)\\ &\quad -2\Re(\mathcal{E}_{\mathbf{n},i}(\theta)) - 2\Re(\mathcal{F}_{\mathbf{n},i}(\theta)) + 2\Re(\mathcal{G}_{\mathbf{n},i}(\theta)) - 2\Re(\mathcal{I}_{\mathbf{n},i}(\theta)) - 2\Re(\mathcal{J}_{\mathbf{n},i}(\theta))\end{aligned}$$

where $\mathcal{E}_{\mathbf{n},i}(\theta)$, $\mathcal{F}_{\mathbf{n},i}(\theta)$, $\mathcal{G}_{\mathbf{n},i}(\theta)$, $\mathcal{I}_{\mathbf{n},i}(\theta)$ and $\mathcal{J}_{\mathbf{n},i}(\theta)$ are the cross-spectra of $\chi_{yz;it} - \chi_{yz;n_z,it}$ and $\psi^{\mathbf{n}}_{y;it}$, $\chi_{yz;it} - \chi_{yz;n_z,it}$ and $\xi^{\mathbf{n}}_{xy;it}$, $\psi^{\mathbf{n}}_{y;it}$ and $\xi^{\mathbf{n}}_{xy;it}$, $\psi_{y;it}$ and $\xi^{\mathbf{n}}_{xy;it}$, and $\psi^{\mathbf{n}}_{y;it}$ and $\xi_{xy;it}$, respectively, and $\Re(z)$ stands for the real part of a complex $z \in \mathbb{C}$. Whe then show (Lemma 3) that those five cross-spectra all pointwise converge to zero, θ-a.e., as $\min(n_y, n_z) \to \infty$. It follows that $\mathcal{B}_{n_z,i}(\theta)$, $\mathcal{C}_{\mathbf{n},i}(\theta)$, and $\mathcal{D}_{\mathbf{n},i}(\theta)$ also pointwise converge to zero θ-a.e. $\min(n_y, n_z) \to \infty$; moreover (Lemma 4), their norms are θ-a.e. bounded. These two facts jointly imply that the corresponding processes tend to zero in quadratic mean, as $\min(n_y, n_z) \to \infty$. This concludes the proof.

LEMMA 2. *For all t, $\underline{\mathbf{H}}^*_{y;n_z,i}(L)\boldsymbol{\xi}_{z;n_z,t}$ tends to zero in quadratic mean, with spectral densities tending to zero pointwise θ-a.e.-uniformly as $n_z \to \infty$.*

PROOF. With the notation of Section 3, the filter $\underline{\mathbf{H}}^*_{y;n_z,i}(L)$ defined in (3.20) can be written as

$$\underline{\mathbf{H}}^*_{y;n_z,i}(L) = \sum_{k=1}^{q_z} \underline{\lambda}^{-1}_{z;n_z,k}(L)\underline{\mathbf{\Sigma}}_{YV;ik}(L)\underline{\mathbf{p}}^*_{z;n_z,k}(L),$$

where $\mathbf{\Sigma}_{YV;ik}(\theta)$ stands for the (scalar) cospectrum of Y_{it} and $V_{z;n_z,kt}$. Then, in view of the mutual orthonormality of eigenvectors,

$$|\mathbf{H}_{y;n_z,i}(\theta)|^2 = \sum_{k=1}^{q_z} |\lambda^{-1}_{z;n_z,k}(\theta)\mathbf{\Sigma}_{YV;ik}(\theta)|^2 = \sum_{k=1}^{q_z} |\lambda^{-1}_{z;n_z,k}(\theta)|^2|\mathbf{\Sigma}_{YV;ik}(\theta)|^2$$

The Cauchy-Schwarz inequality implies that $|\mathbf{\Sigma}_{YV;ik}(\theta)|^2 \leq |\sigma_{ii}(\theta)||\lambda_{z;n_z,k}(\theta)|$. Hence,

$$|\mathbf{H}_{y;n_z,i}(\theta)|^2 \leq c_i \sum_{k=1}^{q_z} |\lambda^{-1}_{z;n_z,k}(\theta)|,$$

a quantity which, in view of Assumption A3(ii), tends to zero as $n_z \to \infty$. The claim then follows from Forni *et al* (2000)'s Lemma 3 and the fact that $\boldsymbol{\xi}_{z;n_z,t}$ is an idiosyncratic process.□

LEMMA 3. *The cross-spectra* $\mathcal{E}_{\mathbf{n},i}(\theta)$ *of* $\chi_{yz;it} - \chi_{yz;n_z,it}$ *and* $\psi^{\mathbf{n}}_{y;it}$, $\mathcal{F}_{\mathbf{n},i}(\theta)$ *of* $\chi_{yz;it} - \chi_{yz;n_z,it}$ *and* $\xi^{\mathbf{n}}_{xy;it}$, $\mathcal{G}_{\mathbf{n},i}(\theta)$ *of* $\psi^{\mathbf{n}}_{y;it}$ *and* $\xi^{\mathbf{n}}_{xy;it}$, $\mathcal{I}_{\mathbf{n},i}(\theta)$ *of* $\psi_{y;it}$ *and* $\xi^{\mathbf{n}}_{xy;it}$, *and* $\mathcal{J}_{\mathbf{n},i}(\theta)$ *of* $\psi^{\mathbf{n}}_{y;it}$ *and* $\xi_{xy;it}$ *tend to zero pointwise,* θ*-a.e., as* $\mathbf{n} \to \infty$.

PROOF. Since $\chi_{yz;it}-\chi_{yz;n_z,it}$ is z-common, it admits a representation of the form $\sum_{j=1}^{q_z} a_{n_z,ij}(L)u_{z;jt}$, where $a_{n_z,ij}(L)$, $j = 1, \dots, q_z$ are square-summable filters and $u_{z;1t}, \dots, u_{z;q_zt}$ are q_z mutually orthogonal white noise processes spanning the z-common space $\mathcal{H}^{\chi}_{z}$ and providing for the Z_{jt}'s a dynamic factor representation of the form (2.4); the existence of such a representation (not its unicity) is guaranteed by Proposition 1.

Contrary to $\chi_{yz;it} - \chi_{yz;n_z,it}$, which belongs to the "exact" z-common space $\mathcal{H}^{\chi}_{z}$, $\psi^{\mathbf{n}}_{y;it}$ is a "reconstructed" quantity, belonging to the orthogonal complement, $\mathcal{H}^{\xi}_{z;n_z}$, say, of the space $\mathcal{H}^{\chi}_{z;n_z}$ spanned by the first q_z dynamic principal components $V^{n_z}_{z;1t}, \dots, V^{n_z}_{z;q_zt}$ of $\boldsymbol{\Sigma}_{z;n_z}(\theta)$ (the spectral densities $\lambda_{z;n_z,1}(\theta), \dots, \lambda_{z;n_z,q_z}(\theta)$ of which diverge). Associated with those dynamic principal components, consider the *normalized* dynamic principal components $W^{n_z}_{z;1t}, \dots, W^{n_z}_{z;q_zt}$, where

$$W^{n_z}_{z;jt} := \lambda^{-1}_{z;n_z,j}(L)V^{n_z}_{z;jt} = \underline{\lambda}^{-1}_{z;n_z,j}(L)\underline{\mathbf{p}}^{*}_{z;n_z,j}(L)\mathbf{Z}_{n_z,t}.$$

For any n_z, the $W^{n_z}_{z;jt}$'s, clearly, are spanning the same *reconstructed* z-common space $\mathcal{H}^{\chi}_{z;n_z}$ as the $V^{n_z}_{z;jt}$'s themselves, but their covariance is a $q_z \times q_z$ unit matrix. The convergence of $\mathcal{H}^{\chi}_{z;n_z}$ to $\mathcal{H}^{\chi}_{z}$ is characterized in the following way (see Lemma 4 of Forni *et al* 2000). Projecting $\mathbf{W}^{n_z}_{z,t} := (W^{n_z}_{z;1t}, \dots, W^{n_z}_{z;q_zt})'$ onto $\mathcal{H}^{\chi}_{z}$ yields

$$\mathbf{W}^{n_z}_{z,t} = \mathbf{A}_{z;n_z}(L)(u_{z;1t}, \dots, u_{z;q_zt})' + \mathbf{R}_{z;n_z,t},$$

where $\mathbf{A}_{z;n_z}(L)$ is an appropriate $n_z \times n_z$ matrix of square-summable filters and the residual $\mathbf{R}_{z;n_z,t}$ is orthogonal to $\mathcal{H}^{\chi}_{z}$. They show that the spectral density matrix of $\mathbf{R}_{z;n_z,t}$ converges to zero θ-a.e., and that $\mathbf{R}_{z;n_z,t}$ itself converges to zero in quadratic mean, as $n_z \to \infty$. Moreover, the projection onto $\mathcal{H}^{\chi}_{z;n_z}$ of $\mathbf{u}_{z,t} := (u_{z;1t}, \dots, u_{z;q_zt})'$ takes the form

$$\mathbf{u}_{z,t} = \mathbf{A}^{*}_{z;n_z}(L^{-1})(W^{n_z}_{z;1t}, \dots, W^{n_z}_{z;q_zt})' + \mathbf{S}_{z;n_z,t}$$

where the spectral density of $\mathbf{S}_{z;n_z,t}$ also converges to zero θ-a.e., and $\mathbf{S}_{z;n_z,t}$ also converges to zero in quadratic mean, as $n_z \to \infty$.

Turning back to the cross-spectrum $\mathcal{E}_{\mathbf{n},i}(\theta)$ of $\chi_{yz;it} - \chi_{yz;n_z,it}$ and $\psi^{\mathbf{n}}_{y;it}$, we thus have

$$\chi_{yz;it} - \chi_{yz;n_z,it} = \mathbf{a}'_i(L)\mathbf{u}_{z,t} = \mathbf{a}'_i(L)\mathbf{A}^{*}_{z;n_z}(L^{-1})\mathbf{W}^{n_z}_{z,t} + \mathbf{a}'_i(L)\mathbf{S}_{z;n_z,t},$$

with $\mathbf{a}'_{n_z,i}(L) := (a_{n_z,i1}(L), \dots, a_{n_z,iq_z}(L))$. Because $\psi^{\mathbf{n}}_{y;it}$ is orthogonal to the space $\mathcal{H}^{\chi}_{z;n_z}$ spanned by $\mathbf{W}^{n_z}_{z,t}$, the cross-spectrum $\mathcal{E}_{\mathbf{n},i}(\theta)$ actually is the cross-spectrum between $\mathbf{a}'_i(L)\mathbf{S}_{z;n_z,t}$ and $\psi^{\mathbf{n}}_{y;it}$. Since $\mathbf{a}'_i(L)\mathbf{S}_{z;n_z,t}$ has spectral density $\mathbf{a}'_i(e^{-\mathrm{i}\theta})\boldsymbol{\Sigma}^{\mathbf{S}}_{z;n_z}(\theta)\mathbf{a}_i(e^{\mathrm{i}\theta})$ tending to zero θ-a.e. as $n_z \to \infty$, and since the spectral density of $\psi^{\mathbf{n}}_{y;it}$ is dominated by that of Y_{it}, the squared modulus of $\mathcal{E}_{\mathbf{n},i}(\theta)$ also tends to zero θ-a.e. as $n_z \to \infty$.

The argument for the cross-spectrum $\mathcal{F}_{\mathbf{n},i}(\theta)$ of $\chi_{yz;it} - \chi_{yz;n_z,it}$ and $\xi^{\mathbf{n}}_{xy;it}$ is entirely similar.

As for the cross-spectrum $\mathcal{G}_{\mathbf{n},i}(\theta)$ of $\psi^{\mathbf{n}}_{y;it}$ and $\xi^{\mathbf{n}}_{xy;it}$, note that, parallel to the decompositions of $\mathcal{H}$ into products of mutually orthogonal subspaces $\mathcal{H} = \mathcal{H}^{\chi} \times \mathcal{H}^{\xi}$, $\mathcal{H} = \mathcal{H}^{\chi}_{z} \times \mathcal{H}^{\xi}_{z}$, $\mathcal{H} = \mathcal{H}^{\chi}_{y} \times \mathcal{H}^{\xi}_{y}$, etc., based on the "exact" common/idiosyncratic components, we have, for each $\mathbf{n}$, decompositions of the form $\mathcal{H} = \mathcal{H}^{\chi}_{\mathbf{n}} \times \mathcal{H}^{\xi}_{\mathbf{n}}$, $\mathcal{H} = \mathcal{H}^{\chi}_{z;n_z} \times \mathcal{H}^{\xi}_{z;n_z}$, $\mathcal{H} = \mathcal{H}^{\chi}_{y;n_y} \times \mathcal{H}^{\xi}_{y;n_y}$, etc., based on the "reconstructed" common/idiosyncratic components. Here, $\xi^{\mathbf{n}}_{xy;it}$ belongs to $\mathcal{H}^{\xi}_{\mathbf{n}}$. On the other hand, $\psi^{\mathbf{n}}_{y;it}$ was defined as

$$\psi^{\mathbf{n}}_{y;it} := \chi^{n_y}_{y;it} - \chi^{n_z}_{yz;it} + \nu^{\mathbf{n}}_{y;it} = \chi^{n_y}_{y;it} - \chi^{n_z}_{yz;it} + \chi^{\mathbf{n}}_{xy;it} - \chi^{n_y}_{y;it} = \chi^{\mathbf{n}}_{xy;it} - \chi^{n_z}_{yz;it}.$$

Since, by construction, $\chi^{\mathbf{n}}_{xy;it} \in \mathcal{H}^{\chi}_{\mathbf{n}}$, and $\chi^{n_z}_{yz;it} \in \mathcal{H}^{\chi}_{z;n_z}$, they both are strictly orthogonal to $\xi^{\mathbf{n}}_{xy;it} \in \mathcal{H}^{\xi}_{\mathbf{n}}$, and the cross-spectrum $\mathcal{G}_{\mathbf{n},i}(\theta)$ is θ-a.e. equal to zero for any $\mathbf{n}$.

The argument for the cross-spectra $\mathcal{I}_{\mathbf{n},i}(\theta)$ of $\psi_{y;it}$ and $\xi^{\mathbf{n}}_{xy;it}$, and the cross-spectra $\mathcal{J}_{\mathbf{n},i}(\theta)$ of $\psi^{\mathbf{n}}_{y;it}$ and $\xi_{xy;it}$ is entirely similar. □

LEMMA 4. *The spectra $\mathcal{B}_{n_z,i}(\theta)$, $\mathcal{C}_{\mathbf{n},i}(\theta)$,and $\mathcal{D}_{\mathbf{n},i}(\theta)$ are θ- a.e. bounded.*

PROOF. We successively consider $\mathcal{B}_{n_z,i}(\theta)$, $\mathcal{C}_{\mathbf{n},i}(\theta)$,and $\mathcal{D}_{\mathbf{n},i}(\theta)$.

(a) The spectral density $\mathcal{B}_{n_z,i}(\theta)$ of $\chi_{yz;it} - \chi_{yz;n_z,it}$ has squared modulus

$$|\mathcal{B}_{n_z,i}(\theta)|^2 = \mathbf{a}'_i(e^{-\mathrm{i}\theta})\mathbf{a}_i(e^{\mathrm{i}\theta}),$$

which is bounded since the $a_{n_z,ij}(L)$'s are square-summable filters.

(b) In order to show that the spectral density $\mathcal{C}_{\mathbf{n},i}(\theta)$ of $\psi_{y;it} - \psi^{\mathbf{n}}_{y;it}$ is θ-a.e. bounded, it is sufficient to show that the spectral densities of $\psi_{y;it}$ and $\psi^{\mathbf{n}}_{y;it}$ are. The spectral density of $\psi_{y;it}$ is dominated by the spectral density of Y_{it} and therefore is θ- a.e. bounded in view of Assumption A2. As for $\psi^{\mathbf{n}}_{y,it}$, we have

$$\begin{aligned}
\psi^{\mathbf{n}}_{y,it} &:= \chi^{n_y}_{y,it} - \phi^{\mathbf{n}}_{y,it} = \chi^{n_y}_{y,it} - \underline{\mathbf{H}}^{*}_{y;n_z,i}(L)\mathbf{Z}_{n_z,t} + \nu^{\mathbf{n}}_{y,it} \\
&= \chi^{n_y}_{y,it} - \underline{\mathbf{H}}^{*}_{y;n_z,i}(L)\mathbf{Z}_{n_z,t} + (\chi^{\mathbf{n}}_{xy;it} - \chi^{n_y}_{y;it}) \\
&= \chi^{\mathbf{n}}_{xy;it} - \underline{\mathbf{H}}^{*}_{y;n_z,i}(L)\mathbf{Z}_{n_z,t} \\
&= \underline{\mathbf{K}}^{*}_{y;\mathbf{n},i}(L)\mathbf{X}_{\mathbf{n},t} - \underline{\mathbf{H}}^{*}_{y;n_z,i}(L)\mathbf{Z}_{n_z,t}.
\end{aligned}$$

The spectral density of $\underline{\mathbf{K}}^{*}_{y;\mathbf{n},i}(L)\mathbf{X}_{\mathbf{n},t}$ is

$$\mathbf{K}^{*}_{y;\mathbf{n},i}(\theta)\Sigma_{\mathbf{n}}(\theta)\mathbf{K}_{y;\mathbf{n},i}(\theta) = \sum_{k=1}^{q} |p_{\mathbf{n},k,i}(\theta)|^2 \lambda_{\mathbf{n},k}(\theta),$$

which is bounded by the spectral density of X_{it} (see Lemma 1 of Forni *et al* 2000). Similarly as in the proof of Lemma 2, the spectral density of $\underline{\mathbf{H}}^{*}_{y;n_z,i}(L)\mathbf{Z}_{n_z,t}$ writes

$$\begin{aligned}
&\sum_{k=1}^{q_z} \left[\mathbf{\Sigma}_{YV;ik}(\theta)\lambda^{-1}_{z;n_z,k}(\theta)\right] \lambda_{z;n_z,k}(\theta) \left[\lambda^{-1}_{z;n_z,k}(\theta)\mathbf{\Sigma}^{*}_{YV;ik}(\theta)\right] \\
&\qquad = \sum_{k=1}^{q_z} |\mathbf{\Sigma}_{YV;ik}(\theta)|^2 \lambda^{-1}_{z;n_z,k}(\theta) \leq q_z\sigma_{ii}(\theta) \leq q_z c_i,
\end{aligned}$$

and therefore is also θ- a.e. bounded; θ- a.e. boundedness of $\mathcal{C}_{\mathbf{n},i}(\theta)$ follows.

(c) Turning to the spectral density $\mathcal{D}_{\mathbf{n},i}(\theta)$ of $\xi_{xy;it} - \xi^{\mathbf{n}}_{xy;it}$, note that the spectral density of $\xi_{xy;it}$, being dominated by that of Y_{it}, is θ- a.e. bounded because of Assumption A2; as for $\xi^{\mathbf{n}}_{xy;it}$, it is of the form $Y_{it} - \chi^{\mathbf{n}}_{xy;it} = Y_{it} - \underline{\mathbf{K}}^{*}_{y;\mathbf{n},i}(L)\mathbf{X}_{\mathbf{n},t}$, where the spectral density of of Y_{it} is θ- a.e. bounded by Assumption A2, while the spectral density of $\underline{\mathbf{K}}^{*}_{y;\mathbf{n},i}(L)\mathbf{X}_{\mathbf{n},t}$ is θ-a.e. bounded because $\underline{\mathbf{K}}_{y;\mathbf{n},i}(L)$ has square-summable coefficients. The claim follows. □

Acknowledgement. The authors thank Marco Lippi and Mario Forni for many stimulating discussions and helpful comments.

References

[1] d'Agostino, A., and D. Giannone (2005), Comparing alternative predictors based on large-panel dynamic factor models, ECB Working Paper #680, Frankfurt, European Central Bank.

[2] Aiolfi, M., L. Catão, and A. Timmerman (2006), Common factors in Latin America's business cycle, Working Paper 06/49, International Monetary Fund.

[3] Altissimo, F., A. Bassanetti, R. Cristadoro, M. Forni, M. Hallin, M. Lippi, and L. Reichlin (2001), A real time coincident indicator of the euro area business cycle. CEPR Discussion Paper 3108, London: Center for Economic Policy Research.

[4] Angelini, E., J. Henry, and R. Mestre (2001). Diffusion index based inflation forecasts for the euro area, ECB Working Paper 61, Frankfurt: European Central Bank.

[5] Artis, M., A. Banerjee, and M. Marcellino (2002). Factor forecasts for the UK, CEPR Discussion Paper 3119, London: Center for Economic Policy Research.

[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. and J. Boivin (2003). Monetary policy in a data rich environment. *Journal of Monetary Economics* 50, 525-546.

[9] Boivin, J. and S. Ng (2006). Are more data always better for factor analysis? *Journal of Econometrics* 132, 169-194.

[10] Brillinger, D.R. (1981). *Time Series : Data Analysis and Theory.* Holden-Day, San Francisco.

[11] Bruneau, C., De Bandt, A., Flageollet, A., and Michaux, E. (2003). Forecasting inflation using economic indicators: the case of France. Working Paper 101, Paris: Banque de France.

[12] Campbell, J.Y., A.W. Lo, and A.C. MacKinlay (1997). *The Econometrics of Financial Markets*, Princeton: Princeton University Press.

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

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

[15] Den Reijer, A.H.J. (2005). Forecasting Dutch GDP using Large Scale Factor Models. DNB Working Paper 28, Amsterdam, De Nederlandsche Bank.

[16] Dreger, C., and Schumacher, C. (2004). Estimating large scale factor models for economic activity in Germany. In A. Wagner, ed., *Jahrbücher für Nationalökonomie und Statistik*, Stuttgart: Lucius & Lucius Verlagsgesellschaft, 731-750.

[17] Engle, R.F. and M.F. Watson (1981). A one-factor multivariate time series model of metropolitan wage rates. *Journal of the American Statistical Association* 76, 774-781.

[18] Favero, C., M. Marcellino, and F. Neglia (2005). Principal components at work: empirical analysis of monetary policy with large data sets. *Journal of Applied Econometrics* 20, 603-620.

[19] Forni, M., D. Giannone, M. Lippi, and L. Reichlin (2005). Opening the black box: identifying shocks and propagation mechanisms in VAR and factor models. IAP Technical Report Series TR # 0482, Université catholique de Louvain.

[20] 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.

[21] Forni, M., Hallin, M., Lippi, M. and L. Reichlin (2003). Do financial variables help forecasting inflation and real activity in the Euro area? *Journal of Monetary Economics* 50, 1243-1255.

[22] Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2004). The generalized dynamic factor model: consistency and rates, *Journal of Econometrics* 119, 231-255.

[23] Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2005). The generalized dynamic factor model: one-sided estimation and forecasting, *Journal of the American Statistical Association* 100, 830-840.

[24] Forni, M., and M. Lippi (2001). The generalized factor model: representation theory, *Econometric Theory* 17, 1113-1141.

[25] Forni, M., and L. Reichlin (1998). Lets get real: a factor analytical approach to disaggregated business cycle dynamics, *Review of Economic Studies* 65, 453-473.

[26] 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*, North-Holland, Amsterdam.

[27] Giannone, D. and T. Matheson (2007). A new core inlaion index for New Zealand. *International Journal of Central Banking* 3, 145-180.

[28] Giannone, D., L. Reichlin, and L. Sala (2005). Monetary policy in real time. In M. Gertler and K. Rogoff, Eds, *NBER Macroeconomic Annual 2004*, Cambridge, M.I.T. Press, 161-200.

[29] Giannone, D., L. Reichlin, and L. Sala (2006). VARs, factor models and the empirical validation of equilibrium business cycle models, *Journal of Econometrics* 132, 257-279.

[30] 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.

[31] Ingersol, J. (1984). Some results in the theory of arbitrage pricing. *The Journal of Finance* 39, 1021-1039.

[32] Lancaster, P. and M. Tismenetsky (1985). *The Theory of Matrices*, 2nd Edition, Orlando: Academic Press.

[33] Marcellino, M., J. Stock, and M. Watson (2003). Macroeconomic forecasting in the euro area: country specific versus area wide information. *European Economic Review* 47, 1-18.

[34] Nieuwenhuyzen, C. (2004). A generalized dynamic factor model for the Belgian economy. *Journal of Business cycle measurement and Analysis* 2, 213-248.

[35] Sargent, T.J. and 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*, Federal Reserve Bank of Minneapolis, Minneapolis.

[36] Schneider, M., and Spitzner, M. (2004). Forecasting Austrian GDP using the generalized dynamic factor model. Working Paper 89, Vienna: Österreichische Nationalbank.

[37] Stock, J. H., and M. W. Watson (1989). New indices of coincident and leading indicators. In O. J. Blanchard and S. Fischer, Eds, *NBER Macroeconomics Annual 1989*, Cambridge: M.I.T. Press.

[38] Stock, J. H. and M. W. Watson (2002a). Macroeconomic forecasting using diffusion indexes, *Journal of Business & Economic Statistics* 20, 147-162.

[39] 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.

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

[41] Yao, T. (2008). Dynamic factors and the source of momentum profits. *Journal of Business & Economic Statistics* 26, 211-226.

NACE code	Description	Germany	France	Italy
E401	Production and distribution of electricity	x	x	x
E402	Manufacture of gas; distribution of gaseous fuels through mains	x	x	x
E403	Steam and hot water supply			x
CA101	Mining and agglomeration of hard coal	x	x	x
CA102	Mining and agglomeration of lignite	x		
CA103	Extraction and agglomeration of peat	x		
CA111	Extraction of crude petroleum and natural gas	x	x	x
CA112	Service activities incidental to oil and gas extraction excluding surveying	x		
CB141	Quarrying of stone	x	x	
CB142	Quarrying of sand and clay	x	x	x
CB143	Mining of chemical and fertilizer minerals	x	x	
CB144	Production of salt	x	x	x
CB145	Other mining and quarrying n.e.c.	x	x	x
DA151	Production, processing, preserving of meat, meat products	x		x
DA152	Processing and preserving of fish and fish products	x	x	x
DA153	Processing and preserving of fruit and vegetables	x	x	x
DA154	Manufacture of vegetable and animal oils and fats	x	x	x
DA155	Manufacture of dairy products	x	x	x
DA156	Manufacture of grain mill products, starches and starch products	x	x	x
DA157	Manufacture of prepared animal feeds	x	x	x
DA158	Manufacture of other food products	x	x	x
DA159	Manufacture of beverages	x	x	x
DB171	Preparation and spinning of textile fibres	x	x	x
DB172	Textile weaving	x	x	x
DB173	Finishing of textiles	x	x	x
DB174	Manufacture of made-up textile articles, except apparel	x	x	x
DB175	Manufacture of other textiles	x	x	x
DB176	Manufacture of knitted and crocheted fabrics	x	x	x
DB177	Manufacture of knitted and crocheted articles	x	x	x
DB181	Manufacture of leather clothes	x	x	
DB182	Manufacture of other wearing apparel and accessories	x	x	x
DB183	Dressing and dyeing of fur; manufacture of articles of fur	x	x	x
DC191	Tanning and dressing of leather	x	x	
DC192	Manufacture of luggage, handbags and the like, saddler	x	x	
DC193	Manufacture of footwear	x	x	x
DD201	Sawmilling and planing of wood, impregnation of wood	x		x
DD202	Manufacture of veneer sheets; manufacture of plywood, laminboard, particle board, fibre board and other panels and boards	x	x	x
DD203	Manufacture of builders' carpentry and joinery	x	x	x
DD204	Manufacture of wooden containers	x	x	x
DD205	Manufacture of other products of wood; manufacture of articles of cork, straw and plaiting materials	x		
DE211	Manufacture of pulp, paper and paperboard	x	x	x
DE212	Manufacture of articles of paper and paperboard	x	x	x
DE221	Publishing	x	x	x
DE222	Printing and service activities related to printing	x	x	
DE223	Reproduction of recorded media	x		
DF231	Manufacture of coke oven products	x	x	x
DF232	Manufacture of refined petroleum products	x	x	x
DF233	Processing of nuclear fuel	x		
DG241	Manufacture of basic chemicals	x	x	x
DG242	Manufacture of pesticides and other agro-chemical products	x	x	x
DG243	Manufacture of paints, varnishes and similar coatings, printing ink and mastics	x	x	x
DG244	Manufacture of pharmaceuticals, medicinal chemicals and botanical products	x	x	x
DG245	Manufacture of soap, detergents, cleaning, polishing	x	x	x
DG246	Manufacture of other chemical products	x	x	x
DG247	Manufacture of man-made fibres	x	x	x
DH251	Manufacture of rubber products	x	x	x
DH252	Manufacture of plastic products	x	x	x
DI261	Manufacture of glass and glass products	x	x	x
DI262	Manufacture of non-refractory ceramic goods other than for construction purposes; manufacture of refractory ceramic products	x	x	x
DI263	Manufacture of ceramic tiles and flags	x	x	x
DI264	Manufacture of bricks, tiles and construction products	x	x	x
DI265	Manufacture of cement, lime and plaster	x	x	x
DI266	Manufacture of articles of concrete, plaster, cement	x		x
DI267	Cutting, shaping and finishing of ornamental and building stone	x	x	
DI268	Manufacture of other non-metallic mineral products	x	x	x
DJ271	Manufacture of basic iron and steel and of ferro-alloys (including production of non-ECSC ferro-alloys in NACE Rev.1.1)	x	x	x
DJ272	Manufacture of tubes	x	x	
DJ273	Other first processing of iron and steel (excluding production of non-ECSC ferro-alloys in NACE Rev.1.1)	x	x	
DJ274	Manufacture of basic precious and non-ferrous metals	x	x	x
DJ275	Casting of metals	x	x	x
DJ281	Manufacture of structural metal products	x	x	x
DJ282	Manufacture of tanks, reservoirs and containers of metal; manufacture of central heating radiators and boilers	x	x	x
DJ283	Manufacture of steam generators, except central heating hot water boilers	x	x	x
DJ284	Forging, pressing, stamping and roll forming of metal; powder metallurgy	x	x	x
DJ285	Treatment and coating of metals; general mechanical engineering	x	x	
DJ286	Manufacture of cutlery, tools and general hardware	x	x	x
DJ287	Manufacture of other fabricated metal products	x	x	x
DK291	Manufacture of machinery for the production and use of mechanical power, except aircraft, vehicle and cycle engines	x	x	x
DK292	Manufacture of other general purpose machinery	x	x	x
DK293	Manufacture of agricultural and forestry machinery	x	x	x
DK294	Manufacture of machine-tools (split into DK2941, DK2942 and DK2943 in NACE Rev.1.1)	x	x	x
DK295	Manufacture of other special purpose machinery	x	x	x
DK296	Manufacture of weapons and ammunition	x	x	x
DK297	Manufacture of domestic appliances n.e.c.	x	x	x
DL311	Manufacture of electric motors, generators and transformers	x	x	x
DL312	Manufacture of electricity distribution and control apparatus	x	x	x
DL313	Manufacture of insulated wire and cable	x	x	x
DL314	Manufacture of accumulators, primary cells and primary batteries	x	x	x
DL315	Manufacture of lighting equipment and electric lamps	x	x	x
DL316	Manufacture of electrical equipment n.e.c.	x	x	x
DL321	Manufacture of electronic valves and tubes and other electronic components	x	x	
DL322	Manufacture of television and radio transmitters and apparatus for line telephony and line telegraphy	x	x	x
DL323	Manufacture of television and radio receivers, sound or video recording or reproducing apparatus and associated goods	x	x	x
DL331	Manufacture of medical and surgical equipment and orthopaedic appliances	x	x	x
DL332	Manufacture of instruments and appliances for measuring, checking, testing, navigating and other purposes, except industrial process control equipment	x	x	x
DL333	Manufacture of industrial process control equipment	x	x	
DL334	Manufacture of optical instruments,photographic equipement	x	x	x
DL335	Manufacture of watches and clocks	x	x	
DM341	Manufacture of motor vehicles	x	x	x
DM342	Manufacture of bodies (coachwork) for motor vehicles; manufacture of trailers and semi-trailers	x	x	x
DM343	Manufacture of parts, accessories for motor vehicles	x	x	x
DM351	Building and repairing of ships and boats	x	x	x
DM352	Manufacture of railway, tramway locomotives, rolling stock	x	x	x
DM353	Manufacture of aircraft and spacecraft	x	x	x
DM354	Manufacture of motorcycles and bicycles	x		x
DM355	Manufacture of other transport equipment n.e.c.	x	x	
DN361	Manufacture of furniture	x	x	x
DN362	Manufacture of jewellery and related articles	x		
DN363	Manufacture of musical instruments	x		x
DN364	Manufacture of sports goods	x		
DN365	Manufacture of games and toys	x		x
DN366	Miscellaneous manufacturing n.e.c.	x		x
DN371	Recycling of metal waste and scrap	x		
DN372	Recycling of non-metal waste and scrap	x		
DD_DE221	Manufacture of wood and paper products, publishing and paper	x		x

Table 2: Data description: the Industrial Production Index based on the 3-digit NACE Rev.1.1 classification, monthly, seasonally adjusted data, from January 1995 through December 2006. Source: Eurostat.