# Is Industrial Production Still the Dominant Factor for the US Economy?*

E. Andreou[†], P. Gagliardini[‡], E. Ghysels[§], M. Rubin [¶]

First version: October, 2014

This version: November 10, 2015

*We thank K. A. Aastveit, G. Barone-Adesi, M. Deistler, M. Del Negro, R. Engle, D. Giannone, T. Götz, C. Hurlin, K. Horl'ovà, A. Levchenko, S. Ng, C. Pérignon, G. Urga, B. Werker for their useful comments as well as participants at the 2013 $(EC)^2$ Conference on "The Econometric Analysis of Mixed Frequency Data" in Nicosia, the 2014 "$8^{th}$ ECB Workshop on Forecasting Techniques" in Frankfurt am Main, the 2014 ESEM in Toulouse, the 2015 "$8^{th}$ Financial Risk International Forum on Scenarios, Stress and Forecast in Finance" in Paris, the 2015 "$2^{nd}$ Workshop on High-Dimensional Time Series in Macroeconomics and Finance" in Vienna, the "Swiss Finance Institute Research Days 2015" in Gerzensee, the 2015 "EABCN/Norges Bank Conference: Econometric methods for business cycle analysis, forecasting and policy simulations" in Oslo, the 2015 "$21^{st}$ International Panel Data Conference" in Budapest, the 2015 "NBER-NSF Time Series Conference" in Vienna, and the 2015 "Econometrics of high-dimensional risk networks" conference at the University of Chicago.

†University of Cyprus and CEPR (elena.andreou@ucy.ac.cy).

‡Università della Svizzera Italiana and Swiss Finance Institute (patrick.gagliardini@usi.ch).

§University of North Carolina - Chapel Hill and CEPR (eghysels@unc.edu).

¶Università della Svizzera Italiana and Swiss Finance Institute (mirco.rubin@usi.ch).

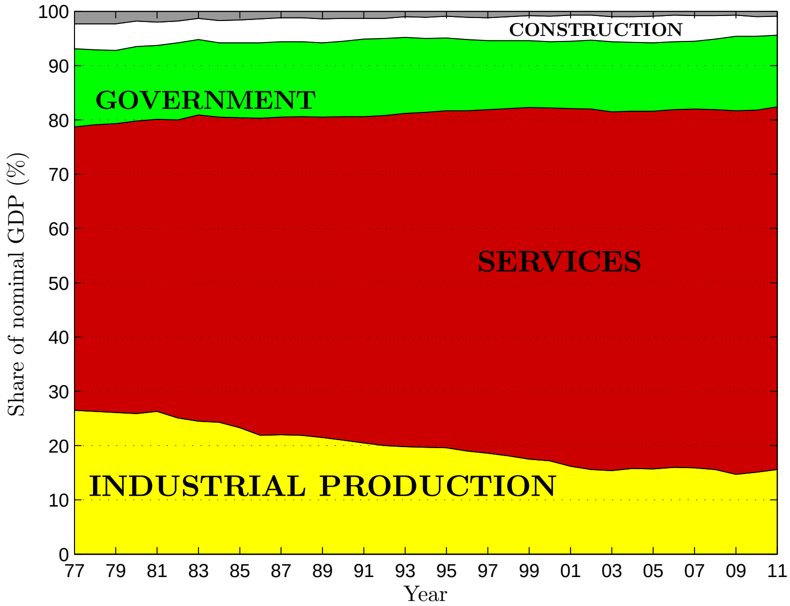
## Abstract

We study the contribution of industrial production (IP) to the decomposition of US output. While the use of factor models has been found convenient, the challenge one faces is that sectoral data beyond IP is only available annually. This imbalance of sampling frequencies poses serious technical problems. We propose a new class of mixed frequency data approximate factor models which enable us to study the full spectrum of quarterly IP sector data combined with the annual non-IP sectors of the economy. We derive the large sample properties of the estimators for the new class of approximate factor models involving mixed frequency data. Using our new approximate factor model, we find that a single common factor explains around 90% of the variability in the aggregate IP output growth index and 60 % of total GDP output growth fluctuations. A single low frequency factor unrelated to manufacturing explains around 14 % of GDP growth fluctuations. The picture with a structural factor model featuring technological innovations is quite different. A factor specific to technological innovations in IP sectors is more important for the IP sector shocks and a low frequency factor which appears to explain variation in information industry as well as professional and business services innovations plays relatively speaking a more important role.

# 1 Introduction

In the public arena it is often claimed that manufacturing has been in decline in the US and most jobs have migrated overseas to lower wage countries. First, we would like to nuance this observation somewhat. It is true, as the figure below clearly shows, that the share of the industrial production sector has been in decline since the late 70's, which is the beginning of our sample period.[1] However, does size matter? The fact that the size shrank does not necessarily exclude the possibility that the industrial production sector still is a key factor, or even the dominant factor, of total US output. We study the validity of this question using novel econometric methods designed to deal with some of the challenging data issues one encounters when trying to address the problem.

Figure 1: Sectoral decomposition of US nominal GDP.




When studying the role of the industrial production sector we face a conundrum. On the one hand, we have fairly extensive data on industrial production (IP) which consists of 117 sectors that make up aggregate IP, each sector roughly corresponding to a four-digit industry classification using NAICS.

[1]The figure displays the evolution from 1977 to 2011 of the sectoral decomposition of US nominal GDP. We aggregate the shares of different sectors available from the website of the US Bureau of Economic Analysis, according to their North American Industry Classification System (NAICS) codes, in 5 different *macro* sectors: Industrial Production (yellow), Services (red), Government (green), Construction (white), Others (grey).

These data are published monthly, and therefore cover a rich time series and cross-section. In our analysis we use the data sampled at quarterly frequency, for reasons explained later in the paper, and consists of over 16,000 data points counting all quarters from 1977 until 2011 (end of our data set) across all sectors. On the other hand, contrary to IP, we do not have monthly or quarterly data about the cross-section of US output across non-IP sectors, but we do so on an annual basis. Indeed, the US Bureau of Economic Analysis provides Gross Domestic Product (GDP) and Gross Output by industry - not only IP sectors - annually. In our empirical analysis we use data on 42 non-IP sectors. If we were to study all sectors annually, we would be left with roughly 4000 data points for IP - a substantial loss of information.

Economists have proposed different models about how various sectors in the economy interact. Some rely on aggregate shocks which affect all sectors at once. Foerster, Sarte, and Watson (2011), who use an approximate factor model estimated with quarterly data, find that nearly all of IP variability is associated with (a small number of) common factors - even a single common factor suffices according to their findings. Does the single common factor which drives IP sectors also affect the rest of the economy, in particular in light of the fact that the services sector grew in relative size? To put it differently, can we maintain a common factor view if we expand beyond IP sectors? Or should we think about sector-specific shocks affecting aggregate US output? If so, are these IP sector shocks, or rather services sector ones?

We propose a new class of factor models able to address these key questions of interest using *all* the data - despite the mixed sampling frequency setting. Empirical research generally avoids the direct use of mixed frequency data by either first aggregating higher frequency series and then performing estimation and testing at the low frequency common across the series, or neglecting the low frequency data and working only on the high frequency series. The literature on large scale factor models is no exception to this practice, see e.g. Forni and Reichlin (1998), Stock and Watson (2002a,b) and Stock and Watson (2010). Using the terminology of the approximate factor model literature, we have a panel consisting of $N_H$ cross-sectional IP sector growth series sampled across $MT$ time periods, where $M$ = 4 for quarterly data and $M$ = 12 for monthly data, with $T$ the number of years. Moreover, we also have a panel of $N_L$ non-IP sectors - such as services and construction for example - which is only observed over $T$ periods. Hence, generically speaking we have a high frequency panel data set of size $N_H \times MT$ and a corresponding low frequency panel data set of size $N_L \times T$. The issue we are interested in can be thought of as follows. There are three types of factors: (1) those which explain variations in both panels - say $g^C$, and therefore are economy-wide factors, (2) those exclusively pertaining to IP sector movements - say $g^H$, and finally (3) those exclusively affecting non-IP, denoted by $g^L$. Hence, we have (1) common, (2) high frequency and (3) low frequency factors. We use superscripts $C$, $H$ and $L$ because the theory we develop is generic and pertains to common (C), high frequency (H) and low frequency (L) factors. The question how to extract common factors from a mixed frequency panel data set is of general interest and has many applications in economics and other fields. In fact our analysis covers an even broader class of group factor models, as will be explained shortly, which is of general interest beyond the mixed frequency setting considered in the empirical application.

The purpose of this paper is to propose large scale approximate factor models in the spirit of Bai and Ng (2002), Stock and Watson (2002a), Bai (2003), Bai and Ng (2006), and extend their analysis to mixed frequency data settings. A number of mixed frequency factor models have been proposed in the literature, although they almost exclusively rely on small cross-sections.[2] Stock and Watson (2002b) in their Appendix A, propose a modification of the EM algorithm of Dempster, Laird, and Rubin (1977) to estimate high frequency factors from potentially large unbalanced panels, with mixed-frequency being a special case.

We approach the problem from a different angle. We start with a setup which identifies factors common to both high and low frequency data panels, the aforementioned $g^C$, and factors specific to the high and low frequency data. Our approach amounts to writing the model as a grouped factor model. The idea to apply grouped factor analysis to mixed frequency data is novel and has many advantages in terms of identification and estimation. In the proposed identification strategy, the groups correspond to panels observed at different sampling frequencies. While there is a literature on how to estimate factors in a grouped model setting, there does not exist a general unifying asymptotic theory for large panel data.[3] We propose estimators for the common and group specific factors, and an inference procedure for the number of common and group specific factors based on canonical correlation analysis of the principal components estimators on each subgroup. One may wonder why we do not apply canonical correlation analysis directly to the high and low frequency data - avoiding the first step of computing

[2]See for example, Mariano and Murasawa (2003), Nunes (2005), Aruoba, Diebold, and Scotti (2009) Frale and Monteforte (2010), Marcellino and Schumacher (2010) and Banbura and Rünstler (2011), among others.

[3]For grouped factor models, see for example Krzanowski (1979), Flury (1984), Kose, Otrok, and Whiteman (2008), Goyal, Pérignon, and Villa (2008), Bekaert, Hodrick, and Zhang (2009), Wang (2012), Hallin and Liska (2011), Moench and Ng (2011), Moench, Ng, and Potter (2013), Ando and Bai (2013) and Breitung and Eickmeier (2014), among others.

principal components since the extra step considerably complicates the asymptotics and actually entails a novel contribution of the paper.[4] What makes the first step of computing principal components necessary is the fact that canonical correlations applied to the raw data may not necessarily uncover pervasive factors.[5] The procedure is therefore general in scope and also of interest in many applications other than the one considered in the current paper.

Our empirical application revisits the analysis of Foerster, Sarte, and Watson (2011) who use factor analytic methods to decompose industrial production (IP) into components arising from aggregate shocks and idiosyncratic sector-specific shocks. They focus exclusively on the industrial production sectors of the US economy. We find that a single common factor explains around 90% of the variability in the aggregate IP output growth index, and a factor specific to IP has very little additional explanatory power. This implies that the single common factor can be interpreted as an Industrial Production factor. Moreover, more than 60% of the variability of GDP output growth in service sectors, such as Transportation and Warehousing services, is also explained by the common factor. A single low frequency factor unrelated to manufacturing, explaining around 14 % of GDP growth fluctuations, drives the comovement of non-IP sectors such as Construction and Government.

We re-examine whether the common factor reflects sectoral shocks that have propagated by way of input-output linkages between service sectors and manufacturing. A structural factor analysis indicates that both low and high frequency aggregate shocks continue to be the dominant source of variation in the US economy. The propagation mechanisms are very different, however, from those identified by Foerster, Sarte, and Watson (2011). Looking at technology shocks instead of output growth, it does not appear that a common factor explaining IP fluctuations is a dominant one for the entire economy. A factor specific to technological innovations in IP sectors is more important for the IP sector shocks and a low frequency factor which appears to explain variation in information industry as well as professional and business services innovations plays relatively speaking a more important role. Hence, when it comes to innovation shocks, IP is no longer the dominant factor.

The rest of the paper is organized as follows. In section 2 we introduce the formal model and

[4]Our work is most closely related to Wang (2012) and Chen (2010, 2012). Yet, there is no comprehensive asymptotic treatment of grouped factor models in a large dimension setting. For example, Wang (2012) proposes an iterative solution from a Least Square (LS) problem. Their procedure is not operational as the resulting equations do not have a unique solution.

[5]A simple example would be to add an anomalous series to one panel and repeat the series to the other one. The canonical correlation analysis applied to the raw data will uncover the presence of the anomalous series in both panels.

discuss identification. In section 3 we study estimation and inference on the number of common factors. The large sample theory appears in section 4. Section 5 covers the empirical application. Section 6 concludes the paper.

Readers who are only interested in the empirical applications can go directly to section 5 which starts with a summary of the novel econometric procedure.

## 2 Model Specification and Identification

We consider a setting where both low and high frequency data are available. Let $t = 1, 2, \ldots, T$ be the low frequency (LF) time units. Each period $(t-1, t]$ is divided into $M$ subperiods with high frequency (HF) dates $t - 1 + m/M$, with $m = 1, \ldots, M$. Moreover, we assume a panel data structure with a cross-section of size $N_H$ of high frequency data and $N_L$ of low frequency data. It will be convenient to use a double time index to differentiate low and high frequency data. Specifically, we let $x^{Hi}_{m,t}$, for $i = 1, \ldots, N_H$, be the high frequency data observation $i$ during subperiod $m$ of low frequency period $t$. Likewise, we let $x^{Li}_t$, with $i = 1, \ldots, N_L$, be the observation of the $i^{th}$ low-frequency series at $t$. These observations are gathered into the $N_H$-dimensional vectors $x^H_{m,t}$, $\forall\, m$, and the $N_L$-dimensional vector $x^L_t$, respectively.

We have a latent factor structure in mind to explain the panel data variation for both the low and high frequency data. To that end, we assume that there are three types of factors, which we denote by respectively $g^C_{m,t}$, $g^H_{m,t}$ and $g^L_{m,t}$. The former represents factors which affect both high and low frequency data (throughout we use superscript $C$ for common), whereas the other two types of factors affect exclusively high (superscript $H$) and low (marked by $L$) frequency data. We denote by $k^C$, $k^H$ and $k^L$, the dimensions of these factors. The latent factor model with high frequency data sampling is:

$$\begin{aligned} x^H_{m,t} &= \Lambda_{HC} g^C_{m,t} + \Lambda_H g^H_{m,t} + e^H_{m,t}, \\ x^{L*}_{m,t} &= \Lambda_{LC} g^C_{m,t} + \Lambda_L g^L_{m,t} + e^L_{m,t}, \end{aligned} \tag{2.1}$$

where $m = 1, ..., M$ and $t = 1, ..., T$, and $\Lambda_{HC}$, $\Lambda_H$, $\Lambda_{LC}$ and $\Lambda_L$ are matrices of factor loadings. The vector $x^{L*}_{m,t}$ is not observable for each high frequency subperiod and the measurements, denoted by $x^L_t$, depend on the observation scheme, which can be either flow sampling or stock sampling (or some general linear scheme). In the remainder of this section we study identification of the model for the case of flow sampling, corresponding to the empirical application covered later in the paper.[6]

In the case of flow sampling, the low frequency observations are the sum (or average) of all $x^{L*}_{m,t}$ in each high frequency subperiod $m$, that is: $x^L_t = \sum_{m=1}^M x^{L*}_{m,t}$. Then, model (2.1) implies:

$$\begin{aligned} x^H_{m,t} &= \Lambda_{HC} g^C_{m,t} + \Lambda_H g^H_{m,t} + e^H_{m,t}, \quad m = 1, ..., M, \\ x^L_t &= \Lambda_{LC} \textstyle\sum_{m=1}^M g^C_{m,t} + \Lambda_L \sum_{m=1}^M g^L_{m,t} + \sum_{m=1}^M e^L_{m,t}. \end{aligned} \tag{2.2}$$

Let us define the aggregated variables and innovations $x^H_t := \sum_{m=1}^M x^H_{m,t}$, $\bar{e}^U_t := \sum_{m=1}^M e^U_{m,t}$, $U = H, L$, and the aggregated factors:

$$\bar{g}^U_t := \sum_{m=1}^M g^U_{m,t}, \qquad U = C, H, L.$$

Then we can stack the observations $x^H_t$ and $x^L_t$ and write:

$$\begin{bmatrix} x^H_t \\ x^L_t \end{bmatrix} = \begin{bmatrix} \Lambda_{HC} & \Lambda_H & 0 \\ \Lambda_{LC} & 0 & \Lambda_L \end{bmatrix} \begin{bmatrix} \bar{g}^C_t \\ \bar{g}^H_t \\ \bar{g}^L_t \end{bmatrix} + \begin{bmatrix} \bar{e}^H_t \\ \bar{e}^L_t \end{bmatrix}. \tag{2.3}$$

The last equation corresponds to a group factor model, with common factor $\bar{g}^C_t$ and "group-specific" factors $\bar{g}^H_t$, $\bar{g}^L_t$.

To further generalize the setup, and draw directly upon the group-factor structure, we will consider the generic specification. To separate the specific from the generic case, we will change notation slightly. Namely, we keep the notation introduced so far with high and low frequency data, temporal aggregation, etc. for the mixed frequency setting further used in the empirical application and use the following for the generic grouped factor model setting:

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} \Lambda^c_1 & \Lambda^s_1 & 0 \\ \Lambda^c_2 & 0 & \Lambda^s_2 \end{bmatrix} \begin{bmatrix} f^c_t \\ f^s_{1,t} \\ f^s_{2,t} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}, \tag{2.4}$$

[6]The identification with stock sampling is discussed in Appendix A.1. It is worth noting though that any sampling scheme leading to a representation of the model analogous to the group-factor model in equation (2.3) or (2.4) - discussed shortly - is compatible with the identification and estimation strategies of this paper.

where $y_{j,t} = [y_{j,1t}, ..., y_{j,N_j t}]'$, $\Lambda_j^c = [\lambda_{j,1}^c, ..., \lambda_{j,N_j}^c]'$, $\Lambda_j^s = [\lambda_{j,1}^s, ..., \lambda_{j,N_j}^s]'$ and $\varepsilon_{j,t} = [\varepsilon_{j,1t}, ..., \varepsilon_{j,N_j t}]'$, with $j = 1, 2$. The dimensions of the common factor $f_t^c$ and the group-specific factors $f_{1,t}^s$, $f_{2,t}^s$ are $k^c$, $k_1^s$ and $k_2^s$, respectively. In the case of no common factors, we set $k^c = 0$, while in the case of no group-specific factors we set $k_j^s = 0$, $j = 1, 2$.[7] The group-specific factors $f_{1,t}^s$ and $f_{2,t}^s$ are orthogonal to the common factor $f_t^c$. Since the unobservable factors can be standardized, we assume:

$$E \begin{bmatrix} f_t^c \\ f_{1,t}^s \\ f_{2,t}^s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \tag{2.5}$$

and

$$V \begin{bmatrix} f_t^c \\ f_{1,t}^s \\ f_{2,t}^s \end{bmatrix} = \begin{bmatrix} I_{k^c} & 0 & 0 \\ 0 & I_{k_1^s} & \Phi \\ 0 & \Phi' & I_{k_2^s} \end{bmatrix}, \tag{2.6}$$

where $\Phi$ is the covariance between the group-specific factors.

## 2.1 Separation of common and group-specific factors

In standard linear latent factor models, the normalization induced by an identity factor variance-covariance matrix identifies the factor process up to a rotation (and change of signs). Let us now show that, under suitable identification conditions, the rotational invariance of model (2.4) - (2.6) allows only for separate rotations among the components of $f_{1,t}^s$, among those of $f_{2,t}^s$, and among those of $f_t^c$. The rotation invariance of model (2.4) - (2.6) therefore maintains the interpretation of common factor and specific factors. More formally, let us consider the following transformation of the stacked factor process:

$$\begin{bmatrix} f_t^c \\ f_{1,t}^s \\ f_{2,t}^s \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \begin{bmatrix} \tilde{f}_t^c \\ \tilde{f}_{1,t}^s \\ \tilde{f}_{2,t}^s \end{bmatrix} \tag{2.7}$$

[7]The case of more than two groups is a relatively straightforward generalization. Note that would also handle situations with more than two sampling frequencies. In the interest of conciseness, we do not consider this type of generalization in the current paper.

where $(\tilde{f}_t^{c\,\prime}, \tilde{f}_{1,t}^{s\,\prime}, \tilde{f}_{2,t}^{s\,\prime})'$ is the transformed stacked factor vector, and the block matrix $A = (A_{ij})$ is nonsingular.

**DEFINITION 1.** *The model is identifiable if: the data $y_{1,t}$ and $y_{2,t}$ satisfy a factor model of the same type as (2.4) - (2.6) with $(f_t^{c\,\prime}, f_{1,t}^{s\,\prime}, f_{2,t}^{s\,\prime})'$ replaced by $(\tilde{f}_t^{c\,\prime}, \tilde{f}_{1,t}^{s\,\prime}, \tilde{f}_{2,t}^{s\,\prime})'$ only when matrix $A$ is a block-diagonal orthogonal matrix.*

The following proposition gives a sufficient condition for the identification of the model with common and group-specific factors.[8]

**PROPOSITION 1.** *If matrices $\Lambda_1 = \begin{bmatrix} \Lambda_1^c \vdots \Lambda_1^s \end{bmatrix}$ and $\Lambda_2 = \begin{bmatrix} \Lambda_2^c \vdots \Lambda_2^s \end{bmatrix}$ are full column-rank (for $N_j$ large enough), then the model is identifiable in the sense of Definition 1.*

**Proof**: See Appendix A.4.1.

Therefore the common factor $f_t^c$ and the group-specific factors $f_{1,t}^s$, $f_{2,t}^s$ and the factor loadings $\Lambda_j^c$, $\Lambda_j^s$, are identifiable up to a linear transformation, since the variables $y_{j,t}$ are observable. By the same token in the mixed frequency setting of equation (2.3), the aggregated factors $\bar{g}_t^C$, $\bar{g}_t^H$, $\bar{g}_t^L$, and the factor loadings $\Lambda_{HC}$, $\Lambda_{LC}$, $\Lambda_H$, $\Lambda_L$, are identified. Once the factor loadings are identified from (2.3), the values of the common and high frequency factors for subperiods $m = 1, ..., M$ are identifiable by cross-sectional regression of the high frequency data on loadings $\Lambda_{HC}$ and $\Lambda_H$ in (2.1). More precisely, $g_{m,t}^C$ and $g_{m,t}^H$ are identified by regressing $x_{m,t}^{Hi}$ on $\lambda_{HC,i}$ and $\lambda_{H,i}$ across $i = 1, 2, ...$, for any $m = 1, ..., M$ and any $t$. Hence, with flow sampling, we can identify the common factor $g_{m,t}^C$ and the high frequency factor $g_{m,t}^H$ at all high frequency subperiods. On the other hand, only $\bar{g}_t^L = \sum_{m=1}^{M} g_{m,t}^L$, i.e. the within-period sum of the low frequency factor, is identifiable by the paired panel data set consisting of $x_t^H$ combined with $x_t^L$. This is not surprising, since we have no HF observation available for the LF process. Note the great advantage of the mixed frequency setting - compared to the single frequency one - in the context of our IP and GDP sector application. The mixed frequency panel setting allows us to identify and estimate the *high frequency* observations of factors common to IP and non-IP sectors. With IP (i.e. high frequency) data only we cannot assess what is common with non-IP. With low frequency data only, we cannot estimate the high frequency common factors.

[8]See also results in e.g. Schott (1999), Wang (2012), Chen (2010, 2012). Proposition 1 is implied by Proposition 1 in Wang (2012).

## 2.2 Identification of the (common) factor space from canonical correlations and directions

In the interest of generality, let us again consider the generic setting of equation (2.4) and let $k_j = k^c + k_j^s$, for $j = 1, 2$, be the dimensions of the factor spaces for the two groups, and define $\underline{k} = \min(k_1, k_2)$. We collect the factors of each group in the $k_j$-dimensional vectors $h_{j,t}$:

$$h_{j,t} := \begin{bmatrix} f_t^c \\ f_{j,t}^s \end{bmatrix}, \qquad j = 1, 2, \qquad t = 1, ..., T, \tag{2.8}$$

and the loadings in the $k_j$-dimensional vectors $\lambda_{j,i}$:

$$\lambda_{j,i} := \begin{bmatrix} \lambda_{j,i}^c \\ \lambda_{j,i}^s \end{bmatrix}, \qquad j = 1, 2, \qquad i = 1, ..., N_j.$$

Using these definitions, model (2.4) can equivalently be written as:

$$y_{j,it} \;=\; \lambda_{j,i}' h_{j,t} + \varepsilon_{j,it}, \qquad j = 1, 2, \qquad i = 1, ..., N_j, \qquad t = 1, ..., T,$$

We also stack the factors $h_{j,t}$, $j = 1, 2$, into the $K$-dimensional vector $h_t = (h_{1,t}', h_{2,t}')'$, with $K = k_1 + k_2$. Moreover, let us express the $(K, K)$-dimensional matrix $V(h_t)$ as:

$$V(h_t) \;=\; \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix}, \tag{2.9}$$

where:

$$V_{j\ell} \;:=\; E(h_{j,t} h_{\ell,t}'), \qquad j, \ell = 1, 2. \tag{2.10}$$

Let us first recall a few basic results from canonical analysis (see e.g. Anderson (2003) and Magnus and Neudecker (2007)). Let $\rho_\ell$, $\ell = 1, ..., \underline{k}$ denote the canonical correlations between $h_{1,t}$ and $h_{2,t}$.

The largest $\underline{k}$ eigenvalues of matrices

$$R = V_{11}^{-1} V_{12} V_{22}^{-1} V_{21}, \qquad \text{and} \quad R^* = V_{22}^{-1} V_{21} V_{11}^{-1} V_{12},$$

are the same, and are equal to the squared canonical correlations $\rho_\ell^2$, $\ell = 1, ..., \underline{k}$ between $h_{1,t}$ and $h_{2,t}$. The associated eigenvectors $w_{1,\ell}$ (resp. $w_{2,\ell}$), with $\ell = 1, ..., \underline{k}$, of matrix $R$ (resp. $R^*$) standardized such that $w_{1,\ell}' V_{11} w_{1,\ell} = 1$ (resp. $w_{2,\ell}' V_{22} w_{2,\ell} = 1$) are the canonical directions which allow to construct the canonical variables from vector $h_{1,t}$ (resp. $h_{2,t}$). The matrices $w_j = (w_{j,1}, ..., w_{j,\underline{k}})$, $j = 1, 2$, are such that $w_j' V_{jj} w_j = I_{\underline{k}}$, $j = 1, 2$. Moreover, if $\rho_\ell \neq 0$, then

$$\begin{aligned} w_{1,\ell} &= \frac{1}{\rho_\ell} V_{11}^{-1} V_{12} w_{2,\ell}, \\ w_{2,\ell} &= \frac{1}{\rho_\ell} V_{22}^{-1} V_{21} w_{1,\ell}. \end{aligned} \tag{2.11}$$

**PROPOSITION 2.** *The following hold:*

*i) If $k^c > 0$, the largest $k^c$ canonical correlations between $h_{1,t}$ and $h_{2,t}$ are equal to $1$, and the remaining $\underline{k} - k^c$ canonical correlations are strictly smaller than $1$.*

*ii) Let $W_j$ be the $(k_j, k^c)$ matrix whose columns are the canonical directions for $h_{j,t}$ associated with the $k^c$ canonical correlations equal to $1$, with $j = 1, 2$. Then, we have $f_t^c = W_j' h_{j,t}$ (up to a rotation matrix), for $j = 1, 2$.*

*iii) If $k^c = 0$, all canonical correlations between $h_{1,t}$ and $h_{2,t}$ are strictly smaller than $1$.*

*iv) Let $W_1^s$ (resp. $W_2^s$) be the $(k_1, k_1^s)$ (resp. $(k_2, k_2^s)$) matrix whose columns are the eigenvectors of matrix $R$ (resp. $R^*$) associated with the smallest $k_1^s$ (resp. $k_2^s$) eigenvalues. Then $f_{j,t}^s = W_j^{s\prime} h_{j,t}$ (up to a rotation matrix) for $j = 1, 2$.*

**Proof**: See Appendix A.4.2.

Proposition 2 shows that the number of common factors $k^c$, the common factor space spanned by $f_t^c$, and the spaces spanned by group specific factors, can be identified from the canonical correlations and canonical variables of $h_{1,t}$ and $h_{2,t}$. Therefore, the dimension $k^c$, and factors $f_t^c$ and $f_{j,t}^s$, $j = 1, 2$, (up to a rotation) are identifiable from information that can be inferred by disjoint principal component analysis (PCA) on the two subgroups. Note that disjoint PCA on the two subgroups allows us to identify $h_{1,t}$ and $h_{2,t}$ up to linear transformations. This fact does not prevent identifiability of the common and group-specific factors from Proposition 2. More precisely, from the subpanel $j$ we can identify the vector $h_{j,t}$ up to a non-singular matrix $U_j$, say, $j = 1, 2$. Under the transformation $h_{j,t} \to U_j h_{j,t}$, the matrices $R$ and $R^*$ are transformed such that $R \to (U_1')^{-1} R U_1'$ and $R^* \to (U_2')^{-1} R^* U_2'$. Therefore, the matrices of canonical directions $W_1$ and $W_2$ are transformed such as $W_j \to (U_j')^{-1} W_j$, $j = 1, 2$. Therefore, the quantities $W_j' h_{j,t}$, $j = 1, 2$, are invariant under such transformations.

Last, but certainly not least, we provide in the Online Appendix to the paper an alternative way for the identification of the common factor space from variance-covariance matrix of stacked factors (see Section OA.1).

# 3 Estimation and inference on the number of common factors

In Section 3.1 we provide estimators of the common and group-specific factors, based on canonical correlations and canonical directions, when the true number of group-specific and common factors are known. In Section 3.2 we propose a sequential testing procedure for determining the number of common factors when only the dimensions $k_1$ and $k_2$ are known. The test statistic is based on the canonical correlations between the estimated factors in each subgroup of observables. In Section 3.3 we explain why the asymptotic results concerning the test statistic and the factors estimators obtained under the assumption that the number of pervasive factors $k_1$ and $k_2$ in each group is known, remain unchanged when the number of pervasive factors is consistently estimated. Finally, in Section 3.4 we use these results to define estimators and test statistics for the mixed frequency factor model.

## 3.1 Estimation of common and group-specific factors when the number of common and group-specific factors is known

Let us assume that the true number of factors $k_j > 0$ in each subgroup, $j = 1, 2$ is known, and also that the true number of common factors $k^c > 0$, is known. Proposition 2 suggests the following estimation procedure for the common factor. Let $h_{1,t}$ and $h_{2,t}$ be estimated (up to a rotation) by extracting the first $k_j$ Principal Components (PCs) from each subpanel $j$, and denote by $\hat{h}_{j,t}$ these PC estimates of the factors, $j = 1, 2$. Let $\hat{H}_j = [\hat{h}_{j,1}, ..., \hat{h}_{j,T}]'$ be the $(T, k_j)$ matrix of estimated PCs extracted from panel $Y_j = [y_{j,1}, ..., y_{j,T}]'$ associated with the largest $k_j$ eigenvalues of matrix $\frac{1}{N_j T} Y_j Y_j'$, $j = 1, 2$. Let

$\hat{V}_{j\ell}$ denote the empirical covariance matrix of the estimated vectors $\hat{h}_{j,t}$ and $\hat{h}_{\ell,t}$, with $j, \ell = 1, 2$:

$$\hat{V}_{j\ell} = \frac{\hat{H}_j' \hat{H}_\ell}{T} = \frac{1}{T} \sum_{t=1}^{T} \hat{h}_{j,t} \hat{h}_{\ell,t}', \qquad j, \ell = 1, 2, \tag{3.1}$$

and let matrices $\hat{R}$ and $\hat{R}^*$ be defined as:

$$\hat{R} := \hat{V}_{11}^{-1} \hat{V}_{12} \hat{V}_{22}^{-1} \hat{V}_{21}, \text{ and } \hat{R}^* := \hat{V}_{22}^{-1} \hat{V}_{21} \hat{V}_{11}^{-1} \hat{V}_{12}. \tag{3.2}$$

Matrices $\hat{R}$ and $\hat{R}^*$ have the same non-zero eigenvalues. From Anderson (2003) and Magnus and Neudecker (2007), we know that the largest $k^c$ eigenvalues of $\hat{R}$ (resp. $\hat{R}^*$), denoted by $\hat{\rho}_\ell^2$, $\ell = 1, ..., k^c$, are the first $k^c$ squared sample canonical correlation between $\hat{h}_{1,t}$ and $\hat{h}_{2,t}$. We also know that the associated $k^c$ canonical directions, collected in the $(k_1, k^c)$ (resp. $(k_2, k^c)$) matrix $\hat{W}_1$ (resp. $\hat{W}_2$), are the eigenvectors associated with the largest $k^c$ eigenvalues of matrix $\hat{R}$ (resp. $\hat{R}^*$), normalized to have length 1 w.r.t. matrix $\hat{V}_{11}$ (resp. $\hat{V}_{22}$). It also holds:

$$\hat{W}_1' \hat{V}_{11} \hat{W}_1 = I_{k^c}, \text{ and } \hat{W}_2' \hat{V}_{22} \hat{W}_2 = I_{k^c}.$$

**DEFINITION 2.** *Two estimators of the common factors vector are* $\hat{f}_t^c = \hat{W}_1' \hat{h}_{1,t}$ *and* $\hat{f}_t^{c*} = \hat{W}_2' \hat{h}_{2,t}$.

Let matrix $\hat{W}_1^s$ (resp. $\hat{W}_2^s$) be the $(k_1, k_1^s)$ (resp. $(k_2, k_2^s)$) matrix collecting $k_1^s$ (resp. $k_2^s$) eigenvectors associated with the $k_1^s$ (resp. $k_2^s$) smallest eigenvalues of matrix $\hat{R}$ (resp. $\hat{R}^*$), normalized to have length 1 w.r.t. matrix $\hat{V}_{11}$ (resp. $\hat{V}_{22}$). It also holds:

$$\hat{W}_1^{s\,\prime} \hat{V}_{11} \hat{W}_1^s = I_{k_1^s}, \text{ and } \hat{W}_2^{s\,\prime} \hat{V}_{22} \hat{W}_2^s = I_{k_2^s}.$$

The estimators of the group-specific factors can be defined analogously to the definition of the common factors.

**DEFINITION 3.** *Two estimators of the specific factors vector are* $\breve{f}_{1,t}^s = \hat{W}_1^{s\,\prime} \hat{h}_{1,t}$ *and* $\breve{f}_{2,t}^s = \hat{W}_2^{s\,\prime} \hat{h}_{2,t}$.

Let $\hat{F}^c = [\hat{f}_1^{c\,\prime}, ..., \hat{f}_T^{c\,\prime}]'$ and $\hat{F}^{c*} = [\hat{f}_1^{c\,*\,\prime}, ..., \hat{f}_T^{c\,*\,\prime}]'$ be the $(T, k^c)$ matrices of estimated common factors, and $\breve{F}_j^s = [\breve{f}_{j,1}^{s\,\prime}, ..., \breve{f}_{j,T}^{s\,\prime}]'$ be the $(T, k_j^s)$, for $j = 1, 2$, be the matrices of estimated group-specific factors. Then, $\hat{F}^c$ (resp. $\hat{F}^{c*}$) and $\breve{F}_1^s$ (resp. $\breve{F}_2^s$) are orthogonal in sample.

An alternative estimator for the group-specific factors $f^s_{1,t}$ (resp. $f^s_{2,t}$) is obtained by computing the first $k^s_1$ (resp. $k^s_2$) principal components of the variance-covariance matrix of the residuals of the regression of $y_{1,t}$ (resp. $y_{2,t}$) on the estimated common factors.[9] More specifically, let $\hat{\Lambda}^c_j = [\hat{\lambda}^c_{j,1}, ..., \hat{\lambda}^c_{j,N_j}]'$ be the $(N_j, k^c)$ matrix collecting the loadings estimators:

$$\hat{\Lambda}^c_j \;=\; Y_j' \hat{F}^c (\hat{F}^{c\,\prime} \hat{F}^c)^{-1}, \qquad j = 1, 2. \tag{3.3}$$

Let $\xi_{j,it} = y_{j,it} - \hat{\lambda}^{c\,\prime}_{j,i} \hat{f}^c_t$ be the residuals of the regression of $y_{j,t}$ on the estimated common factor $\hat{f}^c_t$, and let $\xi_{j,t} = [\xi_{j,1t}, ..., \xi_{j,N_j t}]'$, for $j = 1, 2$. Let $\Xi_j = [\xi_{j,1}, ..., \xi_{j,T}]'$ be the $(T, N_j)$ matrix of the regression residuals, for $j = 1, 2$.

**DEFINITION 4.** *An alternative estimator of the specific factor vector is $\hat{f}^s_{1,t}$ (resp. $\hat{f}^s_{2,t}$), defined as the first $k^s_1$ (resp. $k^s_2$) Principal Components of subpanel $\Xi_1$ (resp. $\Xi_2$).*

We denote by $\hat{F}^s_j = [\hat{f}^{s\,\prime}_{j,1}, ..., \hat{f}^{s\,\prime}_{j,T}]'$ the $(T, k^s_j)$ matrix of estimated group-specific factors, corresponding to the PCs extracted from panel $\Xi_j$ associated with the largest $k^s_j$ eigenvalues of matrix $\frac{1}{N_j T} \Xi_j \Xi_j'$, for $j = 1, 2$. Then, $\hat{F}^c$ is orthogonal in sample both to $\hat{F}^s_1$ and to $\hat{F}^s_2$. Moreover, we define $\hat{\Lambda}^s_j = [\hat{\lambda}^s_{j,1}, ..., \hat{\lambda}^s_{j,N_j}]'$ as the $(N_j, k^s_j)$ matrix collecting the loadings estimators:

$$\hat{\Lambda}^s_j \;=\; Y_j' \hat{F}^s_j (\hat{F}^{s\,\prime}_j \hat{F}^s_j)^{-1} \;=\; \Xi_j' \hat{F}^s_j (\hat{F}^{s\,\prime}_j \hat{F}^s_j)^{-1}, \qquad j = 1, 2, \tag{3.4}$$

where the second equality follows from the in-sample orthogonality of $\hat{F}^c$ and $\hat{F}^s_j$, for $j = 1, 2$.

## 3.2 Inference on the number of common factors based on canonical correlations

Suppose that the number of factors $k_1$ and $k_2$ in each subpanel is known, and hence $\underline{k} = \min(k_1, k_2)$ as well, and let us consider the problem of inferring the dimension $k^c$ of the common factor space. From Proposition 2, this dimension is the number of unit canonical correlations between $h_{1,t}$ and $h_{2,t}$. We

[9]This alternative estimation method for the group-specific factors corresponds to the method proposed by Chen (2012).

consider the following set of hypotheses:

$$\begin{aligned} H(0) &= \{1 > \rho_1 \geq ... \geq \rho_{\underline{k}}\}, \\ H(1) &= \{\rho_1 = 1 > \rho_2 \geq ... \geq \rho_{\underline{k}}\} \\ &... \\ H(k^c) &= \{\rho_1 = ... = \rho_{k^c} = 1 > \rho_{k^c+1} \geq ... \geq \rho_{\underline{k}}\}, \\ &... \\ H(\underline{k}) &= \{\rho_1 = ... = \rho_{\underline{k}} = 1\}, \end{aligned}$$

where $\rho_1, ..., \rho_{\underline{k}}$ are the canonical correlations of $h_{1,t}$ and $h_{2,t}$. Hypothesis $H(0)$ corresponds to the case of no common factor in the two groups of observables $Y_1$ and $Y_2$. Generically, $H(k^c)$ corresponds to the case of $k^c$ common factor and $k_1 - k^c$ and $k_2 - k^c$ group-specific factors in each group. The largest possible number of common factors is the minimum between $k_1$ and $k_2$, i.e. $\underline{k}$, and corresponds to hypothesis $H(\underline{k})$. In order to select the number of common factors, let us consider the following sequence of tests:

$$H_0 = H(k^c) \qquad \text{against} \qquad H_1 = \bigcup_{0 \leq r < k^c} H(r),$$

for each $k^c = \underline{k}, \underline{k} - 1, ..., 1$. We propose the following statistic to test $H_0$ against $H_1$, for any given $k^c = \underline{k},\ \ \underline{k} - 1,\ \ ...,\ \ 1$:

$$\hat{\xi}(k^c) = \sum_{\ell=1}^{k^c} \hat{\rho}_\ell. \tag{3.5}$$

The statistic $\hat{\xi}(k^c)$ corresponds to the sum of the $k^c$ largest sample canonical correlations. We reject the null $H_0 = H(k^c)$ when $\hat{\xi}(k^c) - k^c$ is negative and large. The critical value is deduced by the large sample distribution provided in Section 4.

### 3.3 Inference on the number of common factors when $k_1$ and $k_2$ are unknown

The tests defined in Section 3.2 require the knowledge of the true number of pervasive factors $k_j > 0$ in each subgroup, $j = 1, 2$. When the true number of pervasive factors is not known, but consistent estimators $\hat{k}_1$ and $\hat{k}_2$, say, are available, the asymptotic distributions and rates of convergence for the test statistic $\hat{\xi}(k^c)$ based on $\hat{k}_1$ and $\hat{k}_2$ are the same as those of the test based on the true number of factors. Intuitively, this holds because the consistency of estimators $\hat{k}_j$, implies that $P(\hat{k}_j = k_j) \to 1$ for $j = 1, 2$, which means that the error due to the estimation of the number of pervasive factors is (asymptotically) negligible.[10]

The estimators based on the penalized information criteria of Bai and Ng (2002) applied on the two subgroups, are examples of consistent estimators for the numbers of pervasive factors. Therefore, in the next Section 4, the asymptotic distributions and rates of convergence of the test statistic and factors estimators are derived assuming that the true numbers of factors $k_j > 0$ in each subgroup, $j = 1, 2$, are known.

### 3.4 Estimation and inference in the mixed frequency factor model

The estimators and test statistics defined in Sections 3.1 - 3.3 for the group factor model (2.4) allow to define estimators for the loadings matrices $\Lambda_{HC}$, $\Lambda_H$, $\Lambda_{LC}$, $\Lambda_L$, the aggregated factor values $\overline{g}_t^U$, $U = C, H, L$ and the test statistic for the common factor space dimension $k^C$ in equation (2.3). We denote these estimators $\hat{\Lambda}_{HC}$, $\hat{\Lambda}_H$, $\hat{\Lambda}_{LC}$, $\hat{\Lambda}_L$, $\hat{\overline{g}}_t^U$, and the test statistic $\hat{\xi}(k^C)$. The estimators of the common and high frequency factor values are:

$$\begin{pmatrix} \hat{g}_{m,t}^C \\ \hat{g}_{m,t}^H \end{pmatrix} = \left(\hat{\Lambda}_1'\hat{\Lambda}_1\right)^{-1} \hat{\Lambda}_1' x_{m,t}^H, \qquad m = 1, ..., M, \qquad t = 1, ..., T, \tag{3.6}$$

where $\hat{\Lambda}_1 = [\hat{\Lambda}_{HC} \vdots \hat{\Lambda}_H]$.

[10]This argument is formalized using similar arguments as, for instance, in footnote 5 of Bai (2003).

# 4 Large sample theory

In this section we derive the large sample distributions of the estimators of factor spaces and factor loadings, and of the test statistic for the dimension of the common factor space. We consider the joint asymptotics $N_1, N_2, T \to \infty$ under Assumptions A.1-A.8 provided in Appendices A.2 and A.3. From the asymptotic theory of principal component analysis (PCA) estimators in large panels (see e.g. Bai and Ng (2002), Stock and Watson (2002a), Bai (2003), Bai and Ng (2006)) we know that:

$$\hat{h}_{j,t} \simeq \hat{\mathcal{H}}_j \left( h_{j,t} + \frac{1}{\sqrt{N_j}} u_{j,t} + \frac{1}{T} b_{j,t} \right), \quad j = 1, 2, \tag{4.1}$$

where $b_{j,t}$ is a deterministic bias term, the matrix $\hat{\mathcal{H}}_j$ converges to a non-singular matrix as $N_j, T \to \infty$, and:

$$\begin{aligned} u_{j,t} &:= \left( \frac{1}{N_j} \sum_{i=1}^{N_j} \lambda_{j,i} \lambda'_{j,i} \right)^{-1} \frac{1}{\sqrt{N_j}} \sum_{i=1}^{N_j} \lambda_{j,i} \varepsilon_{j,it} \\ &= \left( \frac{\Lambda'_j \Lambda_j}{N_j} \right)^{-1} \frac{1}{\sqrt{N_j}} \Lambda'_j \varepsilon_{j,t}. \end{aligned} \tag{4.2}$$

Note that the terms $u_{j,t}$ depend also from the cross-sectional dimension $N_j$, but for notational convenience, we omit the index $N_j$ in $u_{j,t}$. From Assumptions A.2 and A.5 *d)* the error terms $u_{j,t}$ are asymptotically Gaussian as $N_j \to \infty$:

$$u_{j,t} \xrightarrow{d} N(0, \Sigma_{u,j}), \tag{4.3}$$

where the asymptotic variance is:

$$\Sigma_{u,j} = \Sigma_{\Lambda,j}^{-1} \Omega_j \Sigma_{\Lambda,j}^{-1}, \tag{4.4}$$

and

$$\Sigma_{\Lambda,j} = \lim_{N_j \to \infty} \frac{1}{N_j} \sum_{i=1}^{N_j} \lambda_{j,i} \lambda_{j,i}', \tag{4.5}$$

$$\Omega_j = \lim_{N_j \to \infty} \frac{1}{N_j} \sum_{i=1}^{N_j} \sum_{\ell=1}^{N_j} \lambda_{j,i} \lambda_{j,\ell}' Cov(\varepsilon_{j,i,t}, \varepsilon_{j,\ell,t}), \quad j = 1, 2. \tag{4.6}$$

Without loss of generality, let $N_2 \leq N_1$. We assume $\sqrt{N_1}/T = o(1)$ (Assumption A.6), which allows to neglect the bias terms $b_{j,t}/T$ in the asymptotic expansion (4.1). We also assume $T/N_2 = o(1)$, which further simplifies the asymptotic distributions derived in the next section.

## 4.1 Main asymptotic results for the group factor model

In this section we collect the main results concerning the asymptotic distributions of estimators and test statistics for the group factor model. Define the matrices:

$$\Omega_{j,k}(h) = \lim_{N_j,\, N_k \to \infty} \frac{1}{\sqrt{N_j N_k}} \sum_{i=1}^{N_j} \sum_{\ell=1}^{N_k} \lambda_{j,i} \lambda_{k,\ell}' Cov(\varepsilon_{j,i,t}, \varepsilon_{k,\ell,t-h}), \tag{4.7}$$

$$\Sigma_{u,jk}(h) = \Sigma_{\Lambda,j}^{-1} \Omega_{jk}(h) \Sigma_{\Lambda,k}^{-1}, \tag{4.8}$$

for $j, k = 1, 2$, and $h = ..., -1, 0, 1, ...$ Matrix $\Sigma_{u,jk}(h)$ is the asymptotic covariance between $u_{j,t}$ and $u_{k,t-h}$. Moreover, we have $\Omega_j \equiv \Omega_{j,j}(0)$ and $\Sigma_{u,j} \equiv \Sigma_{u,jj}(0)$, and similarly we define $\Sigma_{u,12} \equiv \Sigma_{u,12}(0) = \Sigma_{u,21}'$. Let us denote $N = \min\{N_1, N_2\} = N_2$ the minimal cross-sectional dimension among the two groups, and $\mu_N^2 = N_2/N_1 \leq 1$. Let $\mu_N \to \mu$, with $\mu \in [0, 1]$. The boundary value $\mu = 0$ accounts for the possibility that $N_1$ grows faster than $N_2$.

**THEOREM 3.** *Under Assumptions A.1 - A.6, and the null hypothesis $H_0 = H(k^c)$ of $k^c$ common factors, we have:*

$$N\sqrt{T} \left[ \hat{\xi}(k^c) - k^c + \frac{1}{2N} tr \left\{ \tilde{\Sigma}_{cc}^{-1} \Sigma_{U,N} \right\} \right] \xrightarrow{d} N\left(0, \frac{1}{4}\Omega_U\right), \tag{4.9}$$

*where*

$$\tilde{\Sigma}_{cc} = \frac{1}{T}\sum_{t=1}^{T} f_t^c f_t^{c\prime}, \tag{4.10}$$

$$\Omega_U = 2\sum_{h=-\infty}^{\infty} tr\left\{\Sigma_U(h)\Sigma_U(h)'\right\}, \tag{4.11}$$

$$\Sigma_U(h) = \mu^2\Sigma_{u,11}^{(cc)}(h) + \Sigma_{u,22}^{(cc)}(h) - \mu\Sigma_{u,12}^{(cc)}(h) - \mu\Sigma_{u,21}^{(cc)}(h), \tag{4.12}$$

$$\Sigma_{U,N} = \mu_N^2\Sigma_{u,1}^{(cc)} + \Sigma_{u,2}^{(cc)} - \mu_N\Sigma_{u,12}^{(cc)} - \mu_N\Sigma_{u,21}^{(cc)}, \tag{4.13}$$

*and the upper index $(c, c)$ denotes the upper-left $(k^c, k^c)$ block of a matrix.*

**Proof**: See Appendix A.5.

The asymptotic distribution of $\hat{\xi}(k^c) - k^c$ after appropriate recentering and rescaling is Gaussian. The convergence rate is $N\sqrt{T}$. The asymptotic expansion of $\hat{\xi}(k^c) - k^c$ involves a time series average of squared estimation errors on group factors. Since these estimation errors are of order $1/\sqrt{N}$, the expected value of their square will be of order $1/N$, originating a recentering term of the second order analogous to an error-in-variable bias adjustment. Moreover, the averaging over time of the recentered squared estimation errors allows to apply a root-$T$ central limit theorem for weakly dependent processes, originating a total estimation uncertainty for the test statistic of order $1/(N\sqrt{T})$.

**THEOREM 4.** *Under Assumptions A.1 - A.6 we have:*

$$\sqrt{N_1}(\hat{\mathcal{H}}_c\hat{f}_t^c - f_t^c) \xrightarrow{d} N\left(0, \Sigma_{u,1}^{(cc)}\right), \tag{4.14}$$

$$\sqrt{N_2}(\hat{\mathcal{H}}_c^*\hat{f}_t^{c\,*} - f_t^c) \xrightarrow{d} N\left(0, \Sigma_{u,2}^{(cc)}\right), \tag{4.15}$$

$$\sqrt{N_j}\left[\hat{\mathcal{H}}_{s,j}\hat{f}_{j,t}^s - \left(f_{j,t}^s - (F_j^{s\,\prime}F^c)(F^{c\,\prime}F^c)^{-1}f_t^c\right)\right] \xrightarrow{d} N\left(0, (\Sigma_{\Lambda,j}^{(ss)})^{-1}\Omega_j^{(ss)}(\Sigma_{\Lambda,j}^{(ss)})^{-1}\right), \tag{4.16}$$

*for any $j$, $t$, where $\hat{\mathcal{H}}_c$, $\hat{\mathcal{H}}_c^*$ and $\hat{\mathcal{H}}_{s,j}$ are non-singular matrices, $F^c = [f_1^c, ..., f_T^c]'$, $F_j^s = [f_{j,1}^s, ..., f_{j,T}^s]'$ and the upper index $(ss)$ denotes the lower-right $(k_j^s, k_j^s)$ block of a matrix.*

**Proof**: See Appendix A.6.

From Theorem 4 a linear transformation of vector $\hat{f}_t^c$ (resp. $\hat{f}_t^{c*}$) estimates the common factor $f_t^c$ at a rate $1/\sqrt{N_1}$ (resp. $1/\sqrt{N_2}$). The variance of the asymptotic Gaussian distribution is the upper-left $(c, c)$ block of matrix $\Sigma_{u,1}$ (resp. $\Sigma_{u,2}$), i.e. the asymptotic variance of the estimation error $u_{1,t}$

(resp. $u_{2,t}$) for the PC vector in group 1 (resp. group 2). The estimation error for recovering the common factors from the group PC's is of order $1/\sqrt{NT}$, and therefore asymptotically negligible. The estimator $\hat{f}^s_{j,t}$ approximates the residual of the sample projection of the group-$j$ specific factor on the common factor, up to a linear transformation, at rate $1/\sqrt{N_j}$.

Let us now derive the asymptotic distribution of the factor loadings estimators.[11] Define the matrices:

$$\Phi_{j,i} = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \sum_{r=1}^{T} E[f_{j,t} f'_{j,r}] cov(\varepsilon_{j,i,t},\ \varepsilon_{j,i,r}), \tag{4.17}$$

$$\Psi_j = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \sum_{r=1}^{T} E\left[f^s_{j,t} f^{s\prime}_{j,r} \otimes f^c_t f^{c\prime}_r\right]. \tag{4.18}$$

**THEOREM 5.** *Under Assumptions A.1 - A.6 we have:*

$$\sqrt{T}\left[\left(\hat{\mathcal{H}}'_c\right)^{-1} \hat{\lambda}^c_{j,i} - \lambda^c_{j,i}\right] \xrightarrow{d} N\left(0, \Phi^{(cc)}_{j,i} + (\lambda^{s\prime}_{j,i} \otimes I_{k^c}) \Psi_j (\lambda^s_{j,i} \otimes I_{k^c})\right), \tag{4.19}$$

$$\sqrt{T}\left[\left(\hat{\mathcal{H}}'_{s,j}\right)^{-1} \hat{\lambda}^s_{j,i} - \lambda^s_{j,i}\right] \xrightarrow{d} N\left(0, \Phi^{(ss)}_{j,i}\right), \tag{4.20}$$

*for any $j$, $i$, where $\hat{\mathcal{H}}_c$ and $\hat{\mathcal{H}}_{s,j}$, $j = 1, 2$, are the same non-singular matrices of Theorem 4.*

**Proof**: See Appendix A.6.

The factor loadings are estimated at rate $\sqrt{T}$. To get a feasible distributional result for the statistic $\hat{\xi}(k^c)$, we need consistent estimators for the unknown matrices $\tilde{\Sigma}_{cc}$, $\Sigma_{U,N}$ and $\Omega_U$ in Theorem 3. To simplify the analysis, we assume at this stage that the errors $\varepsilon_{j,it}$ are uncorrelated across subpanels $j$, individuals $i$ and dates $t$ (Assumption A.7). [12] Then, we have:

$$\Sigma_{U,N} = \mu^2_N \Sigma^{(cc)}_{u,1} + \Sigma^{(cc)}_{u,2}, \quad \Sigma_U(0) = \mu^2 \Sigma^{(cc)}_{u,1} + \Sigma^{(cc)}_{u,2}, \quad \Omega_U = 2tr\left\{\Sigma_U(0)^2\right\}. \tag{4.21}$$

In Theorem 6 below, we replace $\tilde{\Sigma}_{cc}$, $\Sigma_{U,N}$ and $\Sigma_U(0)$ by consistent estimators, such that the estimation error for $tr(\tilde{\Sigma}^{-1}_{cc} \Sigma_{U,N})$ in the bias adjustment is $o_p(1/\sqrt{T})$. Therefore, the asymptotic distribution of

[11] We assume that $\hat{f}^c_t$ is used for the estimation of the factor loadings. The distribution of the loadings estimators is analogous when using $\hat{f}^{c\,*}_t$ as common factor estimator.

[12] If the errors are weakly correlated across series and/or time, consistent estimation of $\Sigma_{U,N}$ and $\Omega_U$ requires thresholding of estimated cross-sectional covariances and/or HAC-type estimators.

the statistic is unchanged.

**THEOREM 6.** *Let* $\hat{\Sigma}_U = (N_2/N_1)\hat{\Sigma}_{u,1}^{(cc)} + \hat{\Sigma}_{u,2}^{(cc)}$*, with*

$$\hat{\Sigma}_{u,j} = \left(\frac{\hat{\Lambda}_j'\hat{\Lambda}_j}{N_j}\right)^{-1}\left(\frac{1}{N_j}\hat{\Lambda}_j'\hat{\Gamma}_j\hat{\Lambda}_j\right)\left(\frac{\hat{\Lambda}_j'\hat{\Lambda}_j}{N_j}\right)^{-1}, \qquad j = 1, 2, \tag{4.22}$$

*where* $\hat{\Gamma}_j = diag(\hat{\gamma}_{j,ii},\ i = 1, ..., N_j)$*, and* $\hat{\Lambda}_j = [\hat{\Lambda}_j^c \vdots \hat{\Lambda}_j^s]$*, where* $\hat{\Lambda}_j^c$ *and* $\hat{\Lambda}_j^s$*, with* $j = 1, 2$*, are the loadings estimators defined in equations (3.3) and (3.4), and*

$$\hat{\gamma}_{j,ii} = \frac{1}{T}\sum_{t=1}^{T}\hat{\varepsilon}_{j,it}^2, \tag{4.23}$$

*where* $\hat{\varepsilon}_{j,it} = y_{j,it} - \hat{\lambda}_{j,i}^{c\,\prime}\hat{f}_t^c - \hat{\lambda}_{j,i}^{s\,\prime}\hat{f}_{j,t}^s$*. Moreover, let* $\hat{\Sigma}_{cc} = \frac{1}{T}\sum_{t=1}^{T}\hat{f}_t^c\hat{f}_t^{c\,\prime}$ *be the estimator of* $\tilde{\Sigma}_{cc}$*. Then, under Assumptions A.1 - A.7, and the null hypothesis* $H_0 = H(k^c)$ *of* $k^c$ *common factors, we have:*

$$\tilde{\xi}(k^c) := N\sqrt{T}\left(\frac{1}{2}tr\{\hat{\Sigma}_U^2\}\right)^{-1/2}\left[\hat{\xi}(k^c) - k^c + \frac{1}{2N}tr\left\{\hat{\Sigma}_{cc}^{-1}\hat{\Sigma}_U\right\}\right] \xrightarrow{d} N(0,1). \tag{4.24}$$

**Proof**: See Appendix A.7.

The feasible asymptotic distribution in Theorem 6 is the basis for a one-sided test of the null hypothesis of $k^c$ common factors. If $\tilde{\xi}(k^c) < -1.64$, this null hypothesis is rejected at $5\%$ level against the alternative hypothesis of less than $k^c$ common factors.

## 4.2 Main asymptotic results for the mixed frequency factor model

In this section we give the asymptotic distribution for estimators of factor values in the mixed frequency factor model. The asymptotics is for $N_H, N_L, T \to \infty$, such that $N_L \leq N_H$, $\sqrt{N_H}/T = o(1)$, $N_L/T = o(1)$. Define the matrices:

$$\Omega_{\Lambda,m}^* = \lim_{N_H\to\infty}\frac{1}{N_H}\sum_{i=1}^{N_H}\sum_{\ell=1}^{N_H}\lambda_{1,i}\lambda_{1,\ell}'Cov(e_{m,t}^{i,H}, e_{m,t}^{\ell,H}), \qquad m = 1, ..., M, \tag{4.25}$$

where $\lambda_{1,i}'$ is the $i$-th row of the $(N_H, k^C + k^H)$ matrix $\Lambda_1 = [\ \Lambda_{HC} \vdots \Lambda_H\ ]$.

**THEOREM 7.** *Under Assumptions A.1 - A.8 we have:*

$$\sqrt{N_H}(\hat{\mathcal{H}}_c \hat{g}^C_{m,t} - g^C_{m,t}) \overset{d}{\longrightarrow} N\left(0, [\Sigma^{-1}_{\Lambda,1}\Omega^*_{\Lambda,m}\Sigma^{-1}_{\Lambda,1}]^{(CC)}\right), \quad (4.26)$$

$$\sqrt{N_H}\left[\hat{\mathcal{H}}_{1,s}\hat{g}^H_{m,t} - (g^H_{m,t} - (\bar{g}^{H\prime}\bar{g}^C)(\bar{g}^{C\prime}\bar{g}^C)^{-1})g^C_{m,t}\right] \overset{d}{\longrightarrow} N\left(0, [\Sigma^{-1}_{\Lambda,1}\Omega^*_{\Lambda,m}\Sigma^{-1}_{\Lambda,1}]^{(HH)}\right), \quad (4.27)$$

*for any* $m$*,* $t$*, where* $\hat{\mathcal{H}}_c$ *and* $\hat{\mathcal{H}}_{1,s}$ *are the same non-singular matrices of Theorem 4,* $\bar{g}^C = [\bar{g}^C_1, ..., \bar{g}^C_T]'$*,* $\bar{g}^H = [\bar{g}^H_1, ..., \bar{g}^H_T]'$*,* $\Sigma_{\Lambda,1} = \lim_{N_H \to \infty} \frac{1}{N_H}\sum_{i=1}^{N_H} \lambda_{1,i}\lambda'_{1,i}$*, and indices* $(CC)$ *and* $(HH)$ *denote the upper-left* $(k^C, k^C)$ *block and lower-right* $(k^H, k^H)$ *block of a matrix, respectively.*

**Proof**: See Appendix A.8.

From Theorem 7, a linear transformation of vector $\hat{g}^C_{m,t}$, resp. $\hat{g}^H_{m,t}$, estimates the common factor $g^C_{m,t}$, resp. the residual of the sample projection of the high-frequency factor on the common factor. The estimation rate is $\sqrt{N_H}$. There is no asymptotic effect from the error-in-variable problem induced by using estimated factor loadings in the cross-sectional regression when $T/N_H = o(1)$. The asymptotic distribution of the estimator $\hat{\bar{g}}^L_t$ of the aggregated low-frequency factor is deduced from Theorem 4.

# 5 Empirical application

It is worth summarizing the procedure underpinning the empirical analysis, for the benefit of the readers who skipped the previous sections. This is done in a first subsection.

## 5.1 Practical implementation of the procedure

We first assume that $k^C$, $k^H$, $k^L$, the number of respectively common, high and low frequency factors in equation (2.1), are known and all strictly larger than zero. The identification strategy presented in Section 2 directly implies a simple estimation procedure for the factor values and the factor loadings, which consists of the three following steps:

1. *PCA performed on the HF and LF panels separately*
   Define the $(T, N_H)$ matrix of temporally aggregated (in our application flow-sampled) HF observables as $X^H = [x^H_1, ..., x^H_T]'$, and the $(T, N_L)$ matrix of LF observables as $X^L = [x^L_1, ..., x^L_T]'$. The estimated pervasive factors of the HF data, which are collected in $(T, k^C + k^H)$ matrix

$\hat{h}_H = [\hat{h}_{H,1}, ..., \hat{h}_{H,T}]'$, are obtained performing PCA on the HF data:

$$\left(\frac{1}{TN_H} X^H X^{H\prime}\right) \hat{h}_H = \hat{h}_H \hat{V}_H, \tag{5.1}$$

where $\hat{V}_H$ is the diagonal matrix of the eigenvalues of $(TN_H)^{-1} X^H X^{H\prime}$. Analogously, the estimated pervasive factors of the LF data, which are collected in the $(T, k^C + k^L)$ matrix $\hat{h}_L = [\hat{h}_{L,1}, ..., \hat{h}_{L,T}]'$, are obtained performing PCA on the LF data:

$$\left(\frac{1}{TN_L} X^L X^{L\prime}\right) \hat{h}_L = \hat{h}_L \hat{V}_L, \tag{5.2}$$

where $\hat{V}_L$ is the diagonal matrix of the eigenvalues of $(TN_L)^{-1} X^L X^{L\prime}$.

2. *Canonical correlation analysis performed on estimated principal components*
Let $\hat{W}_U^C$ be the $(k^C + k^U, k^C)$ matrix whose columns are the canonical directions for $\hat{h}_{U,t}$ associated with the $k^C$ largest canonical correlations between $\hat{h}_H$ and $\hat{h}_L$, for $U = H, L$. Then, the estimator of the (in our application flow sampled) common factor is $\hat{\bar{g}}_t^C = \hat{W}_U^{C\prime} \hat{h}_{U,t}$, for $U = H, L$ and $t = 1, ..., T$, and the estimated loadings matrices $\hat{\Lambda}_{HC}$ and $\hat{\Lambda}_C$ are obtained from the least squares regressions of $x_t^H$ and $x_t^L$ on estimated factor $\hat{\bar{g}}_t^C$. Collect the residuals of these regressions:

$$\begin{aligned} \hat{\bar{\xi}}_t^H &= x_t^H - \hat{\Lambda}_{HC} \hat{\bar{g}}_t^C, \\ \hat{\bar{\xi}}_t^L &= x_t^L - \hat{\Lambda}_{LC} \hat{\bar{g}}_t^C, \end{aligned}$$

in the following $(T, N_U)$, with $U = H, L$, matrices:

$$\hat{\bar{\Xi}}^U = \left[\hat{\bar{\xi}}_1^{U\prime}, ..., \hat{\bar{\xi}}_T^{U\prime}\right]', \qquad U = H, L.$$

Then the estimators of the HF-specific and LF-specific factors, collected in the $(T, k^U)$, $U = H, L$, matrices:

$$\hat{G}^U = \left[\hat{\bar{g}}_1^{U\prime}, ..., \hat{\bar{g}}_T^{U\prime}\right]', \qquad U = H, L,$$

are obtained extracting the first $k^H$ and $k^L$ PCs from the matrices:

$$\left(\frac{1}{TN_U} \hat{\bar{\Xi}}^U \hat{\bar{\Xi}}^{U\prime}\right) \hat{G}^U = \hat{G}^U \hat{V}_S^U, \qquad U = H, L,$$

where $\hat{V}_S^U$, with $U = H, L$ are the diagonal matrices of the associated eigenvalues. Next, the estimated loadings matrices $\hat{\Lambda}_H$ and $\hat{\Lambda}_C$ are obtained from the least squares regression of $\hat{\xi}_t^H$ and $\hat{\xi}_t^L$ on respectively the estimated factors $\hat{\bar{g}}_t^H$ and $\hat{\bar{g}}_t^L$.

3. *Reconstruction of the common and high frequency-specific factors*
The estimates of the common and HF-specific factors for each HF subperiod, denoted by $\hat{g}_{m,t}^C$ and $\hat{g}_{m,t}^H$, for any $m = 1, ..., M$ and $t = 1, ..., T$, are obtained by cross-sectional regression of

$x_{m,t}$ on the estimated loadings $[\hat{\Lambda}_{HC} \vdots \hat{\Lambda}_H]$ obtained from the second step.

Inference on the number of common, low and high-frequency specific factors proceeds as follows:

- Suppose that $k_X := k^C + k^H$ and $k_Y := k^C + k^L$, i.e. the numbers of pervasive factors in panels $X$ and $Y$, are known (consistent estimators: $IC_{p1}$ and $IC_{p2}$ criteria of Bai and Ng(2002)).
- Let $k^* := \min(k_X, k_Y)$, we develop a test for:

$$H_0 : k^C = r \qquad \text{against} \qquad H_1 : k^C < r,$$

  for any given $r = k^*,\ k^* - 1,\ ...,\ 1$.
- We use the statistic defined in equation (3.5), namely: $\hat{\xi}(r) = \sum_{\ell=1}^{r} \hat{\rho}_\ell$, where $\hat{\rho}_\ell$, $\ell = 1, ..., r$, are the $r$ largest canonical correlations between $\hat{h}_{H,t}$ and $\hat{h}_{L,t}$ (i.e. the empirical analogs of $h_{H,t}$ and $h_{L,t}$).

## 5.2 Data description

The data consists of a combination of IP and non-IP sectors. For industrial production we use the same data on 117 IP sectoral indices considered by Foerster, Sarte, and Watson (2011), sampled at quarterly frequency from 1977.Q1 to 2011.Q4.[13] These indices correspond to the finest level of disaggregation for the sectoral components of the IP aggregate index which can be matched with the available sectors in the *Input-Output* and *Capital Use* tables used in the structural analysis in Section 5.4. The data for all the remaining non-IP sectors consist of the annual growth rates of real GDP for the following 42 sectors: 35 services, Construction, Farms, Forestry-Fishing and related activities, General government (federal), Government enterprises (federal), General government (state and local) and Government enterprises (state and local). These LF data are available from 1977 until 2011 and are published by the Bureau of Economic Analysis (BEA).[14] Moreover, as IP is a Gross Output measure, in the structural analysis it is convenient to consider the yearly growth rates of real Gross Output (GO) for the non-IP sectors. These data are available from 1988 until 2011 and are also published by the BEA. Following the sectoral productivity literature, in the structural analysis we focus exclusively on the private sectors, and therefore exclude four Government Gross Output indices, reducing the sample

[13]The IP data are available also at monthly frequency. Following Foerster, Sarte, and Watson (2011), we focus only on quarterly IP data, as they share the main feature of the monthly ones, but are less noisy.

[14]GDP data are available at quarterly frequency for the aggregate index, but not for sectoral ones. As in the remaining part of the paper we study comovements among different sectors, we consider the panel of yearly GDP sectoral data.

size to 38 non-IP sectors indices. All growth rates refer to seasonally adjusted real output indices, and are expressed in percentage points.[15]

Figure 2: Growth rates of the Industrial Production and Gross Domestic Product indices

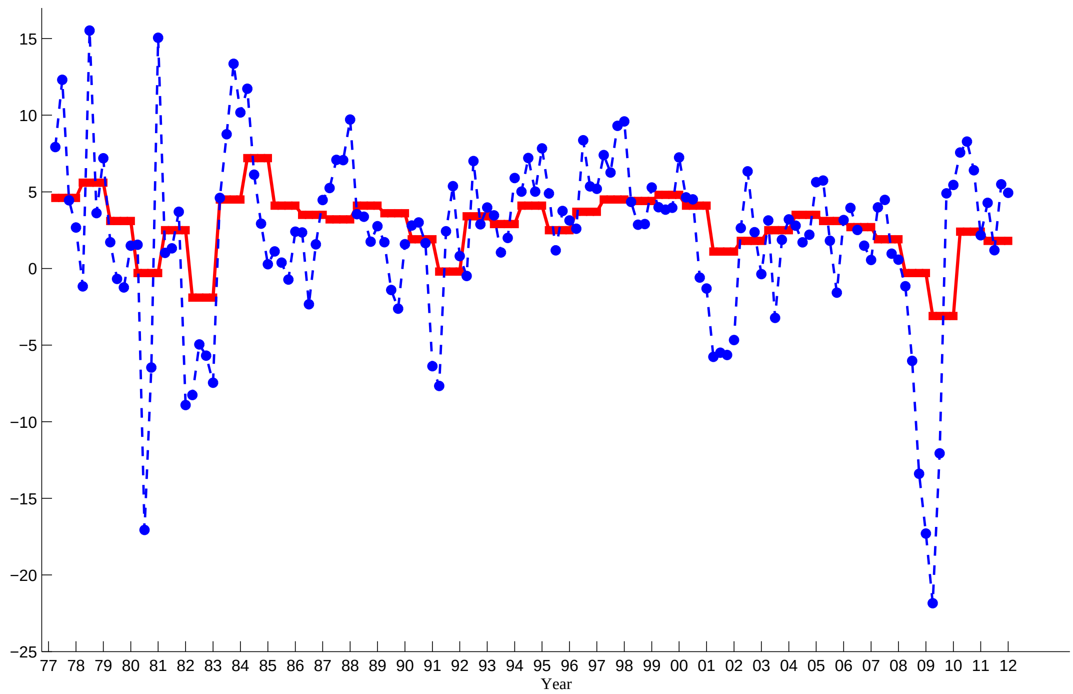



Figure 2 displays the growth rates of the aggregate Industrial Production (dotted (blue) quarterly data) and Gross Domestic Product (solid line (red) annual data) indices over the sample period from 1977 until 2011. The objective of this empirical application is to use our mixed frequency factor model to capture the major sources of comovement among the sectoral constituents of these two indices, which are the most reliable measures of US economic activity.

## 5.3 Factors common to all US sectors

We assume that our dataset follows the factor structure for flow sampling as in equation (2.2), with $x^H_{m,t}$ and $x^L_t$ corresponding to respectively quarterly IP and annual non-IP data. Let $X^H = [x^H_1, ..., x^H_T]'$, with $x^H_t := \sum_{m=1}^{4} x^H_{m,t}$, be the $(T, N_H)$ panel of the yearly observations of the IP indices growth rates (computed as the sum of the quarterly growth rates $x^H_{m,t}$, $m = 1, ..., 4$ for year $t$), and let $X^L =$

[15]A detailed description of the dataset is provided in the Online Appendix OA.2.

$[x_1^L, ..., x_T^L]'$ be the $(T, N_L)$ panel of the yearly growth rates of the non-IP indices. Let also $X_{HF} = [x_{1,1}^H, x_{2,1}^H, ..., x_{m,t}^H, ..., x_{4,T}^H]'$ be the $(4T, N_H)$ panel of IP indices quarterly growth rates.

We start by selecting the number of factors in each subpanel, which are of dimensions $k_X = k^C + k^H$ and $k_Y = k^C + k^L$, respectively. We use the $IC_{p2}$ information criteria of Bai and Ng (2002), and report the results in Table 1. Results for other criteria are in the Online Appendix OA.4.

Table 1: Estimated number of factors

| | $X_{HF}$ | $X^H$ | $X^L$ | $[\ X^H\ \ X^L\ ]$ |
|---|---|---|---|---|
| IP data: 1977.Q1-2011.Q. Non-IP data: Gross Domestic Product, 1977-2011 | | | | |
| $IC_{p2}$ | 1 | 2 | 1 | 1 |
| IP data: 1988.Q1-2011.Q4. Non-IP data: Gross Output, 1988-2011 | | | | |
| $IC_{p2}$ | 1 | 1 | 2 | 2 |

The number of latent pervasive factors selected by the $IC_{p2}$ information criteria is reported for different subpanels. Subpanels $X_{HF}$ and $X^H$ correspond to IP data sampled at quarterly and yearly frequency, respectively. Panels $X^L$ and $[\ X^H\ \ X^L\ ]$ correspond to non-IP data, and the stacked panels of IP and non-IP data, respectively. We use $k_{max} = 15$ as maximum number of factors when computing $IC_{p2}$.

Table 1 corroborates the evidence in Foerster, Sarte, and Watson (2011) suggesting that there is either one or perhaps two pervasive factors in the IP data ($k_X = 1$ or $k_X = 2$). Likewise, for the non-IP data, we also find evidence in favor of either one or two pervasive factors ($k_Y = 1$ or $k_Y = 2$).

Table 2: Canonical Correlations and Tests for Common Factors

| $\hat{\rho}_1$ | $\hat{\rho}_2$ | $\tilde{\xi}(2)$ | $\tilde{\xi}(1)$ |
|---|---|---|---|
| 0.84 | 0.06 | -3.56 | -1.56 |
| 0.80 | 0.11 | - | - |

Top panel: IP data: 1977.Q1-2011.Q4, Non-IP data: GDP, 1977-2011. Lower panel: IP data: 1988.Q1-2011.Q4, Non-IP data: Gross Output, 1988-2011. In rows 1 and 2 we report the canonical correlations of the first two PCs computed in each subpanel of IP and non-IP data, and the values of $\tilde{\xi}(r)$, the feasible standardized value of the test statistic $\hat{\xi}(r)$, for the null hypothesis of $r = 2$ or $r = 1$ common factors, respectively.

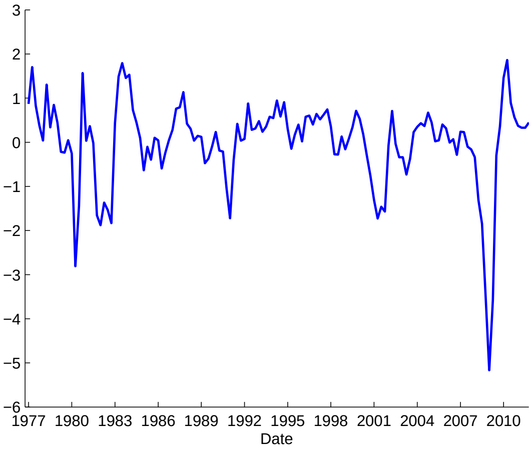
In order to select the number of common and frequency-specific factors, we follow the procedure detailed in Section 5.1. In Table 2 we report the estimated canonical correlations of the first two PC's estimated in each subpanel $X^H$ and $X^L$, which are used to compute the value of the test statistic $\hat{\xi}(r)$, for the null hypothesis of $r$ = 2 or $r$ = 1 common factors.[16] We note that the first canonical correlation is close to one for both datasets, which is consistent with the presence of one common factor in each of the two mixed frequency datasets considered. The tests reject the null hypotheses $r$ = 2, i.e. the presence of two common factors, for any significance level, while we cannot reject the null of one common factor with a 5% significance level. In light of the results in Tables 1 and 2 we select a model with $k^C = k^H = k^L = 1$, for both the panel where the LF data are GDP non-IP indices as well as for the panel in which the LF data are Gross Output non-IP indices. The factors for both datasets are then obtained using the estimation procedure of Section 5.1.

In Figure 3 we plot the estimated factors from the panels of 42 GDP sectors and 117 IP indices on the entire sample going from 1977 to 2011. All factors are standardized to have zero mean and unit variance, and their sign is chosen so that the majority of the associated loadings are positive. A visual inspection of the plots in Figure 3 reveals that the common factor in Panel (a) resembles the IP index of Figure 2, with a large decline corresponding to the Great Recession following the financial crisis of 2007-2008 and the positive spike associated to the recent economic recovery. On the other hand, the LF-specific factor features a less dramatic fall during the Great Recession and actually features a positive spike in 2008, followed by large negative values in the following years. This constitutes preliminary evidence suggesting that some non-IP sectors could feature different responses to the financial crisis of 2007-2008.

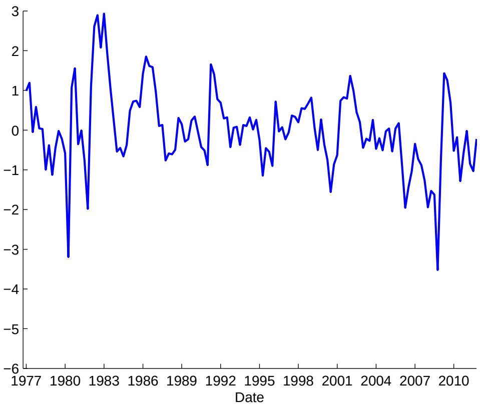
The interpretation of factors is easier when they are used as explanatory variables in standard regression analysis. We start with a disaggregated analysis, and look at the relative importance of the common and frequency specific factors in explaining the variability across all sectoral growth rates. For each sector in the panel, we regress the index growth rates on (i) the common factor only, (ii) on the specific factor only, and (iii) on both common and specific factors. In Table 3 we report the quantiles of the empirical distribution of the adjusted $R^2$ (denoted $\bar{R}^2$) of these regressions. In the first and fourth rows of Panels A and B we report the quantiles of $\bar{R}^2$ of the regressions involving as explanatory variable the common factor only, in the second and fifth rows we report the quantiles of $\bar{R}^2$ when the explanatory variables are the common and frequency-specific factors. Finally, the quantiles of $\bar{R}^2$ in the third and sixth rows refer to regressions where the explanatory variable is the frequency-specific

[16]We extract the first two PC's in each subgroup, compute the matrix $\hat{R}$ as defined in equation 3.2 and compute the canonical correlations as the square root of its two largest eigenvalues.

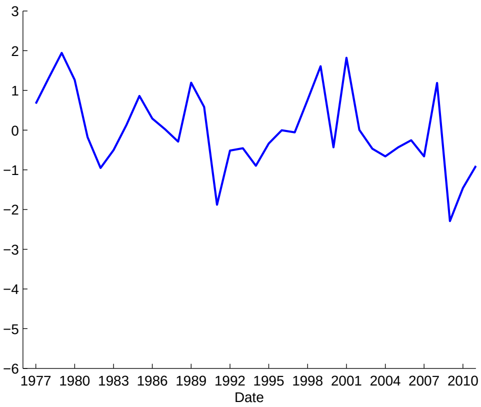
Figure 3: Sample paths of the estimated common and specific factors




(a) Common factor




(b) HF specific factor




(c) LF specific factor

Panel (a) displays the time series plot of the estimated common factor. Panel (b) displays that of the HF-specific factor and finally Panel (c) that of the LF-specific factor. The factors are estimated from the panels of 42 non-IP GDP sectors and 117 IP indices using a mixed frequency factor model with $k^C = k^H = k^L = 1$. The sample period is 1977.Q1-2011.Q4.

factor only.[17]

From the first three lines of Panel A we observe that adding the LF specific factor to the common factor regressions for the non-IP indices yields an increment of the median $\bar{R}^2$ around 14%, going from 11.5% to 25.4%, and for more than 10% of the sectors the $\bar{R}^2$ increases at least by 17%. On the other hand, the HF-specific factor, when added to the common factor, contributes less to the increments in

[17] The regressions in the second and third rows are restricted MIDAS regressions. The regressions in fourth, fifth and sixth rows impose the estimated coefficients of the common and HF-specific factors to be the same for each quarter, as they are estimated as HF regressions.

Table 3: Adjusted $R^2$ of regressions on common factors from indices growth rates

**Panel A**

| Factors | $\bar{R}^2$: Quantile 10% | 25% | 50% | 75% | 90% |
|---|---|---|---|---|---|
| *Observables: Gross Domestic Product, 1977-2011* | | | | | |
| common | -2.2 | -0.5 | 11.5 | 28.9 | 42.9 |
| common, LF-spec. | 0.1 | 9.2 | 25.4 | 34.5 | 60.3 |
| LF-spec. | -2.8 | -2.3 | 5.7 | 15.7 | 22.4 |
| *Observables: IP, 1977.Q1-2011.Q4* | | | | | |
| common | 0.3 | 4.8 | 20.3 | 36.0 | 60.0 |
| common, HF-spec. | 1.1 | 6.8 | 28.7 | 45.3 | 63.4 |
| HF-spec. | -0.7 | -0.1 | 3.0 | 11.2 | 23.5 |

**Panel B**

| Factors | $\bar{R}^2$: Quantile 10% | 25% | 50% | 75% | 90% |
|---|---|---|---|---|---|
| *Observables: Gross Output, 1988-2011* | | | | | |
| common | -2.0 | 6.6 | 28.2 | 45.6 | 64.5 |
| common, LF-spec. | 2.8 | 15.2 | 45.0 | 63.7 | 70.8 |
| LF-spec. | -4.5 | -3.8 | 3.2 | 13.4 | 40.7 |
| *Observables: IP, 1988.Q1-2011.Q4* | | | | | |
| common | 0.1 | 3.5 | 10.5 | 29.8 | 48.2 |
| common, HF-spec. | 0.8 | 7.9 | 28.2 | 43.2 | 65.4 |
| HF-spec. | -0.8 | 2.0 | 10.0 | 21.9 | 33.9 |

*Panel A*. The regressions in the first three lines involve the growth rates of the 42 non-IP sectors as dependent variables, while those in the last tree lines involve the growth rates of the 117 IP indices as dependent variables. The explanatory variables are factors estimated from the same indices using a mixed frequency factor model with $k^C = k^H = k^L = 1$. The sample period for the estimation of both the factor model and the regressions is 1977-2011. *Panel B*. The regressions in the first three lines involve the Gross Output growth rates growth of the 38 non-IP as dependent variables, while those in the last tree lines involve the growth of the 117 IP indices as dependent variables. The explanatory variables are factors estimated from the same indices using a mixed frequency factor model with $k^C = k^H = k^L = 1$. The sample period for the estimation of both the factor model and the regressions is 1988-2011.

$\bar{R}^2$ for the IP sectors. In Panel B we note that for at least 50% of both the IP and non-IP Gross Output sectoral indices, the frequency-specific factors contribute to an increase in $\bar{R}^2$ of at least 15% when added to the common factor. Overall, Table 3 confirms that the common and frequency-specific factors explain a significant part of the variability of output growth for the majority of the sectors of the US economy. Moreover, the common factor is pervasive for most of the IP and non-IP sectors alike.

In order to give economic interpretation to the estimated factors, we list in Table 4 the top and bottom ten GDP non-IP sectors in terms of $\bar{R}^2$ when regressed on the common factor only, and both the common and LF-specific factors. We also report the top and bottom ten GDP non-IP sectors with the highest and lowest absolute increments in $\bar{R}^2$ when the LF-specific factor is added to the common one.[18] From Panel A we first note that the common factor explains most of the variability of service sectors with direct economic links to industrial production sectors like Transportation and Warehousing: for instance, Truck Transportation, Other Transportation & Support Activities, and Warehousing & Storage have an $\bar{R}^2$ of 63%, 43% and 41%, respectively, when regressed on the common factor only. This is a clear indication that the common factor could be interpreted as IP factor.

[18]The entire lists of ordered non-IP sectors for the three panels in Table 4 is available in Tables OA.7-OA.9 in the Online Appendix.

Table 4: Regression of yearly sectoral GDP growth on the common and LF-specific factors: adjusted $R^2$

| **Panel A. Regressor: common factor** | | **Panel B. Regressors: common and LF spec. factors** | | **Panel C. Increment in adjusted $R^2$** | |
|---|---|---|---|---|---|
| **Sector** | $\bar{R}^2$ | Sector | $\bar{R}^2$ | Sector | $\Delta\bar{R}^2$ |
| *Ten sectors with largest $\bar{R}^2$* | | *Ten sectors with largest $\bar{R}^2$* | | *Ten sectors with largest change in $\bar{R}^2$* | |
| Truck transportation | 63.10 | Misc. prof., scient., & tech. serv. | 66.67 | Misc. prof., scient., & tech. serv. | 49.69 |
| Accommodation | 62.43 | Administrative & support services | 62.63 | Government enterprises (state & local) | 34.69 |
| Construction | 44.05 | Truck transportation | 62.51 | Rental & leasing serv. & lessors of int. assets | 29.52 |
| Other transp. & support activ. | 43.31 | Accommodation | 61.48 | General government (state & local) | 24.90 |
| Administrative & support services | 42.69 | Construction | 59.75 | Legal services | 24.32 |
| Other services, except government | 42.53 | Warehousing & storage | 52.53 | Motion picture & sound recording ind. | 22.77 |
| Warehousing & storage | 40.95 | Government enterprises (state & local) | 45.78 | Fed. Reserve banks, credit interm., & rel. activ. | 20.31 |
| Air transportation | 31.58 | Other services, except government | 41.75 | Administrative & support services | 19.95 |
| Retail trade | 30.70 | Other transportation & support activities | 41.71 | Social assistance | 19.91 |
| Amusem., gambling, & recr. ind. | 29.17 | Government enterprises (federal) | 37.78 | Real estate | 18.14 |
| *Ten sectors with smallest $\bar{R}^2$* | | *Ten sectors with smallest $\bar{R}^2$* | | *Ten sectors with smallest change in $\bar{R}^2$* | |
| Funds, trusts, & other finan. vehicles | -1.23 | Ambulatory health care services | 7.76 | Accommodation | -0.96 |
| Motion picture & sound record. ind. | -1.68 | Management of companies & enterprises | 7.52 | Rail transportation | -1.16 |
| Pipeline transportation | -1.74 | Funds, trusts, & other fin. vehicles | 6.15 | Other transportation & support activities | -1.59 |
| Information & data processing services | -1.84 | Information & data processing services | 1.96 | Air transportation | -1.77 |
| Transit & ground passenger transp. | -2.05 | Educational services | 1.35 | Retail trade | -2.15 |
| General government (state & local) | -2.12 | Insurance carriers & related activities | 0.36 | Amusements, gambling, & recreation ind. | -2.15 |
| Forestry, fishing, & related activities | -2.33 | Water transportation | -0.64 | Educational services | -2.62 |
| Water transportation | -2.94 | Farms | -1.87 | Farms | -2.80 |
| Securities, commodity contracts, & investm. | -2.99 | Forestry, fishing, & related activities | -5.31 | Forestry, fishing, & related activities | -2.98 |
| Insurance carriers & related activities | -3.03 | Securities, commodity contracts, & investm. | -5.99 | Securities, commodity contracts, & investm. | -3.00 |

In the table we report the adjusted $R^2$, denoted $\bar{R}^2$, for restricted MIDAS regressions of the growth rates of 42 GDP non-IP sectoral indices on the estimated factors. The factors are estimated from the panel of 42 GDP sectors and 117 IP indices using a mixed frequency factor model with $k^C = k^H = k^L = 1$. The sample period for the estimation of both factor model and regressions is 1977-2011. Regressions in *Panel A* involve a LF explained variable and the estimated common factor. Regressions in *Panel B* involve a LF explained variable and both the common and LF-specific estimated factors. The regressions in both tables are restricted MIDAS regressions. In *Panel C* we report the difference in $\bar{R}^2$ (denoted as $\Delta\bar{R}^2$) between the regressions in *Panel B* and regressions in *Panel A*.

On the other hand, the common factor is completely unrelated to Agriculture, forestry, fishing & hunting, most of the Financial and Information services sectors.

Turning to Panel C, we note that the LF-specific factor explains more than 20% of the variability of output for very heterogeneous services sectors like Miscellaneous professional, scientific, & technical services, Administrative & support services, Legal services, Real Estate, some important financial services like Credit intermediation, & Related activities, Rental & Leasing Services but also Government (state & local). Interpreting these results, we can conclude that the LF-specific factor is completely unrelated to service sectors which depend almost exclusively on IP output, and is a common factor driving the comovement of non-IP sectors such as some Services, Construction and Government.

In Table 4 we highlight further differences in the dynamics of output growth between the two subsectors of the financial services industry which are particularly revealing: "Securities" and "Credit intermediation", extensively studied by Greenwood and Scharfstein (2013). We find that the subsectors "Funds, trusts, & other financial vehicles" and "Securities, commodity contracts, & investments" are unrelated to both the common and LF-specific factors, indicating that their output growth is uncorrelated with the common component of real output growth across the other sectors of the US economy. In contrast, the "credit intermediation" industry comoves with the other IP and non-IP sectors.[19]

Up to this point, we looked at the explanatory power of the factors for sectoral output indices. For both the non-IP GDP and Gross Output, these indices correspond to the finest level of disaggregation of output growth by sector. In Table 5 we report the results of regressions with aggregated indices instead. In particular, we regress the output of each aggregate index either on the estimated common factor or the common and frequency specific factors, and focus on the adjusted $R^2$s of these regressions. It is also important to note that we also include the GDP Manufacturing aggregate index which is *not* used in the estimation of the factors. This will help us with the interpretation of the factors - common and frequency-specific - which we obtained.

[19]See also Tables OA.7 and OA.8 in the Online Appendix.

Table 5: Adj. $R^2$ of aggregate IP and selected GDP indices growth rates on estimated factors

**Panel A** *Quarterly observations, 1977.Q1-2011.Q4*

| Sector | (1) $\bar{R}^2(C)$ | (2) $\bar{R}^2(H)$ | (3) $\bar{R}^2(C+H)$ | (3) - (1) |
|---|---|---|---|---|
| Industrial Production | 89.06 | 5.02 | 90.26 | 1.20 |

**Panel B** *Yearly observations, 1977-2011*

| Sector | (1) $\bar{R}^2(C)$ | (2) $\bar{R}^2(L)$ | (3) $\bar{R}^2(C+L)$ | (3) - (1) |
|---|---|---|---|---|
| GDP | 60.54 | 8.59 | 74.21 | 13.67 |
| GDP - Manufacturing | 81.88 | -3.03 | 81.53 | -0.35 |
| GDP - Agriculture, forestry, fishing, and hunting | 1.43 | -2.52 | -1.26 | -2.69 |
| GDP - Construction | 44.05 | 11.22 | 59.75 | 15.70 |
| GDP - Wholesale trade | 20.35 | 7.90 | 30.83 | 10.48 |
| GDP - Retail trade | 30.70 | -2.86 | 28.56 | -2.15 |
| GDP - Transportation and warehousing | 62.14 | -2.95 | 60.97 | -1.17 |
| GDP - Information | 12.14 | 22.28 | 37.57 | 25.43 |
| GDP - Finance, insurance, real estate, rental, and leasing | -1.42 | 21.22 | 21.11 | 22.53 |
| GDP - Professional and business serv. | 30.02 | 30.21 | 65.61 | 35.59 |
| GDP - Educational serv., health care, and social assist. | -1.38 | 18.38 | 18.18 | 19.56 |
| GDP - Arts, entert., recreat., accomm., and food serv. | 53.51 | -2.23 | 53.70 | 0.18 |
| GDP - Government | -2.12 | 22.37 | 20.47 | 22.59 |

In the table we report the adjusted $R^2$, denoted $\bar{R}^2$, of the regression of growth rates of the aggregate IP index and selected aggregated sectoral GDP non-IP output indices on the common factor (column $\bar{R}^2(C)$), the specific HF and LF factors (columns $\bar{R}^2(H)$ and $\bar{R}^2(L)$) only, and the common and frequency-specific factors together (column (3)). The last column displays the difference between the values in the third and first columns. The factors are estimated from the panel of 42 GDP non-IP sectors and 117 IP indices using a mixed frequency factor model with $k^C = k^H = k^L = 1$. The sample period for the estimation of both factor model and regressions is 1977-2011.

Panel A of Table 5 shows that the common factor explains around 90% of the variability in the aggregate IP index. This implies that the common factor can be interpreted as an Industrial Production factor. This is further corroborated in Panel B where we find an $\bar{R}^2$ around $82\%$ for the regression of the GDP Manufacturing Index on the common factor only. As most of the sectors included in the Industrial Production index are Manufacturing sectors, this result is not surprising, but is still worth noting because, as noted earlier, the GDP data on Manufacturing have not been used in the factor estimation, in order not to double-count these sectors in our mixed frequency sectoral panel.[20] As expected from the results in Table 4, more than 60% of the variability of GDP of Transportation and

[20]A detailed discussion of the difference in the sectoral components of the IP index and the GDP Manufacturing index is provided in Appendix OA.2.

Warehousing services index is explained by the common factor only, and the LF-specific factor has no explanatory power. On the other hand, the HF-specific factor seems not to be important in explaining the aggregate IP index, as the $\bar{R}^2$ increases only by 1% when it is added as a regressor to the common factor.[21] This suggests that the HF-specific factor is pervasive only for a subgroup of IP sectors which have relatively low weights in the index, meaning that their aggregate output is a negligible part of the output of the entire IP sector and, consequently, also the entire US economy.[22]

Looking at the aggregate GDP index, we first note that even if the weight of Industrial Production sectors in the aggregate nominal GDP index has always been below 30%, as evident from Figure 1, still 60% of its total variability can be explained exclusively by the common factor which - as shown in Panel A - is primarily an IP factor. This implies that there must be substantial comovement between IP and some important service sectors. Moreover, it appears from the first entry in Panel B that a relevant part of the variability of the aggregate GDP index not due to the common factor is explained by the LF-specific factor (the $\bar{R}^2$ increases by about 14% to 74%).[23] This indicates that significant comovements are present among the most important sectors of the US economy which are not related to manufacturing. Indeed, Panel B in Table 5 indicates that some services sectors such as Professional & Business Services and Information and Construction load significantly both on the common and the LF-specific factor, while some other sectors like Finance and Government load exclusively on the LF-specific factor.[24]

## 5.4 Structural model and productivity shocks

The macroeconomics literature, with the works of Long and Plosser (1983), Horvath (1998) and Carvalho (2007), among many others, has recognized that input-output linkages in both intermediate materials and capital goods lead to propagation of sector-specific shocks in a way that generates comovements across sectors. An important contribution of the work of Foerster, Sarte, and Watson (2011) was to describe the conditions under which an approximate linear factor structure for sectoral

[21] See also Table OA.10 in Appendix OA.4 for the $\bar{R}^2$ of the regression of all GDP indices on the HF factor only, and all the 3 factors together.

[22] These results corroborate the findings of Foerster, Sarte, and Watson (2011), who claim that the main results of their paper are qualitatively the same when considering either one or two common factors extracted from the same 117 IP indices of our study.

[23] See the results in Table OA.10 in the Online Appendix.

[24] The results change when we look at Finance sector disaggregated in Credit Intermediation, “Securities”, Insurance and Real estate, as evident in Table 4.

output growth arises from standard neoclassical multisector models including those linkages. In particular, they develop a generalized version of the multisector growth model of Horvath (1998), which allows them to filter out the effects of these linkages and reconstruct the time series of productivity shocks of each of the sectors for which data on output growth and input-output tables for intermediate materials and capital goods are available. We can characterize this as statistical versus structural factor analysis.

The main objective of this section is to verify the presence of a common factor in the *innovations* of productivity for *all* the sectors (not just IP) of the US economy by means of our mixed frequency factor model. If a common factor is present also in the productivity shocks, then the factor structure uncovered by the reduced form analysis of output growth in Section 5.3 is not only due to interlinkages in materials and capital use among different sectors.

We rely on the same multi-industry real business cycle model described in Section IV of Foerster, Sarte, and Watson (2011) to extract productivity shocks from the time series of the growth rates of the same 117 IP indices considered in the previous section, and the growth rates of 38 non-IP Gross Output of private sectors, therefore excluding the 4 Government indices considered previously.[25] One challenge due to the mixed frequency nature of our output growth dataset consists in the extraction of mixed frequency technological shocks. In the Online Appendix OA.3 we explain how to adapt the algorithms proposed by Foerster, Sarte, and Watson (2011), and based on the work of King and Watson (2002), to estimate technological shocks for our mixed frequency output series. Specifically, the multi-sector business cycle model that we use to filter out the technological shocks correspond to the "Benchmark" model considered by Foerster, Sarte, and Watson (2011) in their Section IV, while the data on input-output and capital use matrices necessary to estimate the model are built from the BEA's 1997 "use table" and "capital flow table", respectively.[26] Using the extracted productivity shocks for the IP and non-IP sectors, denoted $\hat{\varepsilon}^X_{m,t}$ and $\hat{\varepsilon}^Y_t$, respectively, we estimate our mixed frequency factor model with these productivity shock series. The sample period for the estimation of both the factor model and the regressions is 1989-2011, because the productivity shocks can not be computed for the first year of the sample (see Foerster, Sarte, and Watson (2011), especially their equation (B38) on page 10 of their Appendix B). For a direct comparison between the statistical factor model covered

[25]The exclusion of the public sector from the analysis is a standard choice in the sectoral productivity literature.
[26]The last year for which sectoral capital use tables have been constructed by the BEA is 1997.

in the previous subsection and the structural factor analysis, we need to first re-estimate our model with one common, one HF-specific and one LF-specific factors on the panels of growth rates of annual Gross Output non-IP indices (as opposed to the GDP growth indices in Table 5) and the same 117 quarterly sectoral IP indices. The results are reported in Table 6. We expect some difference with the previous results for at least two reasons. First, the dataset in which the non-IP data are Gross Output indices, refers to shorter time period going from 1988, instead of 1977, to the end of 2011, as Gross Output indices are not available before 1988. Second, as the panel in Table 6 does not include the four governmental sectors, we expect that the common and frequency-specific factors may have different dynamics when compared to those extracted from the panel with GDP non-IP sectors.

Table 6: Adj. $R^2$ of aggregate IP and selected Gross Output indices growth rates on estimated factors

**Panel A** *Quarterly observations, 1988.Q1-2011.Q4*

| **Sector** | (1) $\bar{R}^2(C)$ | (2) $\bar{R}^2(H)$ | (3) $\bar{R}^2(C+H)$ | (3) - (1) |
|---|---|---|---|---|
| Industrial Production | 63.71 | 38.32 | 89.48 | 25.78 |

**Panel B** *Yearly observations, 1988-2011*

| **Sector** | (1) $\bar{R}^2(C)$ | (2) $\bar{R}^2(L)$ | (3) $\bar{R}^2(C+L)$ | (3) - (1) |
|---|---|---|---|---|
| GO (all sectors) | 68.54 | 12.20 | 89.66 | 21.12 |
| GO - Manufacturing | 86.08 | -3.05 | 88.94 | 2.86 |
| GO - Agriculture, forestry, fishing, and hunting | -3.21 | 3.35 | -0.25 | 2.96 |
| GO - Construction | 25.30 | 34.16 | 67.15 | 41.84 |
| GO - Wholesale trade | 80.82 | -3.85 | 79.97 | -0.85 |
| GO - Retail trade | 64.72 | -4.50 | 63.15 | -1.57 |
| GO - Transportation and warehousing | 83.82 | -4.51 | 83.22 | -0.60 |
| GO - Information | 33.70 | 38.59 | 81.54 | 47.84 |
| GO - Finance, insurance, real estate, rental, and leasing | 3.37 | 50.30 | 59.29 | 55.92 |
| GO - Professional and business services | 45.13 | 21.97 | 75.48 | 30.36 |
| GO - Educational serv., health care, and social assist. | -4.19 | -1.58 | -6.17 | -1.98 |
| GO - Arts, entert., recreat., accomm., and food serv. | 71.06 | -3.74 | 71.90 | 0.84 |

In the table we display the adjusted $R^2$, denoted $\bar{R}^2$, of the regressions of growth rates of the aggregate IP index and selected aggregated sectoral Gross Output non-IP output indices on the common factor (column $\bar{R}^2(C)$), the specific HF and LF factors (columns $\bar{R}^2(H)$ and $\bar{R}^2(L)$)only, and the common and frequency-specific factor together (column (3)). The last column displays the difference between the values in the third and first columns. The factors are estimated from the panel of 38 Gross Output non-IP sectors and 117 IP indices using a mixed frequency factor model with $k^C = k^H = k^L = 1$. The sample period for the estimation of both factor model and regressions is 1988-2011.

We obtain qualitatively similar results, as shown in Table 6. There appear to be only two notable differences with the results reported in Table 5. We see an increased importance of the HF-specific factor in explaining the variability of the IP aggregate index (see Panel A in Table 6), at the expense of a lower explanatory power for the common factor. Moreover, there is also an increased importance of both the common and LF-specific factors in explaining the total variability of total aggregate output (measured as total Gross Output, in the first line of Panel B in Table 6). Still the common factor explains roughly 65 % of the variation in the panel of IP data.

What do we learn from the structural analysis with common and frequency-specific factors of productivity shocks? First, it is remarkable to find that again there is one common factor in productivity shocks. Indeed, the selection of the number of common factors is performed as in the previous section, and our testing methodology suggests the presence of one common factor. Therefore we estimate a model for the productivity innovations with $k^C = k^H = k^L = 1$.[27] As in the previous section, we start with a disaggregated analysis and look at the relative importance of the new common and frequency specific factors in explaining the variability of the constituents of the panel of productivity innovations, and the panels of all output growth rates used for the extraction of the productivity innovation themselves. For each sector, we regress both the productivity innovations and the index growth rates on the common factor only, on the specific factor only, and on both common and specific factors. In Table 7 we report the quantiles of the empirical distribution of the adjusted $R^2$ (denoted $\bar{R}^2$) of these regressions.[28] Panel A of Table 7 confirms that both the common and the frequency-specific factors are pervasive for the panels of productivity innovations. From the first two rows we note that the common factor alone explains at least 11% of the variability of half of the non-IP series considered, and this fraction increases to more than 26 % when the LF-specific factor is added as regressor to the common one. On the other hand, from the last three rows of we note that for the panels of IP the high frequency specific factor seems to explain the majority of the variability of the productivity indices, while the explanatory power of the common factor only seem to be significant only for 50% of the IP sectors. Panel B reports the $\bar{R}^2$ of the regressions of the GO indices growth rates on the factors estimated on the panels of productivity shocks themselves. Therefore, they give an indication of the fraction of

[27]The values of the penalized selection criteria of Bai and Ng (2002) performed on different subpanels and the test for the number of common factors are available in Tables OA.12 and OA.13 in the Online Appendix OA.4.

[28]The regressions in the second and third rows are restricted MIDAS regressions. The regressions in fourth, fifth and sixth rows impose the estimated coefficients of the common and HF-specific factors to be the same at each quarter, as they are estimated as HF regressions.

Table 7: Adjusted $R^2$ of regressions on common factors from productivity innovations

**Panel A**

| Factors | Adjusted $R^2$: Quantile 10% | 25% | 50% | 75% | 90% |
|---|---|---|---|---|---|
| *Observables: Gross Output productivity innovations, 1989-2011* | | | | | |
| common | -3.3 | -0.3 | 11.0 | 33.6 | 46.1 |
| common, LF-spec. | -2.6 | 4.8 | 26.3 | 45.0 | 60.7 |
| LF-spec. | -4.2 | -3.6 | -0.1 | 17.7 | 33.1 |
| *Observables: IP productivity innovations, 1989.Q1-2011.Q4* | | | | | |
| common | -1.0 | -0.4 | 1.5 | 12.1 | 22.4 |
| common, HF-spec. | -0.6 | 3.1 | 13.1 | 28.4 | 40.1 |
| HF-spec. | -0.7 | 0.6 | 6.2 | 18.7 | 28.2 |

**Panel B**

| Factors | Adjusted $R^2$: Quantile 10% | 25% | 50% | 75% | 90% |
|---|---|---|---|---|---|
| *Observables: Gross Output, 1988-2011* | | | | | |
| common | -2.4 | 3.7 | 21.2 | 31.5 | 55.8 |
| common, LF-spec. | -0.9 | 7.8 | 28.2 | 56.9 | 68.0 |
| LF-spec. | -4.6 | -3.3 | 1.3 | 20.6 | 43.8 |
| *Observables: IP,1988.Q1-2011.Q4* | | | | | |
| common | -0.8 | 0.2 | 4.5 | 17.7 | 34.7 |
| common, HF-spec. | 1.2 | 5.9 | 25.7 | 40.8 | 63.8 |
| HF-spec. | -0.3 | 2.2 | 14.7 | 29.2 | 37.8 |

Panel A: The regressions in the first three lines involve the productivity innovations of the 38 non-IP sectors as dependent variables, while the regressions in the last tree lines involve the productivity innovations of the 117 IP indices as dependent variables. Productivity innovations are computed using the panel of Gross Output growth rates for the LF observables. The explanatory variables are factors estimated from a mixed frequency factor model with $k^C = k^H = k^L = 1$, on the panels of productivity innovations filtered adapting the procedure of Foerster, Sarte, and Watson (2011). The sample period for the estimation of both the factor model and the regressions is 1989.Q1-2011.Q4. Panel B: The regressions in the first three lines involve the Gross Output growth rates of the 38 non-IP sectors as dependent variables, while the regressions in the last tree lines involve the growth of the 117 IP indices as dependent variables. The explanatory variables are the same factors used in the regressions of Panel A. The sample period for the estimation of both the factor model and the regressions is 1989.Q1-2011.Q4. Productivity innovations are computed using the panel of Gross Output growth rates for the LF observables.

variability of the indices explained by the common components of the output growth which is not due to input-output linkages between sectors, as captured by the structural "Benchmark" of Foerster, Sarte, and Watson (2011). Panel B of Table 7 can be compared with Panel B of Table 3. As expected, as part of the comovement among different sectors is due to input-output and capital use linkages, all the $\bar{R}^2$ in Panel B of Table 7 are strictly lower than those in Table 3, if we exclude the negative ones and those very close to zero. For instance the median $\bar{R}^2$ of regressions including the common only factor for the non-IP sectors decrease from 28% to 21%, and median $\bar{R}^2$ of regressions including the common and LF-specific factors decreases from 45% to 28%. A similar pattern is observed for the higher quantiles, and for the IP indices. Overall, Panel B gives a first indication of the presence of commonality in the comovement on the majority of the sectors of the US economy even when the output growth rates are purged of the input-output linkages in both intermediate materials and capital goods.

We conclude the analysis repeating the same exercise of Table 6, and regress the Industrial Production and aggregate (mostly non-IP) Gross Output indices growth on the factors extracted from productivity innovations and look at the adjusted $R^2$s in Table 8.

Table 8: Adj. $R^2$ aggregate IP and selected Gross Output indices on the estimated factors from productivity innovations

**Panel A** *Quarterly observations, 1988.Q1-2011.Q4*

| Sector | (1) $\bar{R}^2(C)$ | (2) $\bar{R}^2(H)$ | (3) $\bar{R}^2(C+H)$ | (3) - (1) |
|---|---|---|---|---|
| Industrial Production (Q) | 31.21 | 50.15 | 77.25 | 46.05 |

**Panel B** *Yearly observations, 1989-2011*

| Sector | (1) $\bar{R}^2(C)$ | (2) $\bar{R}^2(L)$ | (3) $\bar{R}^2(C+L)$ | (3) - (1) |
|---|---|---|---|---|
| GO (all sectors) | 42.17 | 13.97 | 57.60 | 15.43 |
| GO - Manufacturing | 62.29 | -0.20 | 64.42 | 2.13 |
| GO - Agriculture, forestry, fishing, and hunting | 0.96 | -4.23 | -3.35 | -4.31 |
| GO - Construction | 6.64 | 20.55 | 27.78 | 21.14 |
| GO - Wholesale trade | 74.73 | -3.08 | 74.74 | 0.01 |
| GO - Retail trade | 47.02 | -4.35 | 45.04 | -1.98 |
| GO - Transportation and warehousing | 70.42 | -2.69 | 70.58 | 0.15 |
| GO - Information | 17.78 | 42.45 | 61.76 | 43.98 |
| GO - Finance, insurance, real estate, rental, and leasing | -4.09 | 17.55 | 13.96 | 18.05 |
| GO - Professional and business services | 25.17 | 44.89 | 71.81 | 46.64 |
| GO - Educational services, health care, and social assistance | -4.73 | -4.48 | -9.66 | -4.93 |
| GO - Arts, entert., recreat., accommodation, and food serv. | 55.64 | -2.29 | 55.49 | -0.16 |

In the table we report the adjusted $R^2$, denoted $\bar{R}^2$, of the regressions of growth rates of the aggregate IP index and selected aggregated sectoral Gross Output non-IP output indices on the common factor (column $\bar{R}^2(C)$), the specific HF and LF factors (columns $\bar{R}^2(H)$ and $\bar{R}^2(L)$)only, and the common and frequency-specific factor together (column (3)). The last column displays the difference between the values in the third and first columns. The factors are estimated from the panels of productivity innovations filtered adapting the procedure of Foerster, Sarte, and Watson (2011), using a mixed frequency factor model with $k^C = k^H = k^L = 1$. The sample period for the estimation of both the factor model and the regressions is 1989.Q1-2011.Q4.

From Panel A we observe that the common extracted from productivity innovations explains around 31% of the variability of the aggregate IP index, i.e. around half of the variability explained by the common factor extracted directly from the output series. Moreover, when the high frequency-specific productivity factor is added as explanatory variable, the $\bar{R}^2$ increases to 77% which is also significantly smaller than the 89% $\bar{R}^2$ obtained using as regressors the factors extracted from the output series.[29] Hence, the case of a common pervasive factor in innovation shocks across the entire economy mainly related to IP sector technology shocks is less compelling. From Panel B we observe that 42% of the

[29]See in particular Panel A of Table 6. This result is in line with the findings of Foerster, Sarte, and Watson (2011) in their Section IV C.

variability of the aggregate Gross Output of the US economy can be explained by the common factor of productivity shocks, and when the factor specific to non-IP sector is added, the $\bar{R}^2$ grows to $57\%$.

From this analysis we learn something interesting which Foerster, Sarte, and Watson (2011) were not able to address since they exclusively examined IP sectors. Overall there is a difference in the explanatory power of factors in structural versus non-structural factor models - as they found. However, it seems that looking at technology shocks instead of output, it does not appear that a common factor explaining IP fluctuations is a dominant factor for the entire economy. A factor specific to technological innovations in IP sectors is more important for the IP sector shocks and a low frequency factor which appears to explain variation in information industry as well as professional and business services innovations plays, relatively speaking, a more important role.

## 5.5 Subsample analysis

Our sample covers what is known as the Great Moderation, which refers to a reduction in the volatility of business cycle fluctuations starting in the mid-1980s. We turn therefore to analyzing subsamples. We start by selecting the number of pervasive factors in each subpanel, using the $IC_{p2}$ information criteria, and report the results in Table 9. We consider two subsample configurations: 1984.Q1-2007.Q4 and 1984.Q1-2011.Q4. The former is the Great Moderation sample considered by Foerster, Sarte, and Watson (2011) whereas the second is an augmented subsample including the Great Depression. In light of the results in Tables 9 and 10 we select a model with $k^C = k^H = k^L = 1$, for both subsamples. The factors for both datasets are obtained using the estimation procedure described in Section 5.1. [30]

In Table 11 we report the results of regressions of aggregated version of the indices used for the estimation on the same factors considered in the full samples. This allows us to understand if, and to what extent, the most important sectors of the US economy comoved over the different subsamples. Again, we regress the output of each aggregate index on the estimated common factor only, the common and frequency specific factors, and concentrate our attention on the adjusted $R^2$s of these regressions.

[30]For the shorter sample 1984.Q1-2007.Q4, selecting a model with $k_1 = k_2 = 3$ pervasive factors in each subpanel, we reject the null hypotheses of 3 and 2 common factor, while we cannot reject the null of 1 common factor. Regression results for $k^C = 1$ and $k^H = k^L = 2$ are very similar than those presented in Table 11, i.e. for a model with $k^C = k^H = k^L = 1$ factors, and therefore are omitted.

Table 9: Estimated number of factors for different subsamples

| | $X_{HF}$ | $X^H$ | $X^L$ | $[\, X^H \;\; X^L \,]$ |
|---|---|---|---|---|
| IP data: 1984.Q1-2007.Q4. Non-IP data: Gross Domestic Product, 1984-2007 | | | | |
| $IC_{p2}$ | 1 | 2 | 1 | 1 |
| IP data: 1984.Q1-2011.Q4. Non-IP data: Gross Domestic Product, 1984-2011 | | | | |
| $IC_{p2}$ | 1 | 2 | 1 | 1 |

The number of latent pervasive factors selected by the $IC_{p2}$ information criteria is reported for different subpanels and different sample periods. Subpanels $X_{HF}$ and $X^H$ correspond to IP data sampled at quarterly and yearly frequency, respectively. Panels $X^L$ and $[\; X^H \;\; X^L \;]$ correspond to non-IP data, and the stacked panels of IP and non-IP data, respectively. We use $k_{max} = 15$ as maximum number of factors when computing $IC_{p2}$.

Table 10: Canonical Correlations and Tests for Common Factors

| $\hat{\rho}_1$ | $\hat{\rho}_2$ | $\hat{\rho}_3$ | $\tilde{\xi}(3)$ | $\tilde{\xi}(2)$ | $\tilde{\xi}(1)$ |
|---|---|---|---|---|---|
| IP data: 1984.Q1-2007.Q4. Non-IP data: Gross Domestic Product, 1984-2007 | | | | | |
| 0.81 | 0.13 | - | - | -6.61 | -2.98 |
| 0.87 | 0.57 | 0.45 | -3.15 | -2.74 | -1.03 |
| IP data: 1984.Q1-2011.Q4. Non-IP data: Gross Domestic Product, 1984-2011 | | | | | |
| 0.70 | 0.33 | - | - | -1.67 | -1.28 |

We report the canonical correlations of the first two PCs computed in each subpanel of IP and non-IP data, and the values of $\tilde{\xi}(r)$, the estimated value of the test statistic $\hat{\xi}(r)$, for the null hypothesis of $r = 3, 2, 1$ common factors, respectively.

The results in Table 11 indicate that in general there is a deterioration of the overall fit of approximate factor models during the Great Moderation, i.e. during the sample starting in 1984 and ending 2007 – a finding also reported by Foerster, Sarte, and Watson (2011) – and that the common factor plays a lesser role during the Great Moderation. According to the results in Panel A, the common factor only explains roughly 72 % of the variation across IP sectors, but interestingly when the financial crisis is added to the Great Moderation subsample, we see again a pattern closer to the full sample results reported in the previous subsection. This also transpires from Panels B and C, when examining the total GDP variations projected on the common factor. During the Great Moderation the common factor only explained around 30 %, which goes to 56 % when we add the Great Depression. The other patterns, i.e. the exposure of the various subindices, appear to be similar to those in the full sample.

Table 11: Adj. $R^2$ of aggregate IP and selected GDP indices growth rates on estimated factors

| **Sector** | (1) $\bar{R}^2(C)$ | (2) $\bar{R}^2(H)$ | (3) $\bar{R}^2(C+H)$ | (3) - (1) |
|---|---|---|---|---|
| **Panel A** *Quarterly observations IP* | | | | |
| IP 1984.Q1-2007.Q4 | 72.48 | 10.58 | 80.02 | 7.54 |
| IP 1984.Q1-2011.Q4 | 80.11 | 16.83 | 88.87 | 8.76 |

| **Sector** | (1) $\bar{R}^2(C)$ | (2) $\bar{R}^2(L)$ | (3) $\bar{R}^2(C+L)$ | (3) - (1) |
|---|---|---|---|---|
| **Panel B** *Yearly observations, 1984-2007* | | | | |
| GDP | 29.22 | 39.24 | 76.71 | 47.49 |
| GDP - Manufacturing | 70.69 | -3.85 | 71.18 | 0.50 |
| GDP - Agriculture, forestry, fishing, and hunting | 0.81 | -0.87 | 0.51 | -0.30 |
| GDP - Construction | 13.02 | 50.30 | 70.39 | 57.37 |
| GDP - Wholesale trade | -4.40 | 21.36 | 18.09 | 22.49 |
| GDP - Retail trade | -0.44 | 58.14 | 62.65 | 63.09 |
| GDP - Transportation and warehousing | 41.43 | 11.16 | 52.02 | 10.59 |
| GDP - Information | -4.37 | -4.10 | -8.83 | -4.46 |
| GDP - Finance, insurance, real estate, rental, and leasing | -3.78 | -0.60 | -4.78 | -1.00 |
| GDP - Professional and business services | 4.89 | 56.09 | 67.06 | 62.18 |
| GDP - Educational serv., health care, and social assist. | -3.81 | 3.31 | -0.20 | 3.61 |
| GDP - Arts, entert., recreat., accomm., and food serv. | 13.66 | 37.32 | 57.01 | 43.35 |
| GDP - Government | 0.74 | 14.51 | 14.83 | 14.09 |
| **Panel C** *Yearly observations, 1984-2011* | | | | |
| GDP | 56.33 | 14.88 | 77.87 | 21.55 |
| GDP - Manufacturing | 83.78 | -3.85 | 83.37 | -0.41 |
| GDP - Agriculture, forestry, fishing, and hunting | -3.64 | -2.65 | -6.59 | -2.95 |
| GDP - Construction | 40.54 | 21.76 | 68.61 | 28.07 |
| GDP - Wholesale trade | 23.62 | 10.48 | 37.71 | 14.09 |
| GDP - Retail trade | 20.70 | 6.76 | 30.39 | 9.69 |
| GDP - Transportation and warehousing | 65.17 | 1.10 | 67.14 | 1.97 |
| GDP - Information | 6.20 | 9.23 | 17.35 | 11.14 |
| GDP - Finance, insurance, real estate, rental, and leasing | -1.95 | 5.04 | 3.68 | 5.64 |
| GDP - Professional and business services | 27.59 | 30.75 | 64.39 | 36.80 |
| GDP - Educational serv., health care, and social assist. | -0.73 | -0.90 | -2.00 | -1.27 |
| GDP - Arts, entert., recreat., accomm., and food serv. | 56.94 | 1.56 | 62.97 | 6.03 |
| GDP - Government | 0.50 | 18.75 | 19.03 | 18.53 |

In the table we report the adjusted $R^2$, denoted $\bar{R}^2$, of the regression of growth rates of the aggregate IP index and selected aggregated sectoral GDP non-IP output indices on the common factor (column $\bar{R}^2(C)$), the specific HF and LF factors (columns $\bar{R}^2(H)$ and $\bar{R}^2(L)$) only, and the common and frequency-specific factor together(column (3)). The last column displays the difference between the values in the third and first columns. The factors are estimated from the panel of 42 GDP non-IP sectors and 117 IP indices using a mixed frequency factor model with $k^C = k^H = k^L = 1$. The sample periods for the estimation of both factor model and regressions are 1984-2007 (Great Moderation), and 1984-2011.

# 6 Conclusions

Panels with data sampled at different frequencies are the rule rather than the exception in economic applications. We develop a novel approximate factor modeling approach which allows us to estimate factors which are common across all data regardless of their sample frequency, versus factors which are specific to subpanels stratified by sampling frequency. To develop the generic theoretical framework, we cast our analysis into a group factor structure and develop a unified asymptotic theory for the identification of common and group- or frequency-specific factors, for the determination of the number of common and specific factors, for the estimation of loadings and the factors via principal component analysis in a setting with large dimensional data sets, using asymptotic expansions both in the cross-sections and the time series.

There are a plethora of applications to which our theoretical analysis applies. We selected a specific example based on the work of Foerster, Sarte, and Watson (2011) who analyzed the dynamics of comovements across 117 industrial production sectors using both statistical and structural factor models. We revisit their analysis and incorporate the rest, and most dominant part of the US economy, namely the non-IP sectors which we only observe annually.

Despite the generality of our analysis, we can think of many possible extensions, such as models with loadings which change across subperiods (i.e. periodic loadings) or loading which vary stochastically or feature structural breaks. All these extensions are left for future research.

# References

ANDERSON, T. W. (2003): *An Introduction to Multivariate Statistical Analysis*. Wiley-Interscience.

ANDO, T., AND J. BAI (2013): "Multifactor Asset Pricing with a Large Number of Observable Risk Factors and Unobservable Common and Group-specific Factors," *Working Paper*.

ARUOBA, S. B., F. X. DIEBOLD, AND C. SCOTTI (2009): "Real-time Measurement of Business Conditions," *Journal of Business and Economic Statistics*, 27(4), 417–427.

BAI, J. (2003): "Inferential Theory for Factor Models of Large Dimensions," *Econometrica*, 71, 135–171.

BAI, J., AND S. NG (2002): "Determining the Number of Factors in Approximate Factor Models," *Econometrica*, 70(1), 191–221.

——— (2006): "Confidence Intervals for Diffusion Index Forecasts and Inference for Factor-Augmented Regressions," *Econometrica*, 74(4), 1133–1150.

BANBURA, M., AND G. RÜNSTLER (2011): "A Look into the Factor Model Black Box: Publication Lags and the Role of Hard and Soft Data in Forecasting GDP," *International Journal of Forecasting*, 27, 333–346.

BEKAERT, G., R. J. HODRICK, AND X. ZHANG (2009): "International Stock Return Comovements," *Journal of Finance*, 64, 2591–2626.

BREITUNG, J., AND S. EICKMEIER (2014): "Analyzing business and financial cycles using multi-level factor models," *Deutsche Bundesbank, Research Centre, Discussion Papers 11/2014*.

CARVALHO, V. M. (2007): "Aggregate Fluctuations and the Network Structure of Intersectoral Trade.," *Unpublished Manuscript*.

CHEN, P. (2010): "A Grouped Factor Model," *Working Paper*.

——— (2012): "Common Factors and Specific Factors," *Working Paper*.

DEMPSTER, A., N. LAIRD, AND D. RUBIN (1977): "Maximum likelihood estimation from incomplete data," *Journal of the Royal Statistical Society*, 14, 1–38.

FLURY, B. N. (1984): "Common Principal Components in k Groups," *Journal of the American Statistical Association*, 79, 892–898.

FOERSTER, A. T., P.-D. G. SARTE, AND M. W. WATSON (2011): "Sectoral versus Aggregate Shocks: A Structural Factor Analysis of Industrial Production," *Journal of Political Economy*, 119(1), 1–38.

FORNI, M., AND L. REICHLIN (1998): "Let's Get Real: A Factor Analytical Approach to Disaggregated Business Cycle Dynamics," *The Review of Economic Studies*, 65, 453–473.

FRALE, C., AND L. MONTEFORTE (2010): "FaMIDAS: A Mixed Frequency Factor Model with MIDAS Structure," *Government of the Italian Republic (Italy), Ministry of Economy and Finance, Department of the Treasury Working Paper*, 3.

GOYAL, A., C. PÉRIGNON, AND C. VILLA (2008): "How Common are Common Return Factors across the NYSE and Nasdaq?," *Journal of Financial Economics*, 90, 252–271.

GREENWOOD, R., AND D. SCHARFSTEIN (2013): "The Growth of Finance," *Journal of Economic Perspectives*, 27(2), 3–28.

HALLIN, M., AND R. LISKA (2011): "Dynamic factors in the presence of blocks," *Journal of Econometrics*, 163, 29–41.

HORVATH, M. (1998): "Cyclicality and Sectoral Linkages: Aggregate Fluctuations from Independent Sectoral Shocks," *Review of Economic Dynamics*, 1, 781–808.

KING, R. G., AND M. WATSON (2002): "System Reduction and Model Solution Algorithms for Singular Linear Rational Expectations Models," *Computational Economics*, 20, 57–86.

KOSE, A. M., C. OTROK, AND C. H. WHITEMAN (2008): "Understanding the Evolution of World Business Cycles," *Journal of International Economics*, 75, 110–130.

KRZANOWSKI, W. (1979): "Between-groups Comparison of Principal Components," *Journal of the American Statistical Association*, 74, 703–707.

LONG, J. B. J., AND C. I. PLOSSER (1983): "Real Business Cycles," *Journal of Political Economy*, 91, 39–69.

MAGNUS, J. R., AND H. NEUDECKER (2007): *Matrix Differential Calculus with Applications in Statistics and Econometrics*. John Wiley and Sons: Chichester/New York.

MARCELLINO, M., AND C. SCHUMACHER (2010): "Factor MIDAS for Nowcasting and Forecasting with Ragged-Edge Data: A Model Comparison for German GDP," *Oxford Bulletin of Economics and Statistics*, 72(4), 518–550.

MARIANO, R. S., AND Y. MURASAWA (2003): "A New Coincident Index of Business Cycles Based on Monthly and Quarterly Series," *Journal of Applied Econometrics*, 18(4), 427–443.

MOENCH, E., AND S. NG (2011): "A Hierarchical Factor Analysis of US Housing Market Dynamics," *The Econometrics Journal*, 14, C1–C24.

MOENCH, E., S. NG, AND S. POTTER (2013): "Dynamic Hierarchical Factor Models," *Review of Economics and Statistics*, 95, 1811–1817.

NUNES, L. C. (2005): "Nowcasting Quarterly GDP Growth in a Monthly Coincident Indicator Model," *Journal of Forecasting*, 24(8), 575–592.

SCHOTT, J. R. (1999): "Partial Common Principal Component Subspaces," *Biometrika*, 86, 899–908.

——— (2005): *Matrix Analysis for Statistic*. Wiley, New York, 2 edn.

STOCK, J. H., AND M. W. WATSON (2002a): "Forecasting using principal components from a large number of predictors," *Journal of the American Statistical Association*, 97, 1167 – 1179.

——— (2002b): "Macroeconomic Forecasting Using Diffusion Indexes," *Journal of Business and Economic Statistics*, 20, 147–162.

——— (2010): "Dynamic Factor Models," in *Oxford Handbook of Economic Forecasting*, ed. by E. Graham, C. Granger, and A. Timmerman, pp. 87–115. Michael P. Clements and David F. Hendry (eds), Oxford University Press, Amsterdam.

WANG, P. (2012): "Large Dimensional Factor Models with a Multi-Level Factor Structure: Identifcation, Estimation, and Inference," *Working Paper*.

# Technical Appendices

## A.1 Identification: stock sampling

In the case of stock sampling, the low frequency observations of $x^{L*}_{m,t}$ in the factor model (2.1) are the values of $x^{L*}_{M,t}$, i.e. $x^L_t = y^{L*}_{M,t}$. Then, the model for the observable variables becomes:

$$\begin{aligned} x^H_{m,t} &= \Lambda_{HC} g^C_{m,t} + \Lambda_H g^H_{m,t} + e^H_{m,t}, \quad m = 1, ..., M, \\ x^L_t &= \Lambda_{LC} g^C_{M,t} + \Lambda_L g^L_{M,t} + e^L_{M,t}. \end{aligned}$$

We stack the observations $x_{m,t}$ and $y_t$ of the last high frequency subperiod and write:

$$\begin{bmatrix} x^H_t \\ x^L_t \end{bmatrix} = \begin{bmatrix} \Lambda_{HC} & \Lambda_H & 0 \\ \Lambda_{LC} & 0 & \Lambda_L \end{bmatrix} \begin{bmatrix} g^C_{M,t} \\ g^H_{M,t} \\ g^L_{M,t} \end{bmatrix} + \begin{bmatrix} e^H_{M,t} \\ e^L_{M,t} \end{bmatrix}. \tag{A.1}$$

This equation corresponds to a group factor model, with common factor $g^C_{M,t}$ and "group-specific" factors $g^H_{M,t}$, $g^L_{M,t}$. Therefore, the factor values $g^C_{M,t}$, $f^H_{M,t}$, $f^L_{M,t}$, and the factor loadings $\Lambda_{HC}$, $\Lambda_{LC}$, $\Lambda_H$, $\Lambda_L$, are identifiable up to a sign as proved in Section 2.1 (see also results in e.g. Schott (1999), Wang (2012), Chen (2010, 2012)).

Once the factor loadings are identified from (A.1), the values of the common and high frequency factors for subperiods $m = 1, ..., M-1$ are identifiable by cross-sectional regression of the high frequency data on loadings $\Lambda_{HC}$ and $\Lambda_H$ in (2.1). More precisely, $g^C_{m,t}$ and $g^H_{m,t}$ are identified by regressing $x^{Hi}_{m,t}$ on $\lambda_{HC,i}$ and $\lambda_{H,i}$ across $i = 1, 2, ..., N_H$, for any $m = 1, ..., M-1$ and any $t$. To summarize, with stock sampling, we can identify the common factor $g^C_{m,t}$ and the high frequency factor $g^H_{m,t}$ at all high frequency subperiods. We cannot estimate $g^L_{m,t}$, for $m < M$, as only $g^L_{M,t}$ is identified by the last paired panel data set consisting of $x^H_{M,t}$ combined with $x^L_t$. This is not surprising, since we have no HF observation available for the LF process.

## A.2 Assumptions: group factor model

Let $\|A\| = \sqrt{tr(A'A)}$ denote the Frobenius norm of matrix $A$. Let $k^F = k^c + k^s_1 + k^s_2$, and define the $k_F$-dimensional vector of factors: $F_t = [\ f^{c\,\prime}_t, \quad f^{s\,\prime}_{1,t}, \quad f^{s\,\prime}_{2,t}\ ]'$, and the $(T, k_F)$ matrix $F = [\ F'_1, \ ..., \ F'_T\ ]'$. We make the following assumptions:

**Assumption A.1.** *The unobservable factor process is such that $F'F/T = \Sigma_F + O_p(1/\sqrt{T})$ as $T \to \infty$, where $\Sigma_F$ is a positive definite $(k^F \times k^F)$ matrix defined as:*

$$\Sigma_F = \begin{bmatrix} I_{k^c} & 0 & 0 \\ 0 & I_{k^s_1} & \Phi \\ 0 & \Phi' & I_{k^s_2} \end{bmatrix}. \tag{A.2}$$

**Assumption A.2.** *The loadings matrices $\Lambda_1 = \begin{bmatrix} \Lambda^c_1 \vdots \Lambda^s_1 \end{bmatrix}$ and $\Lambda_2 = \begin{bmatrix} \Lambda^c_2 \vdots \Lambda^s_2 \end{bmatrix}$ are full column-rank, for $N_1, N_2$ large enough. The loadings $\lambda_{j,i}$ are such that:*

$$\frac{\Lambda'_j \Lambda_j}{N_j} = \Sigma_{\Lambda,j} + O\left(\frac{1}{\sqrt{N_j}}\right), \qquad j = 1, 2, \tag{A.3}$$

*where $\Sigma_{\Lambda,j} := \lim_{N_j \to \infty} \left(\frac{\Lambda'_j \Lambda_j}{N_j}\right)$ is a p.d. $(k_j, k_j)$ matrix, for $j = 1, 2$.*

**Assumption A.3.** *The error terms* $(\varepsilon_{1,it} \;\; \varepsilon_{2,it})'$ *are weakly dependent across* $i$ *and* $t$*, and such that* $E[\varepsilon_{j,it}] = 0$.

**Assumption A.4.** *There exists a constant* $C_\varepsilon$ *such that* $E[\varepsilon^4_{j,it}] \le C_\varepsilon$ *for all* $j$*,* $i$ *and* $t$.

**Assumption A.5.** *a) The variables* $F_t$ *and* $\varepsilon_{j,is}$ *are independent, for all* $i, j, t$ *and* $s$.
*b) The processes* $\{\varepsilon_{j,it}\}$ *are stationary, for all* $j, i$.
*c) The process* $\{F_t\}$ *is stationary and weakly dependent over time.*
*d) For each* $j$ *and* $t$*, as* $N_j \to \infty$*, it holds:*

$$\frac{1}{\sqrt{N_j}} \sum_{i=1}^{N_j} \lambda_{j,i}\varepsilon_{j,it} \overset{d}{\to} N(0, \Omega_j), \tag{A.4}$$

*where* $\Omega_j = \lim_{N_j \to \infty} \frac{1}{N_j} \sum_{i=1}^{N_j} \sum_{\ell=1}^{N_j} \lambda_{j,i}\lambda'_{j,\ell} E[\varepsilon_{j,it}\varepsilon_{j,\ell t}]$.

**Assumption A.6.** *The asymptotic analysis is for* $N_1, N_2, T \to \infty$ *such that* $N_2 \le N_1$*,* $T/N_2 = o(1)$*,* $\sqrt{N_1}/T = o(1)$.

The following Assumption A.7 simplifies the derivation of the feasible asymptotic distribution of the statistic used to test the dimension of the common factor space $k^c$.

**Assumption A.7.** *The error terms* $\varepsilon_{j,it}$ *are uncorrelated across* $j$*,* $i$ *and* $t$*, and* $\varepsilon_{j,it} \sim (0, \gamma_{j,ii})$.

Assumption A.7 is a stronger condition than Assumptions A.3 and A.5 *b)*. Moreover, under Assumption A.7, the matrix $\Omega_j$ in Assumption A.5 *d)* simplifies to $\Omega_j = \lim_{N_j \to \infty} \frac{1}{N_j} \sum_{i=1}^{N_j} \lambda_{j,i}\lambda'_{j,i}\gamma_{j,ii}$ .

# A.3 Assumptions: mixed frequency factor model

Let $\lambda'_{1,i}$ be the $i$-th row of the $(N_H, k^C + k^H)$ matrix $\Lambda_1 = [\ \Lambda_{HC} \vdots \Lambda_H\ ]$. We make the following assumption:

**Assumption A.8.** *The variables* $\lambda_{1,i}$ *and* $e^{i,H}_{m,t}$ *are such that:*

$$\frac{1}{\sqrt{N_H}} \sum_{i=1}^{N_H} \lambda_{1,i} e^{i,H}_{m,t} \overset{d}{\to} N(0, \Omega^*_{\Lambda,m}), \tag{A.5}$$

*where*

$$\Omega^*_{\Lambda,m} = \lim_{N_H \to \infty} \frac{1}{N_H} \sum_{i=1}^{N_H} \sum_{\ell=1}^{N_H} \lambda_{1,i}\lambda'_{1,\ell} Cov(e^{i,H}_{m,t}, e^{\ell,H}_{m,t}), \qquad m = 1, ..., M. \tag{A.6}$$

# A.4 Proofs of Theorems and Lemmas

## A.4.1 Proof of Proposition 1

By replacing equation (2.7) into model (2.4), we get

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} \Lambda^c_1 A_{11} + \Lambda^s_1 A_{21} & \Lambda^c_1 A_{12} + \Lambda^s_1 A_{22} & \Lambda^c_1 A_{13} + \Lambda^s_1 A_{23} \\ \Lambda^c_2 A_{11} + \Lambda^s_2 A_{31} & \Lambda^c_2 A_{12} + \Lambda^s_2 A_{32} & \Lambda^c_2 A_{13} + \Lambda^s_2 A_{33} \end{bmatrix} \begin{bmatrix} \tilde{f}^c_t \\ \tilde{f}^s_{1,t} \\ \tilde{f}^s_{2,t} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}. \tag{A.7}$$

This factor model satisfies the restrictions in the loading matrix appearing in equation (2.4) if, and only if,

$$\Lambda_1^c A_{13} + \Lambda_1^s A_{23} = 0, \tag{A.8}$$
$$\Lambda_2^c A_{12} + \Lambda_2^s A_{32} = 0. \tag{A.9}$$

Equations (A.8) and (A.9) can be written as linear homogeneous systems of equations for the elements of matrices $[A_{13}' \; A_{23}']'$ and $[A_{12}' \; A_{32}']'$:

$$\left[\Lambda_1^c \vdots \Lambda_1^s\right] \begin{bmatrix} A_{13} \\ A_{23} \end{bmatrix} = 0, \text{ and } \left[\Lambda_1^c \vdots \Lambda_2^s\right] \begin{bmatrix} A_{12} \\ A_{32} \end{bmatrix} = 0.$$

Since $\left[\Lambda_1^c \vdots \Lambda_1^s\right]$ and $\left[\Lambda_2^c \vdots \Lambda_2^s\right]$ are full column rank, it follows that

$$A_{13} = 0, \; A_{23} = 0, \tag{A.10}$$
$$A_{12} = 0, \; A_{32} = 0. \tag{A.11}$$

Therefore, the transformation of the factors that is compatible with the restrictions on the loading matrix in equation (2.4) is:

$$\begin{bmatrix} f_t^c \\ f_{1,t}^s \\ f_{2,t}^s \end{bmatrix} = \begin{bmatrix} A_{11} & 0 & 0 \\ A_{21} & A_{22} & 0 \\ A_{31} & 0 & A_{33} \end{bmatrix} \begin{bmatrix} \tilde{f}_t^c \\ \tilde{f}_{1,t}^s \\ \tilde{f}_{2,t}^s \end{bmatrix}.$$

We can invert this transformation and write:

$$\begin{aligned} \tilde{f}_t^c &= A_{11}^{-1} f_t^c, \\ \tilde{f}_{1,t}^s &= A_{22}^{-1} f_{1,t}^s - A_{22}^{-1} A_{21} A_{11}^{-1} f_t^c, \\ \tilde{f}_{2,t}^s &= A_{33}^{-1} f_{2,t}^s - A_{33}^{-1} A_{31} A_{11}^{-1} f_t^c. \end{aligned}$$

The transformed factors satisfy the normalization restrictions in (2.6) if, and only if,

$$Cov(\tilde{f}_{1,t}^s, \; \tilde{f}_t^c) = -A_{22}^{-1} A_{21} A_{11}^{-1} (A_{11}^{-1})' = 0, \tag{A.12}$$
$$Cov(\tilde{f}_{2,t}^s, \; \tilde{f}_t^c) = -A_{33}^{-1} A_{31} A_{11}^{-1} (A_{11}^{-1})' = 0, \tag{A.13}$$
$$V(\tilde{f}_t^c) = A_{11}^{-1} (A_{11}^{-1})' = I_{k^c}, \tag{A.14}$$
$$V(\tilde{f}_{1,t}^s) = A_{22}^{-1} (A_{22}^{-1})' + A_{22}^{-1} A_{21} A_{11}^{-1} (A_{11}^{-1})' A_{21}' (A_{22}^{-1})' = I_{k_1^s}, \tag{A.15}$$
$$V(\tilde{f}_{2,t}^s) = A_{33}^{-1} (A_{33}^{-1})' + A_{33}^{-1} A_{31} A_{11}^{-1} (A_{11}^{-1})' A_{31}' (A_{33}^{-1})' = I_{k_2^s}, \tag{A.16}$$

Since the matrices $A_{11}$, $A_{22}$ and $A_{33}$ are nonsingular, equations (A.12) and (A.13) imply

$$A_{21} = 0, \text{ and } A_{31} = 0. \tag{A.17}$$

Then, from equations (A.14) - (A.16), we get that matrices $A_{11}$, $A_{22}$ and $A_{33}$ are orthogonal.

*Q.E.D.*

### A.4.2 Proof of Proposition 2

From equation (2.6) we have

$$R = \begin{pmatrix} I_{k^c} & 0 \\ 0 & \Phi\Phi' \end{pmatrix} \text{ and } R^* = \begin{pmatrix} I_{k^c} & 0 \\ 0 & \Phi'\Phi \end{pmatrix}.$$

Matrix $R$ is block diagonal, and the upper-left block $I_{k^c}$ has eigenvalue 1 with multiplicity $k^c$. The associated eigenspace is $\{(\xi', 0')', \xi \in \mathbb{R}^{k^c}\}$. The lower-right block $\Phi\Phi'$ is a positive semi-definite matrix, and its largest eigenvalue is $\tilde{\rho}^2$, where $\tilde{\rho}^2 = \sup\{\xi_1'\Phi\Phi'\xi_1 : \; \xi_1 \in \mathbb{R}^{k_1^s}, \|\xi_1\| = 1\} < 1$ is the first squared canonical correlation of vectors $f_{1,t}^s$ and $f_{2,t}^s$. Therefore, we deduce that the largest eigenvalue of matrix $R$ is equal to 1, with multiplicity $k^c$, and the associated eigenspace, denoted by $\mathcal{E}_c$, is spanned by vectors $(\xi', 0')'$, with $\xi \in \mathbb{R}^{k^c}$. Let $S_1$ be an orthogonal $(k^c, k^c)$ matrix, then the columns of the $(k_1, k^c)$ matrix

$$W_1 = \begin{pmatrix} S_1 \\ 0_{k_1^s \times k^c} \end{pmatrix}$$

are an orthonormal basis of the eigenspace $\mathcal{E}_c$. We have:

$$W_1' h_{1,t} = S_1' f_t^c. \tag{A.18}$$

Analogous arguments allow to show that the largest eigenvalue of matrix $R^*$ is equal to 1, with multiplicity $k^c$ and that the associated eigenspace , denoted by $\mathcal{E}_c^*$, is spanned by vectors $(\xi^{*\prime}, 0')'$, with $\xi^* \in \mathbb{R}^{k^c}$. Let $S_2$ be an orthogonal $(k^c, k^c)$ matrix. Then, the columns of the $(k_2, k^c)$ matrix

$$W_2 = \begin{pmatrix} S_2 \\ 0_{k_2^s \times k^c} \end{pmatrix}$$

are an orthonormal basis of the eigenspace $\mathcal{E}_c^*$. We have:

$$W_2' h_{2,t} = S_2' f_t^c, \tag{A.19}$$

which yields parts *i)* and *ii)*.

When there is no common factor, the matrix $R$ becomes $R = \Phi\Phi'$, and matrix $R^*$ becomes $R^* = \Phi'\Phi$. By the above arguments, the largest eigenvalue of matrix $R$, which is equal to the largest eigenvalue of matrix $R^*$, is not larger than $\tilde{\rho}^2$, where $\tilde{\rho}^2 < 1$ is the first squared canonical correlation between the two group-specific factors. This yields part *iii)*.

Finally, we prove part *iv)*. We showed that the lower-right block $\Phi\Phi'$ of matrix $R$ is a positive semi-definite matrix and all its $k_1^s = k_1 - k^c$ eigenvalues are strictly smaller than one. These are also eigenvalues of matrix $R$. Let us denote the space spanned by the associated $k_1^s$ eigenvectors of matrix $R$ by $\mathcal{E}_{s,1}$. This space is spanned by vectors $(0', \tilde{\xi}')'$ with $\tilde{\xi} \in \mathbb{R}^{k_1^s}$. We note that, by construction, the vectors $(0', \tilde{\xi}')'$ are linearly independent of the vectors $(\xi', 0')'$ spanning the eigenspace $\mathcal{E}_c$. Let $Q_1$ be an orthogonal $(k_1^s, k_1^s)$ matrix, then the columns of matrix

$$W_1^s = \begin{pmatrix} 0_{k^c \times k_1^s} \\ Q_1 \end{pmatrix}$$

are an orthonormal basis of the eigenspace $\mathcal{E}_{s,1}$. We have:

$$W_1^{s\prime} h_{1,t} = Q_1' f_{1,t}^s. \tag{A.20}$$

Analogously, we have that the lower-right block $\Phi'\Phi$ of matrix $R^*$ is a positive semi-definite matrix and all its $k_2^s = k_2 - k^c$ eigenvalues are strictly smaller than one. These are also eigenvalues of matrix $R^*$. Let us denote the space spanned by the associated $k_2^s$ eigenvectors of matrix $R^*$ by $\mathcal{E}_{s,2}$. This space is spanned by vectors $(0', \tilde{\xi}^{*\prime})'$ with $\tilde{\xi}^* \in \mathbb{R}^{k_2^s}$. We note that, by construction, the vectors $(0', \tilde{\xi}^{*\prime})'$ are linearly independent of the vectors $(\xi^{*\prime}, 0')'$ spanning the eigenspace $\mathcal{E}_c^*$. Let $Q_2$ be an orthogonal $(k_2^s, k_2^s)$ matrix, then the columns of matrix

$$W_2^s = \begin{pmatrix} 0_{k^c \times k_2^s} \\ Q_2 \end{pmatrix}$$

are an orthonormal basis of the eigenspace $\mathcal{E}_{s,2}$. We have:

$$W_2^{s\prime} h_{2,t} = Q_2' f_{2,t}^s. \tag{A.21}$$

*Q.E.D.*

# A.5 Proof of Theorem 3

## A.5.1 Asymptotic expansion of $\hat{R}$

In order to derive the asymptotic distribution of the test statistic $\hat{\xi}(k^c)$ defined in equation (3.5), and common factor estimator introduced in Definition 2, we consider a perturbation of matrix $\hat{R}$ and its eigenvalues and eigenvectors. More precisely, the perturbation of the eigenvalues will allow us to derive the asymptotic distribution of the test statistic $\hat{\xi}(k^c)$, while the perturbation of the eigenvectors will allow us to derive the asymptotic distribution of the common factor estimator.

The canonical correlations and the canonical directions are invariant to one-to-one transformations of the vectors $\hat{h}_{1,t}$ and $\hat{h}_{2,t}$ (see, among others, Anderson (2003)). Therefore, without loss of generality, for the asymptotic analysis of the estimator of the dimension of the common factor space statistic $\hat{\xi}(k^c)$, we can set $\hat{\mathcal{H}}_j = I_{k_j}$, $j = 1, 2$, in approximation (4.1). Moreover, under Assumption A.6 the bias term is negligible, and we get:

$$\hat{h}_{j,t} \simeq h_{j,t} + \frac{1}{\sqrt{N_j}} u_{j,t}, \quad j = 1, 2. \tag{A.22}$$

By using approximation (A.22), and $N_2 = N$, $N_1 = N/{\mu_N}^2$, we have:

$$\begin{aligned}
\hat{V}_{12} &= \frac{1}{T}\sum_{t=1}^{T} \hat{h}_{1,t}\hat{h}'_{2,t} \\
&\simeq \frac{1}{T}\sum_{t=1}^{T} \left(h_{1,t} + \frac{1}{\sqrt{N}}\mu_N u_{1,t}\right)\left(h_{2,t} + \frac{1}{\sqrt{N}} u_{2,t}\right)' \\
&= \tilde{V}_{12} + \hat{X}_{12},
\end{aligned}$$

where:

$$\begin{aligned}
\tilde{V}_{12} &= \frac{1}{T}\sum_{t=1}^{T} h_{1,t}h'_{2,t}, \\
\hat{X}_{12} &= \frac{1}{T\sqrt{N}}\sum_{t=1}^{T}(h_{1,t}u'_{2,t} + \mu_N u_{1,t}h'_{2,t}) + \frac{\mu_N}{TN}\sum_{t=1}^{T} u_{1,t}u'_{2,t}.
\end{aligned} \tag{A.23}$$

Similarly:

$$\begin{aligned}
\hat{V}_{jj} &= \frac{1}{T}\sum_{t=1}^{T} \hat{h}_{j,t}\hat{h}'_{j,t} \\
&\simeq \frac{1}{T}\sum_{t=1}^{T} \left(h_{j,t} + \frac{1}{\sqrt{N_j}} u_{j,t}\right)\left(h_{j,t} + \frac{1}{\sqrt{N_j}} u_{j,t}\right)' \\
&= \tilde{V}_{jj} + \hat{X}_{jj} && \text{(A.24)} \\
&= \tilde{V}_{jj}\left(Id + \tilde{V}_{jj}^{-1}\hat{X}_{jj}\right), \quad j = 1, 2, && \text{(A.25)}
\end{aligned}$$

where:

$$\tilde{V}_{jj} = \frac{1}{T}\sum_{t=1}^{T} h_{j,t}h'_{j,t}, \quad j = 1, 2, \tag{A.26}$$

$$\hat{X}_{11} = \frac{\mu_N}{T\sqrt{N}}\sum_{t=1}^{T}(h_{1,t}u'_{1,t} + u_{1,t}h'_{1,t}) + \frac{\mu_N^2}{TN}\sum_{t=1}^{T} u_{1,t}u'_{1,t}, \tag{A.27}$$

$$\hat{X}_{22} = \frac{1}{T\sqrt{N}}\sum_{t=1}^{T}(h_{2,t}u'_{2,t} + u_{2,t}h'_{2,t}) + \frac{1}{TN}\sum_{t=1}^{T} u_{2,t}u'_{2,t}. \tag{A.28}$$

Therefore, we get:

$$\hat{R} \simeq \left(Id + \tilde{V}_{11}^{-1}\hat{X}_{11}\right)^{-1}\tilde{V}_{11}^{-1}\left(\tilde{V}_{12} + \hat{X}_{12}\right)\left(Id + \tilde{V}_{22}^{-1}\hat{X}_{22}\right)^{-1}\tilde{V}_{22}^{-1}\left(\tilde{V}_{21} + \hat{X}_{21}\right).$$

Let us expand $\hat{R}$ at first order in the $\hat{X}_{j,k} = O_p\left(\frac{1}{\sqrt{NT}}\right)$. By using $(Id + X)^{-1} \simeq Id - X$ for $X \simeq 0$, we have:

$$\begin{aligned}
\hat{R} &\simeq \left(Id - \tilde{V}_{11}^{-1}\hat{X}_{11}\right)\tilde{V}_{11}^{-1}\left(\tilde{V}_{12} + \hat{X}_{12}\right)\left(Id - \tilde{V}_{22}^{-1}\hat{X}_{22}\right)\tilde{V}_{22}^{-1}\left(\tilde{V}_{21} + \hat{X}_{21}\right) \\
&\simeq \tilde{V}_{11}^{-1}\tilde{V}_{12}\tilde{V}_{22}^{-1}\tilde{V}_{21} \\
&\quad -\tilde{V}_{11}^{-1}\hat{X}_{11}\tilde{V}_{11}^{-1}\tilde{V}_{12}\tilde{V}_{22}^{-1}\tilde{V}_{21} + \tilde{V}_{11}^{-1}\hat{X}_{12}\tilde{V}_{22}^{-1}\tilde{V}_{21} - \tilde{V}_{11}^{-1}\tilde{V}_{12}\tilde{V}_{22}^{-1}\hat{X}_{22}\tilde{V}_{22}^{-1}\tilde{V}_{21} + \tilde{V}_{11}^{-1}\tilde{V}_{12}\tilde{V}_{22}^{-1}\hat{X}_{21}.
\end{aligned}$$

Defining the following quantities:

$$\tilde{A} = \tilde{V}_{11}^{-1}\tilde{V}_{12}, \tag{A.29}$$

$$\tilde{B} = \tilde{V}_{22}^{-1}\tilde{V}_{21}, \tag{A.30}$$

$$\tilde{R} = \tilde{V}_{11}^{-1}\tilde{V}_{12}\tilde{V}_{22}^{-1}\tilde{V}_{21} = \tilde{A}\tilde{B}, \tag{A.31}$$

$$\hat{\Psi}^* = -\hat{X}_{11}\tilde{R} + \hat{X}_{12}\tilde{B} - \tilde{B}'\hat{X}_{22}\tilde{B} + \tilde{B}'\hat{X}_{21}, \tag{A.32}$$

$$\hat{\Psi} = \tilde{V}_{11}^{-1}\hat{\Psi}^*, \tag{A.33}$$

we get the asymptotic expansion of matrix $\hat{R}$:

$$\hat{R} = \tilde{R} + \hat{\Psi} + O_p\left(\frac{1}{NT}\right). \tag{A.34}$$

### A.5.2 Matrix $\tilde{R}$ and its eigenvalues and eigenvectors

Let us now compute matrix $\tilde{R}$ and its eigenvalues, that are $\tilde{\rho}_1^2, ..., \tilde{\rho}_{k_1}^2$, i.e. the squared sample canonical correlations of vectors $h_{1,t}$ and $h_{2,t}$, under the null hypothesis of $k^c > 0$ common factors among the 2 groups of observables. Since the vectors $h_{1,t}$ and $h_{2,t}$ have a common component of dimension $k^c$, we know that $\tilde{\rho}_1 = ... = \tilde{\rho}_{k^c} = 1$ a.s.. Using the notation:

$$
\begin{aligned}
\tilde{\Sigma}_{cc} &= \frac{1}{T}\sum_{t=1}^{T} f_t^c f_t^{c\prime}\,, \\
\tilde{\Sigma}_{cj} &= \frac{1}{T}\sum_{t=1}^{T} f_t^c f_{j,t}^{s\prime}\,, \quad \tilde{\Sigma}_{jc} \;=\; \tilde{\Sigma}_{cj}'\,, \quad j=1,2, \\
\tilde{\Sigma}_{jj} &= \frac{1}{T}\sum_{t=1}^{T} f_{j,t}^s f_{j,t}^{s\prime}\,, \quad j=1,2, \\
\tilde{\Sigma}_{12} &= \frac{1}{T}\sum_{t=1}^{T} f_{1,t}^s f_{2,t}^{s\prime}\,,
\end{aligned}
$$

we can write matrices $\tilde{V}_{jj}$, with $j = 1, 2$, and $\tilde{V}_{12}$ as:

$$
\tilde{V}_{jj} = \begin{pmatrix} \tilde{\Sigma}_{cc} & \tilde{\Sigma}_{c,j} \\ \tilde{\Sigma}_{j,c} & \tilde{\Sigma}_{jj} \end{pmatrix}, \quad j=1,2, \tag{A.35}
$$

$$
\tilde{V}_{12} = \begin{pmatrix} \tilde{\Sigma}_{cc} & \tilde{\Sigma}_{c,2} \\ \tilde{\Sigma}_{1,c} & \tilde{\Sigma}_{12} \end{pmatrix} = \tilde{V}_{21}'. \tag{A.36}
$$

By matrix algebra we get:

$$
\tilde{V}_{11}^{-1} = \begin{bmatrix} \Sigma_*^{-1} & -\tilde{\Sigma}_*^{-1}\tilde{\Sigma}_{c1}\tilde{\Sigma}_{11}^{-1} \\ -\tilde{\Sigma}_{11}^{-1}\tilde{\Sigma}_{1c}\tilde{\Sigma}_*^{-1} & \tilde{\Sigma}_{11}^{-1} + \tilde{\Sigma}_{11}^{-1}\tilde{\Sigma}_{1c}\tilde{\Sigma}_*^{-1}\tilde{\Sigma}_{c1}\tilde{\Sigma}_{11}^{-1} \end{bmatrix}, \tag{A.37}
$$

where

$$
\tilde{\Sigma}_* = \tilde{\Sigma}_{cc} - \tilde{\Sigma}_{c1}\tilde{\Sigma}_{11}^{-1}\tilde{\Sigma}_{1c}. \tag{A.38}
$$

From assumption A.1, we have:

$$
\tilde{\Sigma}_{c1} = O_p(1/\sqrt{T}), \tag{A.39}
$$

$$
\tilde{\Sigma}_{cc} = I_{k^c} + O_p(1/\sqrt{T}), \tag{A.40}
$$

$$
\tilde{\Sigma}_{11} = I_{k_1^s} + O_p(1/\sqrt{T}), \tag{A.41}
$$

$$
\tilde{\Sigma}_{22} = I_{k_2^s} + O_p(1/\sqrt{T}), \tag{A.42}
$$

$$
\tilde{\Sigma}_{12} = \Phi + O_p(1/\sqrt{T}), \tag{A.43}
$$

which imply:

$$
\tilde{\Sigma}_* = \tilde{\Sigma}_{cc} + O_p(1/T), \tag{A.44}
$$

$$
\tilde{\Sigma}_*^{-1} = \tilde{\Sigma}_{cc}^{-1} + O_p(1/T), \tag{A.45}
$$

$$
\begin{aligned}
-\tilde{\Sigma}_*^{-1}\tilde{\Sigma}_{c1}\tilde{\Sigma}_{11}^{-1} &= -\tilde{\Sigma}_{cc}^{-1}\tilde{\Sigma}_{c1}\tilde{\Sigma}_{11}^{-1} + O_p(1/T), \\
&= -\tilde{\Sigma}_{c1} + O_p(1/T),
\end{aligned} \tag{A.46}
$$

$$
\tilde{\Sigma}_{11}^{-1}\tilde{\Sigma}_{1c}\tilde{\Sigma}_*^{-1}\tilde{\Sigma}_{c1}\tilde{\Sigma}_{11}^{-1} = O_p(1/T). \tag{A.47}
$$

Substituting results (A.44) - (A.47) into equation (A.37) we get:

$$\tilde{V}_{11}^{-1} = \begin{bmatrix} \tilde{\Sigma}_{cc}^{-1} & -\tilde{\Sigma}_{c1} \\ -\tilde{\Sigma}_{1c} & \tilde{\Sigma}_{11}^{-1} \end{bmatrix} + O_p(1/T). \tag{A.48}$$

Equation (A.37) allows to compute $\tilde{A}$ :

$$\begin{aligned} \tilde{A} &= \tilde{V}_{11}^{-1}\tilde{V}_{12} \\ &= \begin{bmatrix} \tilde{\Sigma}_{*}^{-1} & -\tilde{\Sigma}_{*}^{-1}\tilde{\Sigma}_{c1}\tilde{\Sigma}_{11}^{-1} \\ -\tilde{\Sigma}_{11}^{-1}\tilde{\Sigma}_{1c}\tilde{\Sigma}_{*}^{-1} & \tilde{\Sigma}_{11}^{-1} + \tilde{\Sigma}_{11}^{-1}\tilde{\Sigma}_{1c}\tilde{\Sigma}_{*}^{-1}\tilde{\Sigma}_{c1}\tilde{\Sigma}_{11}^{-1} \end{bmatrix} \begin{bmatrix} \tilde{\Sigma}_{cc} & \tilde{\Sigma}_{c2} \\ \tilde{\Sigma}_{1c} & \tilde{\Sigma}_{12} \end{bmatrix} \\ &= \begin{bmatrix} I_{k^c} & \tilde{A}_{cs} \\ 0 & \tilde{A}_{ss} \end{bmatrix}, \end{aligned} \tag{A.49}$$

where:

$$\tilde{A}_{cs} = \tilde{\Sigma}_{*}^{-1}\tilde{\Sigma}_{c2} - \tilde{\Sigma}_{*}^{-1}\tilde{\Sigma}_{c1}\tilde{\Sigma}_{11}^{-1}\tilde{\Sigma}_{12} = O_p\left(\frac{1}{\sqrt{T}}\right), \tag{A.50}$$

$$\begin{aligned} \tilde{A}_{ss} &= -\tilde{\Sigma}_{11}^{-1}\tilde{\Sigma}_{1c}\tilde{\Sigma}_{*}^{-1}\tilde{\Sigma}_{c2} + \tilde{\Sigma}_{11}^{-1}\tilde{\Sigma}_{12} + \tilde{\Sigma}_{11}^{-1}\tilde{\Sigma}_{1c}\tilde{\Sigma}_{*}^{-1}\tilde{\Sigma}_{c1}\tilde{\Sigma}_{11}^{-1}\tilde{\Sigma}_{12} \\ &= \tilde{\Sigma}_{11}^{-1}\tilde{\Sigma}_{12} + O_p\left(\frac{1}{T}\right) \\ &= \Phi + O_p\left(\frac{1}{\sqrt{T}}\right). \end{aligned} \tag{A.51}$$

**REMARK 1.** *Matrices $\tilde{V}_{12}$ and $\tilde{V}_{11}$ have the same first $k^c$ columns, therefore also matrices $\tilde{V}_{11}^{-1}\tilde{V}_{12}$ and $\tilde{V}_{11}^{-1}\tilde{V}_{11} = I_{k_1}$ have the first $k^c$ columns, which implies:*

$$\tilde{V}_{11}^{-1}\tilde{V}_{12} = \begin{bmatrix} I_{k^c} & * \\ 0 & * \end{bmatrix}.$$

Let us compute:

$$\tilde{V}_{22}^{-1} = \begin{bmatrix} \tilde{\Sigma}_{*2}^{-1} & -\tilde{\Sigma}_{*2}^{-1}\tilde{\Sigma}_{c2}\tilde{\Sigma}_{22}^{-1} \\ -\tilde{\Sigma}_{22}^{-1}\tilde{\Sigma}_{2c}\tilde{\Sigma}_{*2}^{-1} & \tilde{\Sigma}_{22}^{-1} + \tilde{\Sigma}_{22}^{-1}\tilde{\Sigma}_{2c}\tilde{\Sigma}_{*2}^{-1}\tilde{\Sigma}_{c2}\tilde{\Sigma}_{22}^{-1} \end{bmatrix}, \tag{A.52}$$

where

$$\tilde{\Sigma}_{*2} = \tilde{\Sigma}_{cc} - \tilde{\Sigma}_{c2}\tilde{\Sigma}_{22}^{-1}\tilde{\Sigma}_{2c}.$$

Equation (A.52) allows to compute $\tilde{B}$ :

$$\begin{aligned} \tilde{B} &= \tilde{V}_{22}^{-1}\tilde{V}_{21} \\ &= \begin{bmatrix} \tilde{\Sigma}_{*2}^{-1} & -\tilde{\Sigma}_{*2}^{-1}\tilde{\Sigma}_{c2}\tilde{\Sigma}_{22}^{-1} \\ -\tilde{\Sigma}_{22}^{-1}\tilde{\Sigma}_{2c}\tilde{\Sigma}_{*2}^{-1} & \tilde{\Sigma}_{22}^{-1} + \tilde{\Sigma}_{22}^{-1}\tilde{\Sigma}_{2c}\tilde{\Sigma}_{*2}^{-1}\tilde{\Sigma}_{c2}\tilde{\Sigma}_{22}^{-1} \end{bmatrix} \begin{bmatrix} \tilde{\Sigma}_{cc} & \tilde{\Sigma}_{c1} \\ \tilde{\Sigma}_{2c} & \tilde{\Sigma}_{21} \end{bmatrix} \\ &= \begin{bmatrix} I_{k^c} & \tilde{B}_{cs} \\ 0 & \tilde{B}_{ss} \end{bmatrix}, \end{aligned} \tag{A.53}$$

where:

$$\tilde{B}_{cs} = \tilde{\Sigma}_{*2}^{-1}\tilde{\Sigma}_{c1} - \tilde{\Sigma}_{*2}^{-1}\tilde{\Sigma}_{c2}\tilde{\Sigma}_{22}^{-1}\tilde{\Sigma}_{21} = O_p\left(\frac{1}{\sqrt{T}}\right), \tag{A.54}$$

$$\begin{aligned}\tilde{B}_{ss} &= -\tilde{\Sigma}_{22}^{-1}\tilde{\Sigma}_{2c}\tilde{\Sigma}_{*2}^{-1}\tilde{\Sigma}_{c1} + \tilde{\Sigma}_{22}^{-1}\tilde{\Sigma}_{21} + \tilde{\Sigma}_{22}^{-1}\tilde{\Sigma}_{2c}\tilde{\Sigma}_{*}^{-1}\tilde{\Sigma}_{c2}\tilde{\Sigma}_{22}^{-1}\tilde{\Sigma}_{21} \\ &= \tilde{\Sigma}_{22}^{-1}\tilde{\Sigma}_{21} + O_p\left(\frac{1}{T}\right) \\ &= \Phi' + O_p\left(\frac{1}{\sqrt{T}}\right).\end{aligned} \tag{A.55}$$

Finally, using results (A.49) and (A.53) we can compute:

$$\tilde{R} = \tilde{A}\tilde{B} \tag{A.56}$$

$$\begin{aligned} &= \begin{pmatrix} I_{k^c} & \tilde{A}_{cs} \\ 0 & \tilde{A}_{ss}\end{pmatrix}\begin{pmatrix} I_{k^c} & \tilde{B}_{cs} \\ 0 & \tilde{B}_{ss}\end{pmatrix} \\ &= \begin{pmatrix} I_{k^c} & \tilde{R}_{cs} \\ 0 & \tilde{R}_{ss}\end{pmatrix},\end{aligned} \tag{A.57}$$

where

$$\tilde{R}_{cs} = \tilde{B}_{cs} + \tilde{A}_{cs}\tilde{B}_{ss} = O_p(1/\sqrt{T}), \tag{A.58}$$

$$\begin{aligned}\tilde{R}_{ss} &= \tilde{A}_{ss}\tilde{B}_{ss} \\ &= \tilde{\Sigma}_{11}^{-1}\tilde{\Sigma}_{12}\tilde{\Sigma}_{22}^{-1}\tilde{\Sigma}_{21} + O_p(1/T) \\ &= \Phi\Phi' + O_p(1/\sqrt{T}).\end{aligned} \tag{A.59}$$

The eigenvalues of matrix $\tilde{R}$ are $\tilde{\rho}_1^2 = ... = \tilde{\rho}_{k^c}^2 = 1 > \tilde{\rho}_{k^c+1}^2 \geq ... \geq \tilde{\rho}_{k_1}^2$. The eigenvectors associated with the first $k^c$ eigenvalues are spanned by the columns of matrix:

$$\underset{(k_1\times k^c)}{E_c} = \begin{bmatrix} I_{k^c} \\ 0 \end{bmatrix}. \tag{A.60}$$

Define:

$$\underset{(k_1\times (k_1-k^c))}{E_s} = \begin{bmatrix} 0 \\ I_{k_1-k^c} \end{bmatrix}. \tag{A.61}$$

We note:

$$I_{k_1} = \begin{bmatrix} E_c & \vdots & E_s \end{bmatrix},$$

so that the columns of matrices $E_c$ and $E_s$ span the space $\mathbb{R}^{k_1}$. The estimators of the first $k^c$ canonical correlations are such that $\hat{\rho}_\ell^2$, with $\ell = 1, ..., k^c$ are the $k^c$ largest eigenvalues of matrix $\hat{R}$. We derive their asymptotic expansion using perturbations arguments.

### A.5.3 Perturbation of the eigenvalues and eigenvectors of matrix $\hat{R}$

Under the null hypothesis $H(k^c)$, let $\hat{W}_1^*$ be a $(k_1, k^c)$ matrix whose columns are eigenvectors of matrix $\hat{R}$ associated with the eigenvalues $\hat{\rho}_\ell^2$, with $\ell = 1, ..., k^c$. We have:

$$\hat{R}\hat{W}_1^* = \hat{W}_1^*\hat{\Lambda}, \tag{A.62}$$

where:

$$\hat{\Lambda} \quad = \quad diag(\hat{\rho}_\ell^2, \ell = 1, ..., k^c), \tag{A.63}$$

is the $(k^c, k^c)$ diagonal matrix containing the $k^c$ largest eigenvalues of $\hat{R}$. We know from the previous subsection that the eigenspace associated with the largest eigenvalue of $\tilde{R}$ (equal to 1) has dimension $k^c$ and is spanned by the columns of matrix $E_c$. Since the columns of $E_c$ and $E_s$ span $\mathbb{R}^{k_1}$, we can write the following expansions:

$$\hat{W}_1^* \quad = \quad E_c\hat{\mathcal{U}} + E_s\alpha, \tag{A.64}$$

$$\hat{\Lambda} \quad = \quad I_{k^c} + \hat{M}, \tag{A.65}$$

where $E_c$ and $E_s$ are defined in equations (A.60) and (A.61), $\hat{\mathcal{U}}$ is a $(k^c, k^c)$ nonsingular matrix, $\hat{M} = diag(\hat{\mu}_1, ..., \hat{\mu}_{k^c})$, and $\alpha$ is a $(k_1 - k^c, k^c)$ matrix, with $\alpha$, $\hat{\mu}_1$, ..., $\hat{\mu}_{k^c}$ converging to zero as $N_1, N_2, T \to \infty$. Substituting the expansions in equations (A.34) and (A.62) we get:

$$(\tilde{R} + \hat{\Psi})(E_c\hat{\mathcal{U}} + E_s\alpha) \quad \simeq \quad (E_c\hat{\mathcal{U}} + E_s\alpha)(I_{k^c} + \hat{M}),$$

which implies:

$$\tilde{R}E_c\hat{\mathcal{U}} + \tilde{R}E_s\alpha + \hat{\Psi}E_c\hat{\mathcal{U}} + \hat{\Psi}E_s\alpha \quad \simeq \quad E_c\hat{\mathcal{U}} + E_s\alpha + E_c\hat{\mathcal{U}}\hat{M} + E_s\alpha\hat{M}.$$

By using $\tilde{R}E_c = E_c$, and keeping only the terms at first order, we get:

$$\tilde{R}E_s\alpha + \hat{\Psi}E_c\hat{\mathcal{U}} \quad \simeq \quad E_s\alpha + E_c\hat{\mathcal{U}}\hat{M}. \tag{A.66}$$

Pre-multiplying equation (A.66) by $E_c'$, we get:

$$E_c'\tilde{R}E_s\alpha + E_c'\hat{\Psi}E_c\hat{\mathcal{U}} \simeq \hat{\mathcal{U}}\hat{M}$$

$$\Leftrightarrow \hat{M} \simeq \hat{\mathcal{U}}^{-1}\left(\tilde{R}_{cs}\alpha + \hat{\Psi}_{cc}\hat{\mathcal{U}}\right), \tag{A.67}$$

where we use the fact that $\hat{\mathcal{U}}$ is non-singular and

$$\hat{\Psi}_{cc} = E_c'\hat{\Psi}E_c.$$

Pre-multiplying equation (A.66) by $E_s'$, we get:

$$E_s'\tilde{R}E_s\alpha + E_s'\hat{\Psi}E_c\hat{\mathcal{U}} \simeq \alpha$$

$$\Leftrightarrow \alpha \simeq \tilde{R}_{ss}\alpha + \hat{\Psi}_{sc}\hat{\mathcal{U}}, \tag{A.68}$$

where

$$\hat{\Psi}_{sc} = E_s'\hat{\Psi}E_c.$$

This implies:

$$\alpha \quad \simeq \quad (I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\hat{\Psi}_{sc}\hat{\mathcal{U}}. \tag{A.69}$$

Substituting the first order approximation of $\alpha$ from equation (A.69) into equation (A.64) we get:

$$\hat{W}_1^* \quad \simeq \quad \left(E_c + E_s(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\hat{\Psi}_{sc}\right)\hat{\mathcal{U}}. \tag{A.70}$$

The normalized eigenvectors corresponding to the canonical directions are:

$$\hat{W}_1 \quad = \quad \hat{W}_1^* \cdot diag(\hat{W}_1^{*\prime}\hat{V}_{11}\hat{W}_1^*)^{-1/2}. \tag{A.71}$$

Substituting the first order approximation of $\alpha$ from equation (A.69) into (A.67), we get the first order approximation of matrix $\hat{M}$:

$$\hat{M} \simeq \hat{\mathcal{U}}^{-1}\left(\hat{\Psi}_{cc}+\tilde{R}_{cs}(I_{k_1-k^c}-\tilde{R}_{ss})^{-1}\hat{\Psi}_{sc}\right)\hat{\mathcal{U}}. \tag{A.72}$$

Substituting the first order approximation of $\hat{M}$ from equation (A.72) into (A.65), matrix $\hat{\Lambda}$ can be approximated as:

$$\hat{\Lambda} \simeq I_{k^c}+\hat{\mathcal{U}}^{-1}\left(\hat{\Psi}_{cc}+\tilde{R}_{cs}(I_{k_1-k^c}-\tilde{R}_{ss})^{-1}\hat{\Psi}_{sc}\right)\hat{\mathcal{U}}.$$

Note that this first order approximation holds for the terms in the main diagonal, as matrix $\hat{\Lambda}$ has been defined to be diagonal, and the out-of-diagonal terms are of higher order. Up to higher order terms we have:

$$\hat{\Lambda}^{1/2} \simeq I_{k^c}+\frac{1}{2}\hat{\mathcal{U}}^{-1}\left[\hat{\Psi}_{cc}+\tilde{R}_{cs}(I_{k_1-k^c}-\tilde{R}_{ss})^{-1}\hat{\Psi}_{sc}\right]\hat{\mathcal{U}},$$

which implies:

$$\begin{aligned}\sum_{\ell=1}^{k^c}\hat{\rho}_\ell &= tr(\hat{\Lambda}^{1/2})\\ &= k^c+\frac{1}{2}tr\left[\hat{\mathcal{U}}^{-1}\left(\hat{\Psi}_{cc}+\tilde{R}_{cs}(I_{k_1-k^c}-\tilde{R}_{ss})^{-1}\hat{\Psi}_{sc}\right)\hat{\mathcal{U}}\right]+O_p\left(\frac{1}{NT}\right),\\ &= k^c+\frac{1}{2}tr\left[\hat{\Psi}_{cc}+\tilde{R}_{cs}(I_{k_1-k^c}-\tilde{R}_{ss})^{-1}\hat{\Psi}_{sc}\right]+O_p\left(\frac{1}{NT}\right),\end{aligned} \tag{A.73}$$

by the commutative property of the trace.

### A.5.4 Asymptotic distribution of $\sum_{\ell=1}^{k^c}\hat{\rho}_\ell$.

Equation (A.73) can be written as:

$$\begin{aligned}\sum_{\ell=1}^{k^c}\hat{\rho}_\ell &= k^c+\frac{1}{2}tr\left\{\left[I_{k^c}\ \vdots\ \tilde{R}_{cs}(I_{(k_1-k^c)}-\tilde{R}_{ss})^{-1}\right]\hat{\Psi}E_c\right\}+O_p\left(\frac{1}{NT}\right)\\ &= k^c+\frac{1}{2}tr\left\{\left[I_{k^c}\ \vdots\ \tilde{R}_{cs}(I_{(k_1-k^c)}-\tilde{R}_{ss})^{-1}\right]\tilde{V}_{11}^{-1}\hat{\Psi}^*E_c\right\}+O_p\left(\frac{1}{NT}\right).\end{aligned} \tag{A.74}$$

Substituting equation (A.33), we get:

$$\sum_{\ell=1}^{k^c}\hat{\rho}_\ell = k^c+\frac{1}{2}tr\left\{\left[I_{k^c}\ \vdots\ \tilde{R}_{cs}(I_{(k_1-k^c)}-\tilde{R}_{ss})^{-1}\right]\tilde{V}_{11}^{-1}\begin{bmatrix}\hat{\Psi}^*_{cc}\\ \hat{\Psi}^*_{sc}\end{bmatrix}\right\}+O_p\left(\frac{1}{NT}\right) \tag{A.75}$$

where:

$$\hat{\Psi}^*_{cc} = \left[-\hat{X}_{11}\tilde{R}+\hat{X}_{12}\tilde{B}-\tilde{B}'\hat{X}_{22}\tilde{B}+\tilde{B}'\hat{X}_{21}\right]_{(11)}, \tag{A.76}$$

$$\hat{\Psi}^*_{sc} = \left[-\hat{X}_{11}\tilde{R}+\hat{X}_{12}\tilde{B}-\tilde{B}'\hat{X}_{22}\tilde{B}+\tilde{B}'\hat{X}_{21}\right]_{(21)}, \tag{A.77}$$

with $M_{(ij)}$ denoting the block in position $(i,j)$ of matrix $M$. As matrices $\tilde{R}$ and $\tilde{B}$ have the same structure $[\, E_c \,\vdots\, * \,]$, we have:

$$\hat{\Psi}^*_{cc} = \left[-\hat{X}_{11} + \hat{X}_{12} - \tilde{B}'(\hat{X}_{22} - \hat{X}_{21})\right]_{(11)}, \tag{A.78}$$

$$\hat{\Psi}^*_{sc} = \left[-\hat{X}_{11} + \hat{X}_{12} - \tilde{B}'(\hat{X}_{22} - \hat{X}_{21})\right]_{(21)}. \tag{A.79}$$

Moreover as $\tilde{B}' = \begin{bmatrix} I_{k^c} & 0 \\ \tilde{B}'_{cs} & \tilde{B}'_{ss} \end{bmatrix}$, equation (A.78) further simplifies to:

$$\hat{\Psi}^*_{cc} = \left[-\hat{X}_{11} + \hat{X}_{12} - \hat{X}_{22} + \hat{X}_{21}\right]_{(11)}. \tag{A.80}$$

Equations (A.79) and (A.80) allow to perform the asymptotic expansion of terms $\hat{\Psi}^*_{sc}$ and $\hat{\Psi}^*_{cc}$, respectively. Let us compute the asymptotic expansions of the terms $\hat{X}_{11}$, $\hat{X}_{12}$, $\hat{X}_{22}$ and $\hat{X}_{21}$. Vectors $u_{j,t}$, with $j = 1, 2$, can be partitioned into the $k^c$-dimensional vector $u^{(c)}_{jt}$ and the $k^s_j$-dimensional vector $u^{(s)}_{jt}$:

$$u_{jt} = \begin{bmatrix} u^{(c)}_{jt} \\ u^{(s)}_{jt} \end{bmatrix}, \quad j = 1, 2\,, \tag{A.81}$$

and from Assumption A.5 we can express $\Sigma_{u,j}$, $j = 1, 2$, as: [31]

$$\Sigma_{u,j} = E[u_{jt}u'_{jt}] = E\begin{bmatrix} u^{(c)}_{jt}u^{(c)\prime}_{jt} & u^{(c)}_{jt}u^{(s)\prime}_{jt} \\ u^{(s)}_{jt}u^{(c)\prime}_{jt} & u^{(s)}_{jt}u^{(s)\prime}_{jt} \end{bmatrix} = \begin{bmatrix} \Sigma^{(cc)}_{u,j} & \Sigma^{(cs)}_{u,j} \\ \Sigma^{(sc)}_{u,j} & \Sigma^{(ss)}_{u,j} \end{bmatrix}, \quad j = 1, 2\,. \tag{A.82}$$

We also define:

$$\Sigma_{u,12} := E[u_{1t}u'_{2t}] := E\begin{bmatrix} u^{(c)}_{1t}u^{(c)\prime}_{2t} & u^{(c)}_{1t}u^{(s)\prime}_{2t} \\ u^{(s)}_{1t}u^{(c)\prime}_{2t} & u^{(s)}_{1t}u^{(s)\prime}_{2t} \end{bmatrix} = \begin{bmatrix} \Sigma^{(cc)}_{u,12} & \Sigma^{(cs)}_{u,12} \\ \Sigma^{(sc)}_{u,12} & \Sigma^{(ss)}_{u,12} \end{bmatrix}, \tag{A.83}$$

and

$$\Sigma_{u,21} = \Sigma'_{u,12}. \tag{A.84}$$

From equation (A.27) we have:

$$\begin{aligned}
\hat{X}_{11} &= \frac{\mu_N}{T\sqrt{N}}\sum_{t=1}^{T}(h_{1,t}u'_{1,t} + u_{1,t}h'_{1,t}) + \frac{\mu_N^2}{TN}\sum_{t=1}^{T}u_{1,t}u'_{1,t} \\
&= \frac{\mu_N}{T\sqrt{N}}\sum_{t=1}^{T}\left(\begin{bmatrix} f^c_t \\ f^s_{1t} \end{bmatrix}\begin{bmatrix} u^{(c)\prime}_{1t} & u^{(s)\prime}_{1t} \end{bmatrix} + \begin{bmatrix} u^{(c)}_{1t} \\ u^{(s)}_{1t} \end{bmatrix}\begin{bmatrix} f^{c\prime}_t & f^{s\prime}_{1t} \end{bmatrix}\right) \\
&\quad + \frac{\mu_N^2}{TN}\sum_{t=1}^{T}\begin{bmatrix} u^{(c)}_{1t} \\ u^{(s)}_{1t} \end{bmatrix}\begin{bmatrix} u^{(c)\prime}_{1t} & u^{(s)\prime}_{1t} \end{bmatrix} \\
&= \frac{\mu_N}{\sqrt{TN}}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\begin{bmatrix} f^c_t u^{(c)\prime}_{1t} + u^{(c)}_{1t}f^{c\prime}_t & f^c_t u^{(s)\prime}_{1t} + u^{(c)}_{1t}f^{s\prime}_{1t} \\ f^s_{1t}u^{(c)\prime}_{1t} + u^{(s)}_{1t}f^{c\prime}_t & f^s_{1t}u^{(s)\prime}_{1t} + u^{(s)}_{1t}f^{s\prime}_{1t} \end{bmatrix}\right) + \frac{\mu_N^2}{TN}\sum_{t=1}^{T}\begin{bmatrix} u^{(c)}_{1t}u^{(c)\prime}_{1t} & u^{(c)}_{1t}u^{(s)\prime}_{1t} \\ u^{(s)}_{1t}u^{(c)\prime}_{1t} & u^{(s)}_{1t}u^{(s)\prime}_{1t} \end{bmatrix},
\end{aligned}$$

[31] Matrix $\Sigma_{u,j}$ is the asymptotic variance of $u_{j,t}$ as $N_j \to \infty$. We omit the limit for expository purpose.

and from assumption A.5 *b)* we have:

$$\begin{aligned}\hat{X}_{11} &= \frac{\mu_N}{\sqrt{TN}}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[\begin{array}{cc} f_t^c u_{1t}^{(c)\prime} + u_{1t}^{(c)} f_t^{c\prime} & f_t^c u_{1t}^{(s)\prime} + u_{1t}^{(c)} f_{1t}^{s\prime} \\ f_{1t}^s u_{1t}^{(c)\prime} + u_{1t}^{(s)} f_t^{c\prime} & f_{1t}^s u_{1t}^{(s)\prime} + u_{1t}^{(s)} f_{1t}^{s\prime}\end{array}\right]\right) + \frac{\mu_N^2}{N} E\left[\begin{array}{cc} u_{1t}^{(c)} u_{1t}^{(c)\prime} & u_{1t}^{(c)} u_{1t}^{(s)\prime} \\ u_{1t}^{(s)} u_{1t}^{(c)\prime} & u_{1t}^{(s)} u_{1t}^{(s)\prime}\end{array}\right] \\ &+ \frac{\mu_N^2}{N\sqrt{T}}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[\begin{array}{cc} u_{1t}^{(c)} u_{1t}^{(c)\prime} - E[u_{1t}^{(c)} u_{1t}^{(c)\prime}] & u_{1t}^{(c)} u_{1t}^{(s)\prime} - E[u_{1t}^{(c)} u_{1t}^{(s)\prime}] \\ u_{1t}^{(s)} u_{1t}^{(c)\prime} - E[u_{1t}^{(s)} u_{1t}^{(c)\prime}] & u_{1t}^{(s)} u_{1t}^{(s)\prime} - E[u_{1t}^{(c)} u_{1t}^{(s)\prime}]\end{array}\right]\right) \\ &= \frac{\mu_N}{\sqrt{TN}}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[\begin{array}{cc} f_t^c u_{1t}^{(c)\prime} + u_{1t}^{(c)} f_t^{c\prime} & f_t^c u_{1t}^{(s)\prime} + u_{1t}^{(c)} f_{1t}^{s\prime} \\ f_{1t}^s u_{1t}^{(c)\prime} + u_{1t}^{(s)} f_t^{c\prime} & f_{1t}^s u_{1t}^{(s)\prime} + u_{1t}^{(s)} f_{1t}^{s\prime}\end{array}\right]\right) \\ &+ \frac{\mu_N^2}{N}\left[\begin{array}{cc} \Sigma_{u,1}^{(cc)} & \Sigma_{u,1}^{(cs)} \\ \Sigma_{u,1}^{(sc)} & \Sigma_{u,1}^{(ss)}\end{array}\right] + \frac{\mu_N^2}{N\sqrt{T}}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[\begin{array}{cc} u_{1t}^{(c)} u_{1t}^{(c)\prime} - \Sigma_{u,1}^{(cc)} & u_{1t}^{(c)} u_{1t}^{(s)\prime} - \Sigma_{u,1}^{(cs)} \\ u_{1t}^{(s)} u_{1t}^{(c)\prime} - \Sigma_{u,1}^{(sc)} & u_{1t}^{(s)} u_{1t}^{(s)\prime} - \Sigma_{u,1}^{(ss)}\end{array}\right]\right). \qquad \text{(A.85)}\end{aligned}$$

Analogously, from (A.28) we have:

$$\begin{aligned}\hat{X}_{22} &= \frac{1}{\sqrt{TN}}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[\begin{array}{cc} f_t^c u_{2t}^{(c)\prime} + u_{2t}^{(c)} f_t^{c\prime} & f_t^c u_{2t}^{(s)\prime} + u_{2t}^{(c)} f_{2t}^{s\prime} \\ f_{2t}^s u_{2t}^{(c)\prime} + u_{2t}^{(s)} f_t^{c\prime} & f_{2t}^s u_{2t}^{(s)\prime} + u_{2t}^{(s)} f_{2t}^{s\prime}\end{array}\right]\right) \\ &+ \frac{1}{N}\left[\begin{array}{cc} \Sigma_{u,2}^{(cc)} & \Sigma_{u,2}^{(cs)} \\ \Sigma_{u,2}^{(sc)} & \Sigma_{u,2}^{(22)}\end{array}\right] + \frac{1}{N\sqrt{T}}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[\begin{array}{cc} u_{2t}^{(c)} u_{2t}^{(c)\prime} - \Sigma_{u,2}^{(cc)} & u_{2t}^{(c)} u_{2t}^{(s)\prime} - \Sigma_{u,2}^{(cs)} \\ u_{2t}^{(s)} u_{2t}^{(c)\prime} - \Sigma_{u,2}^{(sc)} & u_{2t}^{(s)} u_{2t}^{(s)\prime} - \Sigma_{u,2}^{(ss)}\end{array}\right]\right). \qquad \text{(A.86)}\end{aligned}$$

From equation (A.23), the term $\hat{X}_{12}$ results to be:

$$\begin{aligned}\hat{X}_{12} &= \frac{1}{T\sqrt{N}}\sum_{t=1}^{T}(h_{1,t}u_{2,t}' + \mu_N u_{1,t}h_{2,t}') + \frac{\mu_N}{TN}\sum_{t=1}^{T}u_{1,t}u_{2,t}' \\ &= \frac{1}{T\sqrt{N}}\sum_{t=1}^{T}\left(\left[\begin{array}{c} f_t^c \\ f_{1t}^s\end{array}\right]\left[\begin{array}{cc} u_{2t}^{(c)\prime} & u_{2t}^{(s)\prime}\end{array}\right] + \mu_N\left[\begin{array}{c} u_{1t}^{(c)} \\ u_{1t}^{(s)}\end{array}\right]\left[\begin{array}{cc} f_t^{c\prime} & f_{2t}^{s\prime}\end{array}\right]\right) \\ &+ \frac{\mu_N}{TN}\sum_{t=1}^{T}\left[\begin{array}{c} u_{1t}^{(c)} \\ u_{1t}^{(s)}\end{array}\right]\left[\begin{array}{cc} u_{2t}^{(c)\prime} & u_{2t}^{(s)\prime}\end{array}\right] \\ &= \frac{1}{\sqrt{TN}}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[\begin{array}{cc} f_t^c u_{2t}^{(c)\prime} + \mu_N u_{1t}^{(c)} f_t^{c\prime} & f_t^c u_{2t}^{(s)\prime} + \mu_N u_{1t}^{(c)} f_{2t}^{s\prime} \\ f_{1t}^s u_{2t}^{(c)\prime} + \mu_N u_{1t}^{(s)} f_t^{c\prime} & f_{1t}^s u_{2t}^{(s)\prime} + \mu_N u_{1t}^{(s)} f_{2t}^{s\prime}\end{array}\right]\right) + \frac{\mu_N}{N}\left[\begin{array}{cc} \Sigma_{u,12}^{(cc)} & \Sigma_{u,12}^{(cs)} \\ \Sigma_{u,12}^{(sc)} & \Sigma_{u,12}^{(ss)}\end{array}\right] \\ &+ \frac{\mu_N}{N\sqrt{T}}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[\begin{array}{cc} u_{1t}^{(c)} u_{2t}^{(c)\prime} - \Sigma_{u,12}^{(cc)} & u_{1t}^{(c)} u_{2t}^{(s)\prime} - \Sigma_{u,12}^{(cs)} \\ u_{1t}^{(s)} u_{2t}^{(c)\prime} - \Sigma_{u,12}^{(sc)} & u_{1t}^{(s)} u_{2t}^{(s)\prime} - \Sigma_{u,12}^{(ss)}\end{array}\right]\right). \qquad \text{(A.87)}\end{aligned}$$

Finally we have:

$$\begin{aligned}\hat{X}_{21} &= \hat{X}_{12}' \\ &= \frac{1}{\sqrt{TN}}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[\begin{array}{cc} u_{2t}^{(c)} f_t^{c\prime} + \mu_N f_t^c u_{1t}^{(c)\prime} & u_{2t}^{(c)} f_{1t}^{s\prime} + \mu_N f_t^c u_{1t}^{(s)\prime} \\ u_{2t}^{(s)} f_t^{c\prime} + \mu_N f_{2t}^s u_{1t}^{(c)\prime} & u_{2t}^{(s)} f_{1t}^{s\prime} + \mu_N f_{2t}^s u_{1t}^{(s)\prime}\end{array}\right]\right) + \frac{\mu_N}{N}\left[\begin{array}{cc} \Sigma_{u,21}^{(cc)} & \Sigma_{u,21}^{(cs)} \\ \Sigma_{u,21}^{(sc)} & \Sigma_{u,21}^{(22)}\end{array}\right] \\ &+ \frac{\mu_N}{N\sqrt{T}}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[\begin{array}{cc} u_{2t}^{(c)} u_{1t}^{(c)\prime} - \Sigma_{u,21}^{(cc)} & u_{2t}^{(c)} u_{1t}^{(s)\prime} - \Sigma_{u,21}^{(cs)} \\ u_{2t}^{(s)} u_{1t}^{(c)\prime} - \Sigma_{u,21}^{(sc)} & u_{2t}^{(s)} u_{1t}^{(s)\prime} - \Sigma_{u,21}^{(ss)}\end{array}\right]\right). \qquad \text{(A.88)}\end{aligned}$$

We can now compute directly term $\hat{\Psi}^*_{cc}$. From equation (A.80), we get:

$$
\begin{aligned}
& \hat{\Psi}^*_{cc} && \text{(A.89)}\\
= & \left[-\hat{X}_{11}+\hat{X}_{12}-\hat{X}_{22}+\hat{X}_{21}\right]_{(11)},\\
= & \frac{1}{\sqrt{TN}}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[-\mu_N f^c_t u^{(c)\prime}_{1t}-\mu_N u^{(c)}_{1t}f^{c\prime}_t+f^c_t u^{(c)\prime}_{2t}+\mu_N u^{(c)}_{1t}f^{c\prime}_t-f^c_t u^{(c)\prime}_{2t}-u^{(c)}_{2t}f^{c\prime}_t+u^{(c)}_{2t}f^{c\prime}_t+\mu_N f^c_t u^{(c)\prime}_{1t}\right]\right)\\
& +\frac{1}{N}[-\mu_N^2\Sigma^{(cc)}_{u,1}-\Sigma^{(cc)}_{u,2}+\mu_N\Sigma^{(cc)}_{u,12}+\mu_N\Sigma^{(cc)}_{u,21}]\\
& +\frac{1}{N\sqrt{T}}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[-\mu_N^2[u^{(c)}_{1t}u^{(c)\prime}_{1t}-\Sigma^{(cc)}_{u,1}]+\mu_N[u^{(c)}_{1t}u^{(c)\prime}_{2t}-\Sigma^{(cc)}_{u,12}]-[u^{(c)}_{2t}u^{(c)\prime}_{2t}-\Sigma^{(cc)}_{u,2}]+\mu_N[u^{(c)}_{2t}u^{(c)\prime}_{1t}-\Sigma^{(cc)}_{u,12}]\right]\right)\\
= & -\frac{1}{N}E[(\mu_N u^{(c)}_{1t}-u^{(c)}_{2t})(\mu_N u^{(c)}_{1t}-u^{(c)}_{2t})']\\
& -\frac{1}{N\sqrt{T}}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[(\mu_N u^{(c)}_{1t}-u^{(c)}_{2t})(\mu_N u^{(c)}_{1t}-u^{(c)}_{2t})'-E[(\mu_N u^{(c)}_{1t}-u^{(c)}_{2t})(\mu_N u^{(c)}_{1t}-u^{(c)}_{2t})']\right]\right). && \text{(A.90)}
\end{aligned}
$$

Using the limit $\mu_N \to \mu$, we get:

$$
\begin{aligned}
\hat{\Psi}^*_{cc} = & -\frac{1}{N}E[(\mu_N u^{(c)}_{1t}-u^{(c)}_{2t})(\mu_N u^{(c)}_{1t}-u^{(c)}_{2t})']\\
& -\frac{1}{N\sqrt{T}}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[(\mu_N u^{(c)}_{1t}-u^{(c)}_{2t})(\mu_N u^{(c)}_{1t}-u^{(c)}_{2t})'-E[(\mu u^{(c)}_{1t}-u^{(c)}_{2t})(\mu u^{(c)}_{1t}-u^{(c)}_{2t})']\right]\right)\\
& +o_p\left(\frac{1}{N\sqrt{T}}\right). \qquad \text{(A.91)}
\end{aligned}
$$

Before computing $\hat{\Psi}^*_{sc}$ and substituting it into equation (A.75), we note that some of the terms of this equation can be further simplified. Let us consider the asymptotic expansion of the following term of equation (A.75):

$$
\left[I_{k^c}\ \vdots\ \tilde{R}_{cs}(I_{(k_1-k^c)}-\tilde{R}_{ss})^{-1}\right]\tilde{V}_{11}^{-1}.
$$

Using equation (A.48), we get:

$$
\begin{aligned}
& \left[I_{k^c}\ \vdots\ \tilde{R}_{cs}(I_{k_1-k^c}-\tilde{R}_{ss})^{-1}\right]\tilde{V}_{11}^{-1}\\
= & \left[I_{k^c}\ \vdots\ \tilde{R}_{cs}(I_{k_1-k^c}-\tilde{R}_{ss})^{-1}\right]\begin{bmatrix}\tilde{\Sigma}_{cc}^{-1} & -\tilde{\Sigma}_{c1}\\ -\tilde{\Sigma}_{1c} & \tilde{\Sigma}_{11}^{-1}\end{bmatrix}+O_p\left(\frac{1}{T}\right)\\
= & \left[\tilde{\Sigma}_{cc}^{-1}-\tilde{R}_{cs}(I_{k_1-k^c}-\tilde{R}_{ss})^{-1}\tilde{\Sigma}_{1c}\ \vdots\ -\tilde{\Sigma}_{c1}+\tilde{R}_{cs}(I_{k_1-k^c}-\tilde{R}_{ss})^{-1}\tilde{\Sigma}_{11}^{-1}\right]+O_p\left(\frac{1}{T}\right)\\
= & \left[\tilde{\Sigma}_{cc}^{-1}\ \vdots\ -\tilde{\Sigma}_{c1}+\tilde{R}_{cs}(I_{k_1-k^c}-\tilde{R}_{ss})^{-1}\right]+O_p\left(\frac{1}{T}\right), \qquad \text{(A.92)}
\end{aligned}
$$

where the last equality follows form the fact that $\tilde{R}_{cs}=O_p(1/\sqrt{T})$, $\tilde{\Sigma}_{1c}=O_p(1/\sqrt{T})$ and $\tilde{\Sigma}_{11}=I_{k_1}+O_p(1/\sqrt{T})$. Note that equation (A.92) can be further simplified, considering the asymptotic expansion of term $\tilde{R}_{cs}$. Let us consider the different terms in the equations of $\tilde{R}_{cs}$ and $\tilde{R}_{ss}$:

$$
\begin{aligned}
\tilde{R}_{cs} &= \tilde{B}_{cs}+\tilde{A}_{cs}\tilde{B}_{ss}, && \text{(A.93)}\\
\tilde{R}_{ss} &= \tilde{A}_{ss}\tilde{B}_{ss}, && \text{(A.94)}
\end{aligned}
$$

where:

$$\tilde{A}_{cs} = \tilde{\Sigma}_{*}^{-1}\tilde{\Sigma}_{c2} - \tilde{\Sigma}_{*}^{-1}\tilde{\Sigma}_{c1}\tilde{\Sigma}_{11}^{-1}\tilde{\Sigma}_{12}, \quad (A.95)$$

$$\tilde{A}_{ss} = \Sigma_{11}^{-1}\tilde{\Sigma}_{12} + O_p\left(\frac{1}{T}\right), \quad (A.96)$$

$$\tilde{B}_{cs} = \tilde{\Sigma}_{*2}^{-1}\tilde{\Sigma}_{c1} - \tilde{\Sigma}_{*2}^{-1}\tilde{\Sigma}_{c2}\tilde{\Sigma}_{22}^{-1}\tilde{\Sigma}_{21}, \quad (A.97)$$

$$\tilde{B}_{ss} = \tilde{\Sigma}_{22}^{-1}\tilde{\Sigma}_{21} + O_p\left(\frac{1}{T}\right). \quad (A.98)$$

Substituting equations (A.95) - (A.98) into equations (A.93) and (A.94) we get:

$$\begin{aligned}\tilde{R}_{cs} &= \tilde{\Sigma}_{*2}^{-1}\tilde{\Sigma}_{c1} - \tilde{\Sigma}_{*2}^{-1}\tilde{\Sigma}_{c2}\tilde{\Sigma}_{22}^{-1}\tilde{\Sigma}_{21} + \left[\tilde{\Sigma}_{*}^{-1}\tilde{\Sigma}_{c2} - \tilde{\Sigma}_{*}^{-1}\tilde{\Sigma}_{c1}\tilde{\Sigma}_{11}^{-1}\tilde{\Sigma}_{12}\right]\left[\tilde{\Sigma}_{22}^{-1}\tilde{\Sigma}_{21} + O_p\left(\frac{1}{T}\right)\right] \\ &= \tilde{\Sigma}_{c1}\left[I_{k_1^s} - \tilde{\Sigma}_{11}^{-1}\tilde{\Sigma}_{12}\tilde{\Sigma}_{22}^{-1}\tilde{\Sigma}_{21}\right] + O_p\left(\frac{1}{T}\right), \quad (A.99)\end{aligned}$$

and

$$\begin{aligned}\tilde{R}_{ss} &= \left[\tilde{\Sigma}_{11}^{-1}\tilde{\Sigma}_{12} + O_p\left(\frac{1}{T}\right)\right]\left[\tilde{\Sigma}_{22}^{-1}\tilde{\Sigma}_{21} + O_p\left(\frac{1}{T}\right)\right] \\ &= \tilde{\Sigma}_{11}^{-1}\tilde{\Sigma}_{12}\tilde{\Sigma}_{22}^{-1}\tilde{\Sigma}_{21} + O_p\left(\frac{1}{T}\right). \quad (A.100)\end{aligned}$$

Therefore we have:

$$\tilde{R}_{cs} = \tilde{\Sigma}_{c1}(I_{k_1-k^c} - \tilde{R}_{ss}) + O_p\left(\frac{1}{T}\right), \quad (A.101)$$

which implies :

$$-\tilde{\Sigma}_{c1} + \tilde{R}_{cs}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1} = O_p\left(\frac{1}{T}\right). \quad (A.102)$$

Equation (A.102) and $\hat{\Psi}_{sc}^{*} = O_p\left(\frac{1}{\sqrt{NT}}\right)$, together with the assumption $\sqrt{N}/T = o(1)$, imply:

$$\left[-\tilde{\Sigma}_{c1} + \tilde{R}_{cs}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\right]\hat{\Psi}_{sc}^{*} = o_p\left(\frac{1}{N\sqrt{T}}\right). \quad (A.103)$$

Therefore, substituting results (A.90), (A.92), and (A.103) into equation (A.75), and rearranging terms, we get:

$$\begin{aligned}\sum_{\ell=1}^{k^c}\hat{\rho}_\ell &= k^c - \frac{1}{N}\frac{1}{2}tr\left\{\tilde{\Sigma}_{cc}^{-1}E[(\mu_N u_{1t}^{(c)} - u_{2t}^{(c)})(\mu_N u_{1t}^{(c)} - u_{2t}^{(c)})']\right\} \\ &\quad -\frac{1}{N\sqrt{T}}\frac{1}{2}tr\left\{\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[(\mu u_{1t}^{(c)} - u_{2t}^{(c)})(\mu u_{1t}^{(c)} - u_{2t}^{(c)})' - E[(\mu u_{1t}^{(c)} - u_{2t}^{(c)})(\mu u_{1t}^{(c)} - u_{2t}^{(c)})']\right]\right)\right\} \\ &\quad +o_p\left(\frac{1}{N\sqrt{T}}\right). \quad (A.104)\end{aligned}$$

From the definition of matrix $\Sigma_{U,N}$ we have;

$$E[(\mu_N u_{1t}^{(c)} - u_{2t}^{(c)})(\mu_N u_{1t}^{(c)} - u_{2t}^{(c)})'] = \Sigma_{U,N}. \quad (A.105)$$

Moreover, let us define:

$$U_t \quad := \quad \mu u_{1t}^{(c)} - u_{2t}^{(c)}. \tag{A.106}$$

Definition (A.106) together with the commutativity and linearity properties of the trace operator allow to write the fourth term in the r.h.s. of equation (A.104) as:

$$tr\left\{\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[(\mu u_{1t}^{(c)} - u_{2t}^{(c)})(\mu u_{1t}^{(c)} - u_{2t}^{(c)})' - E[(\mu u_{1t}^{(c)} - u_{2t}^{(c)})(\mu u_{1t}^{(c)} - u_{2t}^{(c)})']\right]\right\}$$
$$= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left\{U_t'U_t - E(U_t'U_t)\right\}. \tag{A.107}$$

Equations (A.105) and (A.107) allow to write equation (A.104) as:

$$\sum_{\ell=1}^{k^c}\hat{\rho}_\ell \quad = \quad k^c - \frac{1}{2N}tr\left\{\tilde{\Sigma}_{cc}^{-1}\Sigma_{U,N}\right\} - \frac{1}{2N\sqrt{T}}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}[U_t'U_t - E(U_t'U_t)]\right) + o_p\left(\frac{1}{N\sqrt{T}}\right). \tag{A.108}$$

By a CLT for weakly dependent data we have:

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}[U_t'U_t - E(U_t'U_t)] \xrightarrow{d} N\left(0, \Omega_U\right), \tag{A.109}$$

where:

$$\Omega_U \quad = \quad \lim_{T\to\infty} V\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}U_t'U_t\right) = \sum_{h=-\infty}^{\infty}Cov(U_t'U_t, U_{t-h}'U_{t-h}). \tag{A.110}$$

From equation (A.109) we get that the asymptotic distribution of $\sum_{\ell=1}^{k^c}\hat{\rho}_\ell$, under the hypothesis of $k^c$ common factors in each group is:

$$N\sqrt{T}\left[\sum_{\ell=1}^{k^c}\hat{\rho}_\ell - k^c + \frac{1}{2N}tr\left\{\tilde{\Sigma}_{cc}^{-1}\Sigma_U\right\}\right] \xrightarrow{d} N\left(0, \frac{1}{4}\Omega_U\right). \tag{A.111}$$

To conclude the proof, let us derive the expression of matrix $\Omega_U$ in equation (4.11). For this purpose, note that vector $(U_t', U_{t-h}')'$ is asymptotically Gaussian for any $h$:

$$\begin{pmatrix} U_t \\ U_{t-h} \end{pmatrix} \xrightarrow{d} N\begin{pmatrix} \Sigma_U(0) & \Sigma_U(h) \\ \Sigma_U(h)' & \Sigma_U(0) \end{pmatrix}. \tag{A.112}$$

We use the following lemma.

**LEMMA A.8.** *Let the $(n,1)$ random vector $x$ and the $(m,1)$ random vector $y$ be such that*

$$\begin{pmatrix} x \\ y \end{pmatrix} \sim N\begin{pmatrix} \Omega_{xx} & \Omega_{xy} \\ \Omega_{xy}' & \Omega_{yy} \end{pmatrix}, \tag{A.113}$$

*and let $A$ and $B$ be symmetric $(n,n)$ and $(m,m)$ matrices, respectively. Then:*

$$\begin{aligned} &i) \quad V[x'Ax] = 2tr\left\{(A\Omega_{xx})^2\right\}, \\ &ii) \quad Cov(x'Ax, y'By) = 2tr\left\{A\Omega_{xy}B\Omega_{xy}'\right\}. \end{aligned}$$

**Proof of Lemma A.8**: For point i), see Theorem 12 p. 284 in Magnus and Neudecker (2007). Point ii) is a consequence of point i) applied to vectors $x$, $y$ and $(x', y')'$, see also Theorem 10.21 in Schott (2005).

From Lemma A.8 we get (asymptotically):

$$Cov(U_t'U_t, U_{t-h}'U_{t-h}) = 2tr\left\{\Sigma_U(h)\Sigma_U(h)'\right\}, \tag{A.114}$$

and the conclusion follows.

*Q.E.D.*

# A.6 Proof of Theorems 4 and 5

## A.6.1 Asymptotic distribution of $\hat{f}_t^c$ and $\hat{f}_t^{c\,*}$

Equation (A.70) and $\hat{\Psi}_{sc} = O_p\left(\frac{1}{\sqrt{NT}}\right)$ imply:

$$\hat{W}_1^* = E_c\,\hat{\mathcal{U}} + O_p\left(\frac{1}{\sqrt{NT}}\right). \tag{A.115}$$

Recall from equation (A.71) that the normalized eigenvectors corresponding to the canonical directions are:

$$\hat{W}_1 = \hat{W}_1^*\hat{D},$$

where $\hat{D} = diag(\hat{W}_1^{*\,\prime}\hat{V}_{11}\hat{W}_1^*)^{-1/2}$. Then, we get:

$$\begin{aligned}
\hat{f}_t^c &= \hat{W}_1'\hat{h}_{1,t} \\
&= \hat{D}\hat{\mathcal{U}}'E_c'\left(h_{1,t} + \frac{1}{\sqrt{N_1}}u_{1,t}\right) + O_p\left(\frac{1}{\sqrt{NT}}\right) \\
&= \hat{D}\hat{\mathcal{U}}'\left(f_t^c + \frac{1}{\sqrt{N_1}}u_{1,t}^{(c)}\right) + O_p\left(\frac{1}{\sqrt{NT}}\right).
\end{aligned} \tag{A.116}$$

Therefore the estimated factor can be written as:

$$\hat{f}_t^c = \hat{\mathcal{H}}_c^{-1}\left(f_t^c + \frac{1}{\sqrt{N_1}}u_{1,t}^{(c)}\right) + O_p\left(\frac{1}{\sqrt{NT}}\right), \tag{A.117}$$

where $\hat{\mathcal{H}}_c^{-1} = \hat{D}\hat{\mathcal{U}}'$. Equation (A.117) implies:

$$\sqrt{N_1}\left(\hat{\mathcal{H}}_c\hat{f}_t^c - f_t^c\right) = u_{1,t}^{(c)} + o_p(1) \xrightarrow{d} N\left(0, \Sigma_{u,1}^{(cc)}\right).$$

The derivation of the asymptotic distribution of $\sqrt{N_2}\left(\hat{\mathcal{H}}_c^*\hat{f}_t^{c\,*} - f_t^c\right)$ obtained from the canonical direction $\hat{W}_2$ is analogous, and therefore is omitted.

## A.6.2 Asymptotic distribution of $\hat{\lambda}_{j,i}^c$

Let us derive the asymptotic expansion of the loading estimator $\hat{\lambda}_{j,i}^c = (\hat{F}^{c\prime}\hat{F}^c)^{-1}\hat{F}^{c\prime}y_{j,i}$, where $y_{j,i}$ is the $i$-th column of matrix $Y_j$. From equation (A.117) we can express $\hat{F}^c = [\hat{f}_1^c, ..., \hat{f}_T^c]'$ as:

$$\begin{aligned}
\hat{F}^c &= \left(F^c + \frac{1}{\sqrt{N_1}}U_1^{(c)}\right)\left(\hat{\mathcal{H}}_c^{-1}\right)' + O_p\left(\frac{1}{\sqrt{NT}}\right) \\
&= F^c\left(\hat{\mathcal{H}}_c^{-1}\right)' + \frac{1}{\sqrt{N_1}}U_1^{(c)}\left(\hat{\mathcal{H}}_c^{-1}\right)' + O_p\left(\frac{1}{\sqrt{NT}}\right),
\end{aligned} \tag{A.118}$$

where $U_1^{(c)} = [u_{1,1}^{(c)}, ..., u_{1,T}^{(c)}]'$. Equation (A.118) implies:

$$\hat{F}^c \hat{\mathcal{H}}_c' - F^c = \frac{1}{\sqrt{N_1}} U_1^{(c)} + O_p\left(\frac{1}{\sqrt{NT}}\right). \tag{A.119}$$

Then, denoting with $\xi_{j,i}$ the $i$-th column of matrix $\Xi_j$, we get:

$$\begin{aligned}
\hat{\lambda}_{j,i}^c &= (\hat{F}^{c\prime}\hat{F}^c)^{-1}\hat{F}^{c\,\prime} y_{j,i} \\
&= (\hat{F}^{c\prime}\hat{F}^c)^{-1}\hat{F}^{c\,\prime}\left(F^c\lambda_{j,i}^c + F_j^s\lambda_{j,i}^s + \varepsilon_{j,i}\right) \\
&= (\hat{F}^{c\prime}\hat{F}^c)^{-1}\hat{F}^{c\,\prime}\left[\left(F^c - \hat{F}^c\hat{\mathcal{H}}_c' + \hat{F}^c\hat{\mathcal{H}}_c'\right)\lambda_{j,i}^c + F_j^s\lambda_{j,i}^s + \varepsilon_{j,i}\right] \\
&= \hat{\mathcal{H}}_c'\lambda_{j,i}^c + (\hat{F}^{c\prime}\hat{F}^c)^{-1}\hat{F}^{c\,\prime}\varepsilon_{j,i} \\
&\quad + (\hat{F}^{c\prime}\hat{F}^c)^{-1}\hat{F}^{c\,\prime}\left(F^c - \hat{F}^c\hat{\mathcal{H}}_c'\right)\lambda_{j,i}^c + (\hat{F}^{c\prime}\hat{F}^c)^{-1}\hat{F}^{c\,\prime}F_j^s\lambda_{j,i}^s, \qquad j = 1, 2.
\end{aligned} \tag{A.120}$$

We first note that

$$\begin{aligned}
\frac{\hat{F}^{c\prime}\hat{F}^c}{T} &= \frac{1}{T}\hat{\mathcal{H}}_c^{-1}\left(F^c + \frac{1}{\sqrt{N_1}}U_1^{(c)}\right)'\left(F^c + \frac{1}{\sqrt{N_1}}U_1^{(c)}\right)(\hat{\mathcal{H}}_c^{-1})' + O_p\left(\frac{1}{\sqrt{NT}}\right) \\
&= \hat{\mathcal{H}}_c^{-1}\frac{F^{c\prime}F^c}{T}(\hat{\mathcal{H}}_c^{-1})' + \frac{1}{\sqrt{N_1}}\hat{\mathcal{H}}_c^{-1}\frac{U_1^{(c)\prime}F_t^c}{T}(\hat{\mathcal{H}}_c^{-1})' \\
&\quad + \frac{1}{\sqrt{N_1}}\hat{\mathcal{H}}_c^{-1}\frac{F^{c\prime}U_1^{(c)}}{T}(\hat{\mathcal{H}}_c^{-1})' + \frac{1}{N_1}\hat{\mathcal{H}}_c^{-1}\frac{U_1^{(c)\prime}U_1^{(c)}}{T}(\hat{\mathcal{H}}_c^{-1})' + O_p\left(\frac{1}{\sqrt{NT}}\right) \\
&= \hat{\mathcal{H}}_c^{-1}\frac{F^{c\prime}F^c}{T}(\hat{\mathcal{H}}_c^{-1})' + O_p\left(\frac{1}{\sqrt{NT}}\right),
\end{aligned}$$

where we use $\frac{1}{\sqrt{T}}F^{c\prime}U_1^{(c)} = O_p(1)$, $\frac{1}{T}U_1^{(c)\prime}U_1^{(c)} = O_p(1)$ and $T/N_1 = o(1)$. We also have:

$$\left(\frac{\hat{F}^{c\prime}\hat{F}^c}{T}\right)^{-1} = \hat{\mathcal{H}}_c'\left(\frac{F^{c\prime}F^c}{T}\right)^{-1}\hat{\mathcal{H}}_c + O_p\left(\frac{1}{\sqrt{TN}}\right). \tag{A.121}$$

Equations (A.118) and (A.119) allow to compute:

$$\begin{aligned}
\frac{1}{T}\hat{F}^{c\,\prime}\left(F^c - \hat{F}^c\hat{\mathcal{H}}_c'\right) &\simeq -\frac{1}{T\sqrt{N_1}}\hat{\mathcal{H}}_c^{-1}F^{c\,\prime}U_1^{(c)} - \frac{1}{N_1 T}\hat{\mathcal{H}}_c^{-1}U_1^{(c)\,\prime}U_1^{(c)} \\
&= O_p\left(\frac{1}{\sqrt{NT}}\right),
\end{aligned} \tag{A.122}$$

and:

$$\begin{aligned}
\frac{1}{T}\hat{F}^{c\,\prime}\varepsilon_{j,i} &= \hat{\mathcal{H}}_c^{-1}\left(\frac{1}{T}F^{c\prime}\varepsilon_{j,i} + \frac{1}{T\sqrt{N_1}}U_1^{(c)\prime}\varepsilon_{j,i}\right) \\
&= \hat{\mathcal{H}}_c^{-1}\frac{1}{T}F^{c\prime}\varepsilon_{j,i} + O_p\left(\frac{1}{\sqrt{NT}}\right).
\end{aligned} \tag{A.123}$$

We also have:

$$\frac{1}{T}\hat{F}^{c\,\prime}F_j^s = \hat{\mathcal{H}}_c^{-1}\left(\frac{1}{T}F^{c\prime}F_j^s + \frac{1}{T\sqrt{N_1}}U_1^{(c)\prime}F_j^s\right)$$
$$= \hat{\mathcal{H}}_c^{-1}\frac{1}{T}F^{c\prime}F_j^s + O_p\left(\frac{1}{\sqrt{NT}}\right). \qquad \text{(A.124)}$$

Substituting approximations (A.121) - (A.124) into equation (A.120) we get:

$$\hat{\lambda}_{j,i}^c \simeq \hat{\mathcal{H}}_c'\lambda_{j,i}^c + \hat{\mathcal{H}}_c'\left(\frac{F^{c\prime}F^c}{T}\right)^{-1}\frac{1}{T}F^{c\,\prime}\varepsilon_{j,i}$$
$$+\hat{\mathcal{H}}_c'\left(\frac{F^{c\prime}F^c}{T}\right)^{-1}\frac{1}{T}F^{c\,\prime}F_j^s\lambda_{j,i}^s + O_p\left(\frac{1}{\sqrt{NT}}\right).$$

The last equation implies:

$$\sqrt{T}\left[\left(\hat{\mathcal{H}}_c'\right)^{-1}\hat{\lambda}_{j,i}^c - \lambda_{j,i}^c\right] = \varphi_{j,i} + K_j\lambda_{j,i}^s + o_p(1), \qquad \text{(A.125)}$$

where:

$$\varphi_{j,i} = \left(\frac{F^{c\prime}F^c}{T}\right)^{-1}\frac{1}{\sqrt{T}}F^{c\,\prime}\varepsilon_{j,i}, \qquad \text{(A.126)}$$
$$K_j = \left(\frac{F^{c\prime}F^c}{T}\right)^{-1}\frac{1}{\sqrt{T}}F^{c\,\prime}F_j^s. \qquad \text{(A.127)}$$

Since $(F^{c\,\prime}F^c/T)^{-1} = I_{k^c} + o_p(1)$, the r.h.s. of equation (A.125) can be rewritten to get:

$$\sqrt{T}\left[\left(\hat{\mathcal{H}}_c'\right)^{-1}\hat{\lambda}_{j,i}^c - \lambda_{j,i}^c\right] = \frac{1}{\sqrt{T}}\sum_{t=1}^{T} f_t^c(\varepsilon_{j,it} + f_{j,t}^{s\,\prime}\lambda_{j,i}^s) + o_p(1) \equiv w_{j,i}^c + o_p(1). \qquad \text{(A.128)}$$

Then, since the errors and the factors are independent (Assumption A.5 *a)*) , a CLT for weakly dependent data yields equation (4.19).

### A.6.3 Asymptotic distribution of $\hat{f}_{j,t}^s$ and $\hat{\lambda}_{j,i}^s$

Let us now derive the asymptotic expansion of term $\hat{f}_{j,t}^s$. We start by computing the asymptotic expansion of the regression residuals $y_{j,it} - \hat{f}_t^{c\,\prime}\hat{\lambda}_{j,i}^c$:

$$y_{j,it} - \hat{f}_t^{c\,\prime}\hat{\lambda}_{j,i}^c = f_{j,t}^{s\,\prime}\lambda_{j,i}^s + \varepsilon_{j,it} - \left(\hat{f}_t^{c\,\prime}\hat{\lambda}_{j,i}^c - f_t^{c\,\prime}\lambda_{j,i}^c\right)$$
$$= f_{j,t}^{s\,\prime}\lambda_{j,i}^s + \varepsilon_{j,it} - \left[\left(f_t^c + \frac{1}{\sqrt{N_1}}u_{1,t}^{(c)}\right)'\left(\lambda_{j,i}^c + \frac{1}{\sqrt{T}}\varphi_{j,i} + \frac{1}{\sqrt{T}}K_j\lambda_{j,i}^s\right) - f_t^{c\,\prime}\lambda_{j,i}^c\right]$$
$$\simeq g_{j,t}'\lambda_{j,i}^s + e_{j,it}, \qquad \text{(A.129)}$$

where:

$$g_{j,t} := f_{j,t}^s - \frac{1}{\sqrt{T}}K_j'f_t^c = f_{j,t}^s - (F_j^{s\,\prime}F^c)(F^{c\,\prime}F^c)^{-1}f_t^c, \qquad \text{(A.130)}$$
$$e_{j,it} := \varepsilon_{j,it} - \frac{1}{\sqrt{T}}f_t^{c\prime}\varphi_{j,i}. \qquad \text{(A.131)}$$

Then, the residuals $y_{j,it} - \hat{f}_t^{c\,\prime}\hat{\lambda}_{j,i}^c$ satisfy an approximate factor structure with factors $g_{j,t}$ and errors $e_{j,it}$. From asymptotic theory of the PC estimators in large panels, we know that:

$$\sqrt{N}\left[\hat{\mathcal{H}}_{s,j}\hat{f}_{j,t}^s - g_{j,t}\right] \;=\; v_{j,t}^{*s} + o_p(1), \qquad j = 1, 2, \tag{A.132}$$

where $\hat{\mathcal{H}}_{s,j}$, $j = 1, 2$, is a non-singular matrix and:

$$\begin{aligned} v_{j,t}^{*s} &= \left(\frac{\Lambda_j^{s\prime}\Lambda_j^s}{N_j}\right)^{-1}\frac{1}{\sqrt{N_j}}\Lambda_j^{s\prime}e_{j,t} \\ &= \left(\frac{\Lambda_j^{s\prime}\Lambda_j^s}{N_j}\right)^{-1}\frac{1}{\sqrt{N_j}}\sum_{i=1}^{N_j}\lambda_{j,i}^s\varepsilon_{j,it} - \left(\frac{\Lambda_j^{s\prime}\Lambda_j^s}{N_j}\right)^{-1}\frac{1}{\sqrt{NT}}\sum_{i=1}^{N_j}\lambda_{j,i}^s f_t^{c\prime}\left(\frac{1}{\sqrt{T}}\sum_{r=1}^{T}f_r^c\varepsilon_{j,ir}\right) \\ &= \left(\frac{\Lambda_j^{s\prime}\Lambda_j^s}{N_j}\right)^{-1}\frac{1}{\sqrt{N_j}}\sum_{i=1}^{N_j}\lambda_{j,i}^s\varepsilon_{j,it} + o_p(1). \end{aligned}$$

Therefore we have

$$\sqrt{N}\left[\hat{\mathcal{H}}_{s,j}\hat{f}_{j,t}^s - (f_{j,t}^s - (F_j^{s\,\prime}F^c)(F^{c\,\prime}F^c)^{-1}f_t^c)\right] \;=\; v_{j,t}^s + o_p(1), \qquad j = 1, 2, \tag{A.133}$$

where $v_{j,t}^s = \left(\frac{\Lambda_j^{s\prime}\Lambda_j^s}{N_j}\right)^{-1}\frac{1}{\sqrt{N_j}}\sum_{i=1}^{N_j}\lambda_{j,i}^s\varepsilon_{j,it}$, which proves equation (4.16).

From asymptotic theory of the PC estimators in large panels, we also know that the following result must hold for the loadings estimator of factor model (A.129):

$$\sqrt{T}\left[\left(\hat{\mathcal{H}}_{s,j}'\right)^{-1}\hat{\lambda}_{j,i}^s - \lambda_{j,i}^s\right] \;=\; w_{j,i}^{*s} + o_p(1), \qquad j = 1, 2 \tag{A.134}$$

where $\hat{\mathcal{H}}_{s,j}$, $j = 1, 2$ are the same non-singular matrices in equation (A.132), and

$$\begin{aligned} w_{j,i}^{*s} &= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(f_{j,t}^s + \frac{1}{\sqrt{T}}K_j' f_t^c\right)e_{j,it}, \\ &= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(f_{j,t}^s + \frac{1}{\sqrt{T}}K_j' f_t^c\right)\left(\varepsilon_{j,it} - \frac{1}{\sqrt{T}}f_t^{c\prime}\varphi_{j,i}\right), \\ &= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}f_{j,t}^s\varepsilon_{j,it} - \frac{1}{T}\sum_{t=1}^{T}f_{j,t}^s f_t^{c\prime}\varphi_{j,i} \\ &\quad + K_j'\frac{1}{T}\sum_{t=1}^{T}f_t^{c\prime}\varepsilon_{j,it} - K_j'\frac{1}{T\sqrt{T}}\sum_{t=1}^{T}f_t^c f_t^{c\prime}\varphi_{j,i} \\ &= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}f_{j,t}^s\varepsilon_{j,it} + o_p(1), \end{aligned} \tag{A.135}$$

since $\frac{1}{T}\sum_{t=1}^{T}f_{j,t}^s f_t^{c\prime} = o_p(1)$. Therefore, we get:

$$\sqrt{T}\left[\left(\hat{\mathcal{H}}_{s,j}'\right)^{-1}\hat{\lambda}_{j,i}^s - \lambda_{j,i}^s\right] = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}f_{j,t}^s\varepsilon_{j,it} + o_p(1) \equiv w_{j,i}^s + o_p(1), \tag{A.136}$$

which yields equation (4.20).

*Q.E.D.*

# A.7 Proof of Theorem 6

Theorem 6 follows from Theorem 3 since we have:

$$tr\left\{\hat{\Sigma}_{cc}^{-1}\hat{\Sigma}_U\right\} = tr\left\{\tilde{\Sigma}_{cc}^{-1}\Sigma_{U,N}\right\} + o_p(1/\sqrt{T}), \tag{A.137}$$

$$tr\left\{\hat{\Sigma}_U^2\right\} = tr\left\{\Sigma_U(0)^2\right\} + o_p(1). \tag{A.138}$$

These expansions are proved next.

## A.7.1 Asymptotic expansion of $\hat{\Sigma}_{cc}^{-1}$

Substituting the expression of $\hat{f}_t^c$ from equation (A.117) into $\hat{\Sigma}_{cc} = \frac{1}{T}\sum_{t=1}^{T}\hat{f}_t^c\hat{f}_t^{c\prime}$ we get:

$$\begin{aligned}\hat{\Sigma}_{cc} &= \frac{1}{T}\sum_{t=1}^{T}\hat{\mathcal{H}}_c^{-1}\left(f_t^c + \frac{1}{\sqrt{N_1}}u_{j,t}^{(c)}\right)\left(f_t^c + \frac{1}{\sqrt{N_1}}u_{j,t}^{(c)}\right)'\left(\hat{\mathcal{H}}_c^{-1}\right)' + O_p\left(\frac{1}{\sqrt{NT}}\right)\\ &= \hat{\mathcal{H}}_c^{-1}\tilde{\Sigma}_{cc}\left(\hat{\mathcal{H}}_c^{-1}\right)' + O_p\left(\frac{1}{\sqrt{NT}}\right).\end{aligned}$$

This implies:

$$\hat{\Sigma}_{cc}^{-1} = \hat{\mathcal{H}}_c'\tilde{\Sigma}_{cc}^{-1}\hat{\mathcal{H}}_c + O_p\left(\frac{1}{\sqrt{NT}}\right). \tag{A.139}$$

## A.7.2 Asymptotic expansion of $\hat{\Sigma}_U$

### i) Asymptotic expansion of $\frac{\hat{\Lambda}_j'\hat{\Lambda}_j}{N_j}$

To derive the asymptotic expansion of matrix $\hat{\Lambda}_j'\hat{\Lambda}_j/N_j$, it is useful to write the matrix versions of the quantities defined in equations (A.128) and (A.136). Stacking the loadings $\hat{\lambda}_{j,i}^c$ in matrix $\hat{\Lambda}_j^c = [\hat{\lambda}_{j,1}^c, ..., \hat{\lambda}_{j,N}^c]'$ we get:

$$\hat{\Lambda}_j^c = \left[\Lambda_j^c + \frac{1}{\sqrt{T}}G_j^c\right]\hat{\mathcal{H}}_c + o_p\left(\frac{1}{\sqrt{T}}\right),$$

where

$$G_j^c = \frac{1}{\sqrt{T}}\varepsilon_j'F^c + \Lambda_j^s\left(\frac{1}{\sqrt{T}}F_j^{s\prime}F^c\right) \tag{A.140}$$

$$= \frac{1}{\sqrt{T}}\varepsilon_j'F^c + \Lambda_j^s\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}f_{j,t}^s f_t^{c\prime}\right). \tag{A.141}$$

Similarly, stacking the loadings $\hat{\lambda}_{j,i}^s$ in matrix $\hat{\Lambda}_j^s = [\hat{\lambda}_{j,1}^s, ..., \hat{\lambda}_{j,N}^s]'$ we get:

$$\hat{\Lambda}_j^s = \left[\Lambda_j^c + \frac{1}{\sqrt{T}}G_j^s\right]\hat{\mathcal{H}}_{j,s} + o_p\left(\frac{1}{\sqrt{T}}\right),$$

where

$$G_j^s = \frac{1}{\sqrt{T}}\varepsilon_j' F_j^s. \tag{A.142}$$

By gathering these expansions, we get:

$$\hat{\Lambda}_j \simeq \left(\Lambda_j + \frac{1}{\sqrt{T}}G_j\right)\hat{\mathcal{U}}_j, \qquad j = 1, 2, \tag{A.143}$$

where

$$G_j = \left[\begin{array}{ccc} G_j^c & \vdots & G_j^s \end{array}\right], \tag{A.144}$$

$$\hat{\mathcal{U}}_j = \left[\begin{array}{cc} \hat{\mathcal{H}}_c & 0 \\ 0 & \hat{\mathcal{H}}_{s,j} \end{array}\right]. \tag{A.145}$$

We start by computing the asymptotic expansion of $\frac{\hat{\Lambda}_j'\hat{\Lambda}_j}{N_j}$. From Assumptions A.1, A.2 and A.5 we get:

$$\frac{1}{N_j}\left[\Lambda_j + \frac{1}{\sqrt{T}}G_j\right]'\left[\Lambda_j + \frac{1}{\sqrt{T}}G_j\right] \simeq \frac{1}{N_j}\Lambda_j'\Lambda_j + \frac{1}{N\sqrt{T}}\left(\Lambda_j'G_j + G_j'\Lambda_j\right) + \frac{1}{NT}G_j'G_j. \tag{A.146}$$

Let us compute the asymptotic expansion of $\frac{1}{N\sqrt{T}}\Lambda_j'G_j$:

$$\frac{1}{N_j\sqrt{T}}\Lambda_j'G_j = \frac{1}{N_j\sqrt{T}}\left[\begin{array}{cc} \Lambda_j^{c\prime}G_j^c & \Lambda_j^{c\prime}G_j^s \\ \Lambda_j^{s\prime}G_j^c & \Lambda_j^{s\prime}G_j^s \end{array}\right]. \tag{A.147}$$

Using equation (A.140) we get:

$$\begin{aligned}\frac{1}{N_j\sqrt{T}}\Lambda_j^{c\prime}G_j^c &= \frac{1}{N_j\sqrt{T}}\Lambda_j^{c\prime}\left[\frac{1}{\sqrt{T}}\varepsilon_j'F^c + \Lambda_j^s\left(\frac{1}{\sqrt{T}}F_j^{s\prime}F^c\right)\right] \\ &= \frac{1}{N_jT}\Lambda_j^{c\prime}\varepsilon_j'F^c + \frac{1}{N_jT}\Lambda_j^{c\prime}\Lambda_j^s\left(F_j^{s\prime}F^c\right) \\ &= \left(\frac{\Lambda_j^{c\prime}\Lambda_j^s}{N_j}\right)\frac{1}{T}\sum_{t=1}^{T} f_{j,t}^s f_t^{c\prime} + O_p\left(\frac{1}{\sqrt{N_jT}}\right),\end{aligned} \tag{A.148}$$

Using analogous arguments and equation (A.142), we get:

$$\frac{1}{N_j\sqrt{T}}\Lambda_j^{s\prime}G_j^c = \left(\frac{\Lambda_j^{s\prime}\Lambda_j^s}{N_j}\right)\frac{1}{T}\sum_{t=1}^{T} f_{j,t}^s f_t^{c\prime} + O_p\left(\frac{1}{\sqrt{N_jT}}\right), \tag{A.149}$$

$$\frac{1}{N_j\sqrt{T}}\Lambda_j^{c\prime}G_j^s = \frac{1}{N_j\sqrt{T}}\Lambda_j^{c\prime}\varepsilon_j'F^s = O_p\left(\frac{1}{\sqrt{N_jT}}\right), \tag{A.150}$$

$$\frac{1}{N_j\sqrt{T}}\Lambda_j^{s\prime}G_j^s = \frac{1}{N_j\sqrt{T}}\Lambda_j^{s\prime}\varepsilon_j'F^s = O_p\left(\frac{1}{\sqrt{N_jT}}\right). \tag{A.151}$$

The last four equations imply:

$$\frac{1}{N_j\sqrt{T}}\Lambda_j' G_j = \begin{bmatrix} \left(\frac{\Lambda_j^{c\prime}\Lambda_j^s}{N_j}\right)\frac{1}{T}\sum_{t=1}^T f_{j,t}^s f_t^{c\prime} & 0 \\ \left(\frac{\Lambda_j^{s\prime}\Lambda_j^s}{N_j}\right)\frac{1}{T}\sum_{t=1}^T f_{j,t}^s f_t^{c\prime} & 0 \end{bmatrix} + O_p\left(\frac{1}{\sqrt{N_j T}}\right)$$

$$= \left[ \left(\frac{\Lambda_j'\Lambda_j^s}{N_j}\right)\frac{1}{T}\sum_{t=1}^T f_{j,t}^s f_t^{c\prime} \ \vdots \ 0_{(k_j\times k_j^s)} \right] + O_p\left(\frac{1}{\sqrt{N_j T}}\right). \quad (A.152)$$

Using analogous arguments, we have:

$$\frac{1}{N_j T}G_j^{c\prime}G_j^c = \frac{1}{N_j T}\left[\frac{1}{\sqrt{T}}\varepsilon_j' F^c + \Lambda_j^s\left(\frac{1}{\sqrt{T}}F_j^{s\prime}F^c\right)\right]'\left[\frac{1}{\sqrt{T}}\varepsilon_j' F^c + \Lambda_j^s\left(\frac{1}{\sqrt{T}}F_j^{s\prime}F^c\right)\right]$$

$$= o_p\left(\frac{1}{\sqrt{T}}\right) \quad (A.153)$$

and

$$\frac{1}{N_j T}G_j' G_j = o_p\left(\frac{1}{\sqrt{T}}\right). \quad (A.154)$$

Substituting (A.152) and (A.154) into equation (A.146) we get:

$$\frac{1}{N_j}\left[\Lambda_j + \frac{1}{\sqrt{T}}G_j\right]'\left[\Lambda_j + \frac{1}{\sqrt{T}}G_j\right] \simeq \Sigma_{\Lambda,j} + \frac{1}{\sqrt{T}}\left(L_{1,j} + L_{1,j}'\right) + O_p\left(\frac{1}{\sqrt{N}}\right) \quad (A.155)$$

where

$$L_{1,j} = \left[ \left(\frac{\Lambda_j'\Lambda_j^s}{N_j}\right)\left(\frac{1}{\sqrt{T}}F_j^{s\prime}F^c\right) \ \vdots \ 0_{(k_j\times k_j^s)} \right]. \quad (A.156)$$

Therefore we have:

$$\frac{\hat{\Lambda}_j'\hat{\Lambda}_j}{N_j} = \hat{\mathcal{U}}_j'\left[\Sigma_{\Lambda,j} + \frac{1}{\sqrt{T}}\left(L_{1,j} + L_{1,j}'\right)\right]\hat{\mathcal{U}}_j + o_p\left(\frac{1}{\sqrt{T}}\right). \quad (A.157)$$

## ii) Asymptotic expansion of $\hat{\Gamma}_j$

The approximations in Propositions 4 and 5 allow to compute the asymptotic expansion of $\hat{\varepsilon}_{j,it}$:

$$\begin{aligned} \hat{\varepsilon}_{j,it} &= y_{j,it} - \hat{\lambda}_{j,i}^{c\,\prime}\hat{f}_t^c - \hat{\lambda}_{j,i}^{s\,\prime}\hat{f}_{j,t}^s \\ &= \varepsilon_{j,it} - \left[\hat{\lambda}_{j,i}^{c\,\prime}\hat{f}_t^c - \lambda_{j,i}^{c\,\prime}f_t^c\right] - \left[\hat{\lambda}_{j,i}^{s\,\prime}\hat{f}_{j,t}^s - \lambda_{j,i}^{s\,\prime}f_{j,t}^s\right] \\ &\simeq \varepsilon_{j,it} - \left[\left(\lambda_{j,i}^c + \frac{1}{\sqrt{T}}w_{j,i}^c\right)'\left(f_t^c + \frac{1}{\sqrt{N_1}}u_{1,t}^{(c)}\right) - \lambda_{j,i}^{c\,\prime}f_t^c\right] \\ &\quad - \left[\left(\lambda_{j,i}^s + \frac{1}{\sqrt{T}}w_{j,i}^s\right)'\left(f_{j,t}^s - \frac{1}{\sqrt{T}}K_j' f_t^c + \frac{1}{\sqrt{N_j}}v_{j,t}^s\right) - \lambda_{j,i}^{s\,\prime}f_{j,t}^s\right] \\ &\simeq \varepsilon_{j,it} - \left(\frac{1}{\sqrt{N_1}}\lambda_{j,i}^{c\,\prime}u_{1,t}^{(c)} + \frac{1}{\sqrt{T}}w_{j,i}^{c\prime}f_t^c\right) - \left(\frac{1}{\sqrt{N_j}}\lambda_{j,i}^{s\,\prime}v_{j,t}^s + \frac{1}{\sqrt{T}}w_{j,i}^{s\prime}f_{j,t}^s\right) \\ &\quad + \lambda_{j,i}^{s\,\prime}\frac{1}{\sqrt{T}}K_j' f_t^c. \end{aligned} \quad (A.158)$$

Since $T/N_j = o(1)$, we keep only the terms of order $1/\sqrt{T}$ in equation (A.158), and we get:

$$\hat{\varepsilon}_{j,it} = \varepsilon_{j,it} - \frac{1}{\sqrt{T}}\left(w_{j,i}^{c\prime} f_t^c + w_{j,i}^{s\prime} f_{j,t}^s\right) + \lambda_{j,i}^{s\,\prime} \frac{1}{\sqrt{T}} K_j' f_t^c + o_p\left(\frac{1}{\sqrt{T}}\right). \tag{A.159}$$

From the definition of $w_{j,i}^c$ in Proposition 5 we get:

$$w_{j,i}^{c\prime} f_t^c = \frac{1}{\sqrt{T}}\left(\sum_{r=1}^{T} \varepsilon_{j,ir} f_r^{c\prime}\right) f_t^c + \lambda_{j,i}^{s\prime} K_j' f_t^c, \tag{A.160}$$

which implies:

$$\hat{\varepsilon}_{j,it} = \varepsilon_{j,it} - \frac{1}{\sqrt{T}}\left(\tilde{w}_{j,i}^{c\prime} f_t^c + w_{j,i}^{s\prime} f_{j,t}^s\right) + o_p\left(\frac{1}{\sqrt{T}}\right), \tag{A.161}$$

where:

$$\tilde{w}_{j,i}^c = \frac{1}{\sqrt{T}} \sum_{r=1}^{T} f_r^c \varepsilon_{j,ir}. \tag{A.162}$$

Equation (A.159) allows us to compute:

$$\begin{aligned}
\hat{\gamma}_{j,ii} &= \frac{1}{T}\sum_{t=1}^{T} \hat{\varepsilon}_{j,it}^2 \\
&\simeq \frac{1}{T}\sum_{t=1}^{T}\left[\varepsilon_{j,it} - \frac{1}{\sqrt{T}}\left(\tilde{w}_{j,i}^{c\prime} f_t^c + w_{j,i}^{s\prime} f_{j,t}^s\right)\right]^2 \\
&= \frac{1}{T}\sum_{t=1}^{T} \varepsilon_{j,it}^2 - \frac{2}{T\sqrt{T}}\sum_{t=1}^{T} \varepsilon_{j,it}\left(\tilde{w}_{j,i}^{c\,\prime} f_t^c + w_{j,i}^{s\,\prime} f_{j,t}^s\right) + \frac{1}{T^2}\sum_{t=1}^{T}\left(\tilde{w}_{j,i}^{c\,\prime} f_t^c + w_{j,i}^{s\,\prime} f_{j,t}^s\right)^2.
\end{aligned} \tag{A.163}$$

Using $\frac{1}{\sqrt{T}}\sum_{t=1}^{T} \varepsilon_{j,it} f_t^c = O_p(1)$ and $\frac{1}{\sqrt{T}}\sum_{t=1}^{T} \varepsilon_{j,it} f_{j,t}^s = O_p(1)$ we get:

$$\hat{\gamma}_{j,ii} = \frac{1}{T}\sum_{t=1}^{T} \varepsilon_{j,it}^2 + O_p\left(\frac{1}{T}\right), \tag{A.164}$$

which implies:

$$\begin{aligned}
\hat{\gamma}_{j,ii} &= \frac{1}{T}\sum_{t=1}^{T} \varepsilon_{j,it}^2 + o_p\left(\frac{1}{\sqrt{T}}\right) \\
&= \gamma_{j,ii} + \frac{1}{\sqrt{T}} w_{j,i} + o_p\left(\frac{1}{\sqrt{T}}\right),
\end{aligned} \tag{A.165}$$

where

$$w_{j,i} = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}(\varepsilon_{j,it}^2 - \gamma_{j,ii}) = O_p(1), \tag{A.166}$$

from Assumptions A.4 and A.7. Therefore, we have:

$$\hat{\Gamma}_j \simeq \Gamma_j + \frac{1}{\sqrt{T}} W_j. \tag{A.167}$$

where $\Gamma_j = diag(\gamma_{j,ii},\ i = 1, ..., N)$ and $W_j = diag(w_{j,i},\ i = 1, ..., N)$, for $j = 1, 2$.

## iii) Asymptotic expansion of $\frac{1}{N_j}\hat{\Lambda}_j'\hat{\Gamma}_j\hat{\Lambda}_j$

Let us define

$$\begin{aligned}
\hat{\Omega}_j^* &:= \frac{1}{N_j}\left(\Lambda_j + \frac{1}{\sqrt{T}}G_j\right)'\hat{\Gamma}\left(\Lambda_j + \frac{1}{\sqrt{T}}G_j\right) \\
&= \frac{1}{N_j}\left(\Lambda_j + \frac{1}{\sqrt{T}}G_j\right)'\left(\Gamma_j + \frac{1}{\sqrt{T}}W_j\right)\left(\Lambda_j + \frac{1}{\sqrt{T}}G_j\right) \\
&= \frac{1}{N_j}\Lambda_j'\Gamma_j\Lambda_j + \hat{\Omega}_{j,I}^* + \hat{\Omega}_{j,II}^* + \hat{\Omega}_{j,III}^* + \hat{\Omega}_{j,II}^{*}{}' + \hat{\Omega}_{j,III}^{*}{}' + \hat{\Omega}_{j,IV}^* + \hat{\Omega}_{j,V}^*, \qquad \text{(A.168)}
\end{aligned}$$

where

$$\hat{\Omega}_{j,I}^* = \frac{1}{N_j\sqrt{T}}\Lambda_j'W_j\Lambda_j = O_p\left(\frac{1}{\sqrt{NT}}\right), \tag{A.169}$$

$$\hat{\Omega}_{j,III}^* = \frac{1}{N_jT}\Lambda_j'W_jG_j = O_p\left(\frac{1}{T}\right), \tag{A.170}$$

$$\hat{\Omega}_{j,IV}^* = \frac{1}{N_jT}G_j'\Gamma_jG_j = O_p\left(\frac{1}{T}\right), \tag{A.171}$$

$$\hat{\Omega}_{j,V}^* = \frac{1}{N_jT\sqrt{T}}G_j'W_jG_j = O_p\left(\frac{1}{T\sqrt{T}}\right). \tag{A.172}$$

Moreover, similarly as for (A.152) we have:

$$\begin{aligned}
\hat{\Omega}_{j,II}^* &= \frac{1}{N_j\sqrt{T}}\Lambda_j'\Gamma_jG_j \\
&= \left[\ \frac{1}{N_j}\Lambda_j'\Gamma_j\Lambda_j^s\left(\frac{1}{T}\sum_{t=1}^{T} f_{j,t}^s f_t^{c\prime}\right)\ \vdots\ 0_{(k_j\times k_j^s)}\ \right] + o_p\left(\frac{1}{\sqrt{T}}\right), \qquad \text{(A.173)} \\
&= \frac{1}{\sqrt{T}}\left[\ \left(\frac{1}{N_j}\Lambda_j'\Gamma_j\Lambda_j^s\right)\left(\frac{1}{\sqrt{T}}F_j^s{}'F_c\right)\ \vdots\ 0_{(k_j\times k_j^s)}\ \right] + o_p\left(\frac{1}{\sqrt{T}}\right), \qquad \text{(A.174)} \\
&= \frac{1}{\sqrt{T}}L_{2,j} + o_p\left(\frac{1}{\sqrt{T}}\right), \qquad \text{(A.175)} \\
& \qquad \text{(A.176)}
\end{aligned}$$

where

$$L_{2,j} = \left[\ \left(\frac{1}{N}\Lambda_j'\Gamma_j\Lambda_j^s\right)\left(\frac{1}{\sqrt{T}}F_j^s{}'F_c\right)\ \vdots\ 0_{(k_j\times k_j^s)}\ \right].$$

Collecting the previous results, using $T/N = o_p(1)$, and defining $\Omega_j^* = \lim_{N\to\infty} \frac{1}{N}\Lambda_j'\Gamma_j\Lambda_j$ we get:

$$\begin{aligned}\hat{\Omega}_j^* &= \frac{1}{N}\Lambda_j'\Gamma_j\Lambda_j + \frac{1}{\sqrt{T}}\left(L_{2,j} + L_{2,j}'\right) + o_p\left(\frac{1}{\sqrt{T}}\right) \\ &= \Omega_j^* + \frac{1}{\sqrt{T}}\left(L_{2,j} + L_{2,j}'\right) + o_p\left(\frac{1}{\sqrt{T}}\right). \qquad \text{(A.177)}\end{aligned}$$

Substituting equation (A.143) into $\dfrac{1}{N_j}\hat{\Lambda}_j'\hat{\Gamma}_j\hat{\Lambda}_j$, and using equation (A.177) we get:

$$\begin{aligned}\hat{\Omega}_j &= \hat{\mathcal{U}}_j'\hat{\Omega}_j^*\hat{\mathcal{U}}_j \\ &= \hat{\mathcal{U}}_j'\left[\Omega_j^* + \frac{1}{\sqrt{T}}\left(L_{2,j} + L_{2,j}'\right)\right]\hat{\mathcal{U}}_j + o_p\left(\frac{1}{\sqrt{T}}\right), \qquad j = 1,2. \qquad \text{(A.178)}\end{aligned}$$

## iv) Asymptotic expansion of $\hat{\Sigma}_U$

The estimator of $\Sigma_{u,j}$ is given in equation (4.22). Equation (A.157) allows to compute the asymptotic approximation of $\left(\dfrac{\hat{\Lambda}_j'\hat{\Lambda}_j}{N_j}\right)^{-1}$:

$$\left(\frac{\hat{\Lambda}_j'\hat{\Lambda}_j}{N_j}\right)^{-1} \simeq \hat{\mathcal{U}}_j^{-1}\left[\Sigma_{\Lambda,j}^{-1} - \frac{1}{\sqrt{T}}\Sigma_{\Lambda,j}^{-1}\left(L_{1,j} + L_{1,j}'\right)\Sigma_{\Lambda,j}^{-1}\right]\left(\hat{\mathcal{U}}_j'\right)^{-1}. \qquad \text{(A.179)}$$

Substituting equations (A.179) and (A.178) into equation (4.22), we get:

$$\begin{aligned}\hat{\Sigma}_{u,j} &\simeq \hat{\mathcal{U}}_j^{-1}\left[\Sigma_{\Lambda,j}^{-1} - \frac{1}{\sqrt{T}}\Sigma_{\Lambda,j}^{-1}\left(L_{1,j} + L_{1,j}'\right)\Sigma_{\Lambda,j}^{-1}\right]\left[\Omega_j^* + \frac{1}{\sqrt{T}}\left(L_{2,j} + L_{2,j}'\right)\right] \\ &\quad \times\left[\Sigma_{\Lambda,j}^{-1} - \frac{1}{\sqrt{T}}\Sigma_{\Lambda,j}^{-1}\left(L_{1,j} + L_{1,j}'\right)\Sigma_{\Lambda,j}^{-1}\right]\left(\hat{\mathcal{U}}_j'\right)^{-1} \\ &\simeq \hat{\mathcal{U}}_j^{-1}\Sigma_{\Lambda,j}^{-1}\left[I - \frac{1}{\sqrt{T}}\left(L_{1,j} + L_{1,j}'\right)\Sigma_{\Lambda,j}^{-1}\right]\left[\Omega_j^* + \frac{1}{\sqrt{T}}\left(L_{2,j} + L_{2,j}'\right)\right] \\ &\quad \times\left[I - \frac{1}{\sqrt{T}}\Sigma_{\Lambda,j}^{-1}\left(L_{1,j} + L_{1,j}'\right)\right]\Sigma_{\Lambda,j}^{-1}\left(\hat{\mathcal{U}}_j'\right)^{-1} \\ &\simeq \hat{\mathcal{U}}_j^{-1}\Sigma_{\Lambda,j}^{-1}\left[\Omega_j^* + \frac{1}{\sqrt{T}}\left(L_{2,j} + L_{2,j}'\right) - \frac{1}{\sqrt{T}}\Omega_j^*\Sigma_{\Lambda,j}^{-1}\left(L_{1,j} + L_{1,j}'\right) - \frac{1}{\sqrt{T}}\left(L_{1,j} + L_{1,j}'\right)\Sigma_{\Lambda,j}^{-1}\Omega_j^*\right] \\ &\quad \times\Sigma_{\Lambda,j}^{-1}\left(\hat{\mathcal{U}}_j'\right)^{-1},\end{aligned}$$

which implies:

$$\hat{\Sigma}_{u,j} = \hat{\mathcal{U}}_j^{-1}\Sigma_{u,j}\left(\hat{\mathcal{U}}_j'\right)^{-1} + \frac{1}{\sqrt{T}}\hat{\mathcal{U}}_j^{-1}L_{3,j}\left(\hat{\mathcal{U}}_j'\right)^{-1} + o_p\left(\frac{1}{\sqrt{T}}\right),$$

where

$$L_{3,j} = \Sigma_{\Lambda,j}^{-1}\left[\left(L_{2,j} + L_{2,j}'\right) - \Omega_j^*\Sigma_{\Lambda,j}^{-1}\left(L_{1,j} + L_{1,j}'\right) - \left(L_{1,j} + L_{1,j}'\right)\Sigma_{\Lambda,j}^{-1}\Omega_j^*\right]\Sigma_{\Lambda,j}^{-1}. \qquad \text{(A.180)}$$

From equation (A.145) we have:

$$\begin{aligned}\hat{\Sigma}_U &= \mu_N^2 \hat{\Sigma}_{u,1}^{(cc)} + \hat{\Sigma}_{u,2}^{(cc)} \\ &= \hat{\mathcal{H}}_c^{-1}\left[\mu_N^2\Sigma_{u,1}+\Sigma_{u,2}\right]^{(cc)}\left(\hat{\mathcal{H}}_c'\right)^{-1} + \frac{1}{\sqrt{T}}\hat{\mathcal{H}}_c^{-1}\left(\mu_N^2 L_{3,1}+L_{3,2}\right)^{(cc)}\left(\hat{\mathcal{H}}_c'\right)^{-1} + o_p\left(\frac{1}{\sqrt{T}}\right) \\ &= \hat{\mathcal{H}}_c^{-1}\Sigma_{U,N}\left(\hat{\mathcal{H}}_c'\right)^{-1} + \frac{1}{\sqrt{T}}\hat{\mathcal{H}}_c^{-1}\left(\mu_N^2 L_{3,1}+L_{3,2}\right)^{(cc)}\left(\hat{\mathcal{H}}_c'\right)^{-1} + o_p\left(\frac{1}{\sqrt{T}}\right). \end{aligned} \tag{A.181}$$

This expansion, the convergence $\Sigma_{U,N} \to \Sigma_U(0)$ and the commutative property of the trace, imply equation (A.138).

### A.7.3 Asymptotic expansion of $tr\left\{\tilde{\Sigma}_{cc}^{-1}\hat{\Sigma}_U\right\}$

Results (A.139) and (A.181), and the commutative property of the trace, imply:

$$tr\left\{\hat{\Sigma}_{cc}^{-1}\hat{\Sigma}_U\right\} = tr\left\{\tilde{\Sigma}_{cc}^{-1}\Sigma_{U,N}\right\} + \frac{1}{\sqrt{T}}tr\left\{\tilde{\Sigma}_{cc}^{-1}\left(\mu_N^2 L_{3,1}+L_{3,2}\right)^{(cc)}\right\} + o_p\left(\frac{1}{\sqrt{T}}\right).$$

Noting that $L_{3,j} = O_p(1)$, for $j = 1, 2$, and recalling that $\tilde{\Sigma}_{cc} = I_{k^c} + O_p(1/\sqrt{T})$ and $\mu_N = \mu + o(1)$, the last equation can be further simplified to

$$tr\left\{\hat{\Sigma}_{cc}^{-1}\hat{\Sigma}_U\right\} = tr\left\{\tilde{\Sigma}_{cc}^{-1}\Sigma_{U,N}\right\} + \frac{1}{\sqrt{T}}tr\left\{\left(\mu^2 L_{3,1}+L_{3,2}\right)^{(cc)}\right\} + o_p\left(\frac{1}{\sqrt{T}}\right). \tag{A.182}$$

Let us compute $L_{3,j}$ explicitly. From equation (A.156) we get:

$$\begin{aligned} L_{1,j} &= \begin{bmatrix} \left(\frac{\Lambda_j^{c\prime}\Lambda_j^s}{N}\right)\left(\frac{1}{\sqrt{T}}F_j^{s\prime}F^c\right) & 0_{(k^c\times k_j^s)} \\ \left(\frac{\Lambda_j^{s\prime}\Lambda_j^s}{N}\right)\left(\frac{1}{\sqrt{T}}F_j^{s\prime}F^c\right) & 0_{(k_j^s\times k_j^s)} \end{bmatrix} \\ &= \begin{bmatrix} \Sigma_{\Lambda,j,cs}\left(\frac{1}{\sqrt{T}}F_j^{s\prime}F^c\right) & 0 \\ \Sigma_{\Lambda,j,ss}\left(\frac{1}{\sqrt{T}}F_j^{s\prime}F^c\right) & 0 \end{bmatrix} + O_p\left(\frac{1}{\sqrt{N}}\right) \\ &= \Sigma_{\Lambda,j}\begin{bmatrix} 0_{(k^c\times k^c)} & 0_{(k^c\times k_j^s)} \\ K_j' & 0_{(k_j^s\times k_j^s)} \end{bmatrix} + O_p\left(\frac{1}{\sqrt{N}}\right). \end{aligned} \tag{A.183}$$

Equation (A.183) implies:

$$\Omega_j^*\Sigma_{\Lambda,j}^{-1}L_{1,j} = L_{2,j} + O_p\left(\frac{1}{\sqrt{N}}\right). \tag{A.184}$$

Substituting results (A.183) and (A.184) into equation (A.180) we get:

$$\begin{aligned} L_{3,j} &= -\Sigma_{\Lambda,j}^{-1}\left[\Omega_j^*\Sigma_{\Lambda,j}^{-1}L_{1,j}' + L_{1,j}\Sigma_{\Lambda,j}^{-1}\Omega_j^*\right]\Sigma_{\Lambda,j}^{-1} \\ &= -\Sigma_{u,j}L_{1,j}'\Sigma_{\Lambda,j}^{-1} - \Sigma_{\Lambda,j}^{-1}L_{1,j}\Sigma_{u,j} + O_p\left(\frac{1}{\sqrt{N}}\right). \end{aligned} \tag{A.185}$$

Moreover, noting that:

$$\Sigma_{\Lambda,j}^{-1}L_{1,j} = \begin{bmatrix} 0_{(k^c\times k^c)} & 0_{(k^c\times k_j^s)} \\ \left(\frac{1}{\sqrt{T}}F_j^{s\prime}F^c\right) & 0_{(k_j^s\times k_j^s)} \end{bmatrix} + O_p\left(\frac{1}{\sqrt{N}}\right), \tag{A.186}$$

we get:

$$\Sigma_{\Lambda,j}^{-1} L_{1,j} \Sigma_{u,j} = \begin{bmatrix} 0_{(k^c \times k^c)} & 0_{(k^c \times k_j^s)} \\ * & * \end{bmatrix} + O_p\left(\frac{1}{\sqrt{N}}\right). \tag{A.187}$$

Equation (A.187) implies:

$$(L_{3,j})^{(cc)} = O_p\left(\frac{1}{\sqrt{N}}\right). \tag{A.188}$$

Finally, substituting result (A.188) into equation (A.182), equation (A.137) follows.

*Q.E.D.*

# A.8 Proof of Theorem 7

Let us re-write the model for the high frequency observables $x_{m,t}^H$, where $m = 1, ..., M$, and $t = 1, ..., T$ in equation (2.1) as:

$$\begin{aligned} x_{m,t}^H &= \Lambda_{HC} g_{m,t}^C + \Lambda_H g_{m,t}^H + e_{m,t}^H, \\ &= \Lambda_1 g_{m,t} + e_{m,t}^H, \\ &= \hat{\Lambda}_1 \hat{\mathcal{U}}_1^{-1} g_{m,t} - \left(\hat{\Lambda}_1 \hat{\mathcal{U}}_1^{-1} - \Lambda_1\right) g_{m,t} + e_{m,t}^H, \end{aligned} \tag{A.189}$$

where $g_{m,t} = [\, g_{m,t}^{C\,\prime} \vdots g_{m,t}^{H\,\prime} \,]'$, $\Lambda_1 = [\Lambda_{HC} \vdots \Lambda_H] = [\Lambda_1^c \vdots \Lambda_1^s]$, $\hat{\Lambda}_1 = [\hat{\Lambda}_{HC} \vdots \hat{\Lambda}_H] = [\hat{\Lambda}_1^c \vdots \hat{\Lambda}_1^s]$, and $\hat{\mathcal{U}}_1$ has been defined in equation (A.145). Let us also define the estimator $\hat{g}_{m,t} = [\, \hat{g}_{m,t}^{C\,\prime} \vdots \hat{g}_{m,t}^{H\,\prime} \,]'$ as in equation (3.6):

$$\hat{g}_{m,t} = \begin{bmatrix} \hat{g}_{m,t}^C \\ \hat{g}_{m,t}^H \end{bmatrix} = \left(\hat{\Lambda}_1' \hat{\Lambda}_1\right)^{-1} \hat{\Lambda}_1' x_{m,t}^H, \qquad m = 1, ..., M, \qquad t = 1, ..., T. \tag{A.190}$$

Substituting equation (A.189) into equation (A.190), and rearranging terms, we get:

$$\hat{g}_{m,t} = \hat{\mathcal{U}}_1^{-1} g_{m,t} - \left(\frac{\hat{\Lambda}_1' \hat{\Lambda}_1}{N_H}\right)^{-1} \frac{1}{N_H} \hat{\Lambda}_1' \left(\hat{\Lambda}_1 \hat{\mathcal{U}}_1^{-1} - \Lambda_1\right) g_{m,t} + \left(\frac{\hat{\Lambda}_1' \hat{\Lambda}_1}{N_H}\right)^{-1} \frac{1}{N_H} \hat{\Lambda}_1' e_{m,t}^H. \tag{A.191}$$

From equations (A.156) and (A.157) we have:

$$\frac{\hat{\Lambda}_1' \hat{\Lambda}_1}{N_H} = \hat{\mathcal{U}}_1' \Sigma_{\Lambda,1} \hat{\mathcal{U}}_1 + O_p\left(\frac{1}{\sqrt{T}}\right),$$

which implies:

$$\left(\frac{\hat{\Lambda}_1' \hat{\Lambda}_1}{N_H}\right)^{-1} = \hat{\mathcal{U}}_1^{-1} \Sigma_{\Lambda,1}^{-1} \left(\hat{\mathcal{U}}_1'\right)^{-1} + O_p\left(\frac{1}{\sqrt{T}}\right). \tag{A.192}$$

From equations (A.140) - (A.144) we get:

$$\hat{\Lambda}_1 \hat{\mathcal{U}}_1^{-1} - \Lambda_1 \simeq \frac{1}{\sqrt{T}} G_1, \tag{A.193}$$

where

$$G_1 = \left[ G_1^c \vdots G_1^s \right], \tag{A.194}$$

with

$$G_1^c = \frac{1}{\sqrt{T}}\bar{e}^{H\prime}\bar{g}^C + \Lambda_H\left(\frac{1}{\sqrt{T}}\bar{g}^{H\prime}\bar{g}^C\right), \tag{A.195}$$

$$G_1^s = \frac{1}{\sqrt{T}}\bar{e}^{H\prime}\bar{g}^H, \tag{A.196}$$

$\bar{e}^H = \left[\bar{e}_1^H, ..., \bar{e}_T^H\right]'$, $\bar{g}^C = \left[\bar{g}_1^C, ..., \bar{g}_T^C\right]'$ and $\bar{g}^H = \left[\bar{g}_1^H, ..., \bar{g}_T^H\right]'$. Moreover, we have:

$$\hat{\Lambda}_1 \simeq \Lambda_1\hat{\mathcal{U}}_1 + \frac{1}{\sqrt{T}}G_1\hat{\mathcal{U}}_1. \tag{A.197}$$

From equations (A.193) and (A.197) it follows:

$$\begin{aligned}\frac{1}{N_H}\hat{\Lambda}_1'\left(\hat{\Lambda}_1\hat{\mathcal{U}}_1^{-1} - \Lambda_1\right) &\simeq \frac{1}{N_H}\left(\Lambda_1\hat{\mathcal{U}}_1 + \frac{1}{\sqrt{T}}G_1\hat{\mathcal{U}}_1\right)'\frac{1}{\sqrt{T}}G_1 \\ &= \frac{1}{N_H\sqrt{T}}\hat{\mathcal{U}}_1'\Lambda_1'G_1 + \frac{1}{N_H T}\hat{\mathcal{U}}_1'G_1'G_1. \end{aligned}\tag{A.198}$$

Equations (A.192) and (A.198) allow to express the second term in the r.h.s. of equation (A.191) as:

$$\left(\frac{\hat{\Lambda}_1'\hat{\Lambda}_1}{N_H}\right)^{-1}\frac{1}{N_H}\hat{\Lambda}_1'\left(\hat{\Lambda}_1'\hat{\mathcal{U}}_1^{-1} - \Lambda_1\right)g_{m,t} \simeq \hat{\mathcal{U}}_1^{-1}\Sigma_{\Lambda,1}^{-1}\frac{1}{N_H\sqrt{T}}\Lambda_1'G_1 g_{m,t} + \hat{\mathcal{U}}_1^{-1}\Sigma_{\Lambda,1}^{-1}\frac{1}{N_H T}G_1'G_1 g_{m,t} \tag{A.199}$$

From equation (A.152) we have:

$$\frac{1}{N_H\sqrt{T}}\Lambda_1'G_1 = \left[\begin{array}{ccc}\left(\frac{\Lambda_1'\Lambda_H}{N_H}\right)\frac{1}{T}\sum_{t=1}^T \bar{g}_t^H\bar{g}_t^{C\prime} & \vdots & 0_{(k_1\times k^H)}\end{array}\right] + O_p\left(\frac{1}{\sqrt{N_H T}}\right), \tag{A.200}$$

where $k_1 = k^C + k^H$. From equation (A.194) we have:

$$\frac{1}{N_H T}G_1'G_1 = \frac{1}{N_H T}\left[\begin{array}{cc} G_1^{c\prime}G_1^c & G_1^{c\prime}G_1^s \\ G_1^{s\prime}G_1^c & G_1^{s\prime}G_1^s\end{array}\right]. \tag{A.201}$$

Equation (A.195) implies:

$$\begin{aligned}\frac{1}{N_H T}G_1^{c\prime}G_1^c &= \frac{1}{N_H T}\left[\frac{1}{\sqrt{T}}\bar{e}^{H\prime}\bar{g}^C + \Lambda_H\left(\frac{1}{\sqrt{T}}\bar{g}^{H\prime}\bar{g}^C\right)\right]'\left[\frac{1}{\sqrt{T}}\bar{e}^{H\prime}\bar{g}^C + \Lambda_H\left(\frac{1}{\sqrt{T}}\bar{g}^{H\prime}\bar{g}^C\right)\right] \\ &= \frac{1}{N_H T^2}\bar{g}^{C\prime}\bar{e}^H\bar{e}^{H\prime}\bar{g}^C + \frac{1}{N_H T\sqrt{T}}\bar{g}^{C\prime}\bar{e}^H\Lambda_H\left(\frac{1}{\sqrt{T}}\bar{g}^{H\prime}\bar{g}^C\right) \\ &\quad + \frac{1}{N_H T\sqrt{T}}\left(\frac{1}{\sqrt{T}}\bar{g}^{H\prime}\bar{g}^C\right)'\Lambda_H'\bar{e}^{H\prime}\bar{g}^C + \frac{1}{N_H T}\left(\frac{1}{\sqrt{T}}\bar{g}^{H\prime}\bar{g}^C\right)'\Lambda_H'\Lambda_H\left(\frac{1}{\sqrt{T}}\bar{g}^{H\prime}\bar{g}^C\right) \\ &= O_p\left(\frac{1}{T}\right), \end{aligned}\tag{A.202}$$

where the last equality follows from the assumption $T/N_H = o(1)$. Equation (A.202) and the assumption $\sqrt{N_H}/T = o(1)$ imply:

$$\frac{1}{N_H T}G_1^{c\prime}G_1^c = o_p\left(\frac{1}{\sqrt{N_H}}\right). \tag{A.203}$$

Similar arguments applied to the other blocks of the matrix in the r.h.s. of (A.201) yield:

$$\frac{1}{N_H T} G_1' G_1 \quad = \quad o_p\left(\frac{1}{\sqrt{N_H}}\right). \tag{A.204}$$

Substituting equations (A.200) and (A.204) into equation (A.199) we get:

$$\left(\frac{\hat{\Lambda}_1'\hat{\Lambda}_1}{N_H}\right)^{-1} \frac{1}{N_H}\hat{\Lambda}_1'\left(\hat{\Lambda}_1\hat{\mathcal{U}}_1^{-1}-\Lambda_1\right) g_{m,t} \quad \simeq \quad \hat{\mathcal{U}}_1^{-1}\Sigma_{\Lambda,1}^{-1}\left(\frac{\Lambda_1'\Lambda_H}{N_H}\right)\left(\frac{1}{T}\sum_{t=1}^{T}\bar{g}_t^H\bar{g}_t^{C\prime}\right) g_{m,t}^C + o_p\left(\frac{1}{\sqrt{N_H}}\right). \tag{A.205}$$

Let us now focus on the third term in the r.h.s. of equation (A.191). From equation (A.197) we have:

$$\begin{aligned}\frac{1}{N_H}\hat{\Lambda}_1' e_{m,t}^H \quad &\simeq \quad \frac{1}{N_H}\left(\Lambda_1\hat{\mathcal{U}}_1 + \frac{1}{\sqrt{T}}G_1\hat{\mathcal{U}}_1\right)' e_{m,t}^H \\ &= \quad \hat{\mathcal{U}}_1'\frac{1}{N_H}\Lambda_1' e_{m,t}^H + \hat{\mathcal{U}}_1'\frac{1}{N_H\sqrt{T}}G_1' e_{m,t}^H.\end{aligned} \tag{A.206}$$

The second term in the r.h.s. of equation (A.206) can be written as:

$$\frac{1}{N_H\sqrt{T}}G_1' e_{m,t}^H \quad = \quad \frac{1}{N_H\sqrt{T}}\begin{bmatrix} G_1^{c\prime} e_{m,t}^H \\ G_1^{s\prime} e_{m,t}^H\end{bmatrix}. \tag{A.207}$$

Using equation (A.195) we get:

$$\begin{aligned}\frac{1}{N_H\sqrt{T}}G_1' e_{m,t}^H \quad &= \quad \frac{1}{N_H T}\bar{g}^{C\prime}\bar{e}^H e_{m,t}^H + \frac{1}{N_H\sqrt{T}}\left(\frac{1}{\sqrt{T}}\bar{g}^{C\prime}\bar{g}^H\right)\Lambda_H' e_{m,t}^H \\ &= \quad O_p\left(\frac{1}{\sqrt{N_H T}}\right).\end{aligned} \tag{A.208}$$

Equation (A.196) implies:

$$\frac{1}{N_H\sqrt{T}}G_1^{s\prime} e_{m,t}^H \quad = \quad \frac{1}{N_H T}\bar{g}^{H\prime}\bar{e}^H e_{m,t}^H \; = \; O_p\left(\frac{1}{\sqrt{N_H T}}\right). \tag{A.209}$$

Substituting results (A.208) and (A.209) into equations (A.207) and (A.206) we get:

$$\frac{1}{N_H}\hat{\Lambda}_1' e_{m,t}^H \quad = \quad \hat{\mathcal{U}}_1'\frac{1}{N_H}\Lambda_1' e_{m,t}^H + O_p\left(\frac{1}{\sqrt{N_H T}}\right). \tag{A.210}$$

Substituting results (A.192), (A.205), and (A.210) into equation (A.191), and rearranging terms we get:

$$\hat{\mathcal{U}}_1\hat{g}_{m,t} - g_{m,t} \quad = \quad -\Sigma_{\Lambda,1}^{-1}\left(\frac{\Lambda_1'\Lambda_H}{N_H}\right)\left(\frac{1}{T}\bar{g}^{H\prime}\bar{g}^C\right) g_{m,t}^C + \Sigma_{\Lambda,1}^{-1}\frac{1}{N_H}\Lambda_1' e_{m,t}^H + o_p\left(\frac{1}{\sqrt{N_H}}\right). \tag{A.211}$$

Let us denote the last $k^H$ columns of matrix $\Sigma_{\Lambda,1}$ as $\Sigma_{\Lambda,1}^{(\cdot\ s)}$. The term $\frac{\Lambda_1'\Lambda_H}{N_H}$ in equation (A.211) can be written as:

$$\begin{aligned}\frac{\Lambda_1'\Lambda_H}{N_H} \quad &= \quad \Sigma_{\Lambda,1}^{(\cdot\ s)} + \frac{1}{N_H}\sum_{i=1}^{N_H}\lambda_{1,i}\lambda_{H,i}' - \Sigma_{\Lambda,1}^{(\cdot\ s)} \\ &= \quad \Sigma_{\Lambda,1}^{(\cdot\ s)} + O_p\left(\frac{1}{\sqrt{N_H}}\right),\end{aligned} \tag{A.212}$$

where the last equality follows from Assumption A.2. Equation (A.212) implies:

$$\Sigma_{\Lambda,1}^{-1}\left(\frac{\Lambda_1'\Lambda_H}{N_H}\right) = \begin{bmatrix} 0_{(k^C\times k^H)} \\ I_{k^H} \end{bmatrix} + O_p\left(\frac{1}{\sqrt{N_H}}\right). \tag{A.213}$$

Substituting equation (A.213) into equation (A.211) we have:

$$\hat{\mathcal{U}}_1\hat{g}_{m,t} - g_{m,t} = -\begin{bmatrix} 0_{(k^C\times k^H)} \\ I_{k^H} \end{bmatrix}\left(\frac{1}{T}\bar{g}^{H\prime}\bar{g}^C\right) g_{m,t}^C + \Sigma_{\Lambda,1}^{-1}\frac{1}{N_H}\Lambda_1' e_{m,t}^H + o_p\left(\frac{1}{\sqrt{N_H}}\right). \tag{A.214}$$

Recalling the expression of $\hat{\mathcal{U}}_1$ from equation (A.145):

$$\hat{\mathcal{U}}_1 = \begin{bmatrix} \hat{\mathcal{H}}_c & 0 \\ 0 & \hat{\mathcal{H}}_{s,1} \end{bmatrix}, \tag{A.215}$$

from equation (A.214) we get the asymptotic expansions:

$$\hat{\mathcal{H}}_c\hat{g}_{m,t}^C - g_{m,t}^C \simeq \left[\Sigma_{\Lambda,1}^{-1}\frac{1}{N_H}\Lambda_1' e_{m,t}^H\right]^{(C)}, \tag{A.216}$$

$$\hat{\mathcal{H}}_{1,s}\hat{g}_{m,t}^H - g_{m,t}^H \simeq -\left(\frac{1}{T}\bar{g}^{H\prime}\bar{g}^C\right) g_{m,t}^C + \left[\Sigma_{\Lambda,1}^{-1}\frac{1}{N_H}\Lambda_1' e_{m,t}^H\right]^{(H)}, \tag{A.217}$$

where $\left[\Sigma_{\Lambda,1}^{-1}\frac{1}{N_H}\Lambda_1' e_{m,t}^H\right]^{(C)}$ and $\left[\Sigma_{\Lambda,1}^{-1}\frac{1}{N_H}\Lambda_1' e_{m,t}^H\right]^{(H)}$ denote the upper $k^C$ rows, resp. the lower $k^H$ rows, of vector $\Sigma_{\Lambda,1}^{-1}\frac{1}{N_H}\Lambda_1' e_{m,t}^H$. Since $\bar{g}^{C\prime}\bar{g}^C/T = I_{k^C} + o_p(1)$, we can rewrite equation (A.217) as:

$$\hat{\mathcal{H}}_{1,s}\hat{g}_{m,t}^H - (g_{m,t}^H - (\bar{g}^{H\prime}\bar{g}^C)(\bar{g}^{C\prime}\bar{g}^C)^{-1}g_{m,t}^C) \simeq [\Sigma_{\Lambda,1}^{-1}\frac{1}{N_H}\Lambda_1' e_{m,t}^H]^{(H)}. \tag{A.218}$$

From Assumption A.8 we have:

$$\frac{1}{\sqrt{N_H}}\Lambda_1' e_{m,t}^H \xrightarrow{d} N(0,\Omega_{\Lambda,m}^*), \tag{A.219}$$

where

$$\Omega_{\Lambda,m}^* = \lim_{N_H\to\infty}\frac{1}{N_H}\sum_{i=1}^{N_H}\sum_{\ell=1}^{N_H}\lambda_{1,i}\lambda_{1,\ell}' Cov(e_{m,t}^{i,H}, e_{m,t}^{\ell,H}). \tag{A.220}$$

Equations (A.216) and (A.219) imply:

$$\sqrt{N_H}\left(\hat{\mathcal{H}}_c\hat{g}_{m,t}^C - g_{m,t}^C\right) \xrightarrow{d} N\left(0,\left[\Sigma_{\Lambda,1}^{-1}\Omega_{\Lambda,m}^*\Sigma_{\Lambda,1}^{-1}\right]^{(CC)}\right).$$

Similarly, equation (A.218) and (A.219) imply:

$$\sqrt{N_H}\left[\hat{\mathcal{H}}_{1,s}\hat{g}_{m,t}^H - (g_{m,t}^H - (\bar{g}^{H\prime}\bar{g}^C)(\bar{g}^{C\prime}\bar{g}^C)^{-1})g_{m,t}^C\right] \xrightarrow{d} N\left(0,\left[\Sigma_{\Lambda,1}^{-1}\Omega_{\Lambda,m}^*\Sigma_{\Lambda,1}^{-1}\right]^{(HH)}\right). \tag{A.221}$$

*Q.E.D.*