Supplement to:

"Inference in Group Factor Models with an Application to Mixed Frequency Data"

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

First version: October, 2014

This version: April 16, 2019

[*] University of Cyprus and CEPR (elena.andreou@ucy.ac.cy).
[†] Università della Svizzera Italiana (USI, Lugano) and Swiss Finance Institute (patrick.gagliardini@usi.ch).
[‡] University of North Carolina - Chapel Hill and CEPR (eghysels@unc.edu).
[§] (mircorubin@gmail.com).

Contents

Introduction

In this Online Appendix we provide material supplementary to Andreou, Gagliardini, Ghysels, and Rubin (2019). In Section C we provide proofs of the Propositions 1, 2, 3 and technical Lemmas B.1-B.9 appearing in the paper.

Section D provides additional theoretical results on identification, including the separation of common and group-specific factors, an alternative identification strategy, different from the canonical correlation analysis proposed in Section 2 of Andreou, Gagliardini, Ghysels, and Rubin (2019), for the common and group-specific factor spaces in a group-factor model, and a discussion of identification of the mixed frequency factor model in the cases of stock-sampling, and of general linear aggregation schemes for the LF observables. This section also provides the asymptotic distribution of our factors and loadings estimators in group factor models, a digression on some regularity conditions, and contains a discussion of properties of an iterative PCA estimator for group factor models. Moreover, Section D also contains an exhaustive description of the dataset used in the empirical application of Section 7 and presents additional empirical results.

Online Appendix E describes the Monte Carlo (MC) simulation study used to assess the finite sample size and power properties of tests of hypotheses on the number of common factors k^c based on the test statistics introduced in Theorems 1 and 2. MC simulations are also used to compare the performance of the sequential testing procedure for the selection of the number of common factors introduced in Proposition 2 with alternatives adopted from earlier literature. Finally, we also compare the quantiles of the cross-sectional distribution of R^2 and adjusted R^2 of regression of simulated observables on factors, when the number of common factors is either correctly specified, or overestimated, for a DGP in which specific factors are highly correlated.

C Proofs of Propositions 1, 2, 3 and Lemmas B.1-B.9

C.1 Proof of Proposition 1

From equation (2.2) we have

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

Matrix R is block diagonal, and the upper-left block I_{k^c} has eigenvalue 1 with multiplicity k^c. The associated eigenspace is $\{(\xi', 0')', \xi \in \mathbb{R}^{k^c}\}$. From the positive-definite character of matrix Σ_F in Assumption A.2, the lower-right block $\Phi\Phi'$ is a positive semi-definite matrix and its largest eigenvalue is $\tilde{\rho}^2 < 1$, where $\tilde{\rho}^2 = \sup\left\{\xi_1'\Phi\Phi'\xi_1 : \xi_1 \in \mathbb{R}^{k_1^s}, \|\xi_1\| = 1\right\}$ 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:

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

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 $\xi^* \in \mathbb{R}^{k^c}$. We have

$\mathcal{E}_c^* = \mathcal{E}_c$. 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:

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

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 whose columns are eigenvectors of $\Phi\Phi'$, 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: $W_1^{s\prime} h_{1,t} = Q_1' f_{1,t}^s$.

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 whose columns are eigenvectors of $\Phi'\Phi$, 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 $W_2^{s\prime} h_{2,t} = Q_2' f_{2,t}^s$. ■

C.2 Proof of Proposition 2

Let us define the events $\Omega_{r,\alpha_{N,T}} \equiv \{\tilde{\xi}(r) < z_{\alpha_{N,T}}\}$, for $r = 1, ..., \underline{k}$, and their complementary events $\Omega^c_{r,\alpha_{N,T}} = \{\tilde{\xi}(r) \geq z_{\alpha_{N,T}}\}$. For any integer $k^* \leq \underline{k}$ we can write the event $\{\hat{k}^c = k^*\}$ as:

$$\{\hat{k}^c = k^*\} = \begin{cases} \Omega^c_{k^*,\alpha_{N,T}}, & \text{if } k^* = \underline{k}, \\ \left(\bigcap_{r=k^*+1}^{\underline{k}} \Omega_{r,\alpha_{N,T}}\right) \bigcap \Omega^c_{k^*,\alpha_{N,T}}, & \text{if } 0 < k^* < \underline{k}, \\ \bigcap_{r=k^*+1}^{\underline{k}} \Omega_{r,\alpha_{N,T}}, & \text{if } k^* = 0. \end{cases} \tag{C.3}$$

We prove Proposition 2 by distinguishing three cases according to the true number of common factors: $k_0^c = \underline{k}$, $0 < k_0^c < \underline{k}$, $k_0^c = 0$. Moreover, we use the convergence results:

$$
\begin{aligned}
P(\Omega_{r,\alpha_{N,T}}) &\to 1, \quad r > k_0^c, &\text{(C.4)}\\
P(\Omega_{r,\alpha_{N,T}}) &\to 0, \quad r = k_0^c, &\text{(C.5)}
\end{aligned}
$$

which are proved at the end of the section.

i) Case $k_0^c = \underline{k}$. We have $P(\hat{k}^c = k_0^c) = P(\Omega^c_{\underline{k},\alpha_{N,T}}) = 1 - P(\Omega_{\underline{k},\alpha_{N,T}}) \to 1$, from equation (C.5).

ii) Case $0 < k_0^c < \underline{k}$. From (C.3), we have $\{\hat{k}^c = k_0^c\} = \left(\bigcap_{r=k_0^c+1}^{\underline{k}} \Omega_{r,\alpha_{N,T}}\right) \bigcap \Omega^c_{k_0^c,\alpha_{N,T}}$. The events $\Omega_{r,\alpha_{N,T}}$, for $r = k_0^c+1, ..., \underline{k}$, have all probability tending to 1 from equation (C.4), and so do events $\bigcap_{r=k_0^c+1}^{\underline{k}} \Omega_{r,\alpha_{N,T}}$ and $\left(\bigcap_{r=k_0^c+1}^{\underline{k}} \Omega_{r,\alpha_{N,T}}\right) \bigcup \Omega^c_{k_0^c,\alpha_{N,T}}$. Moreover, $P(\Omega^c_{k_0^c,\alpha_{N,T}}) = 1 - P(\Omega_{k_0^c,\alpha_{N,T}}) \to 1$ from equation (C.5). therefore, we get:

$$
\begin{aligned}
P(\hat{k}^c = k_0^c) &= P\left[\left(\bigcap_{r=k_0^c+1}^{\underline{k}} \Omega_{r,\alpha_{N,T}}\right) \bigcap \Omega^c_{k_0^c,\alpha_{N,T}}\right]\\
&= P\left(\bigcap_{r=k_0^c+1}^{\underline{k}} \Omega_{r,\alpha_{N,T}}\right) + P\left(\Omega^c_{k_0^c,\alpha_{N,T}}\right) - P\left[\left(\bigcap_{r=k_0^c+1}^{\underline{k}} \Omega_{r,\alpha_{N,T}}\right) \bigcup \Omega^c_{k_0^c,\alpha_{N,T}}\right] \to 1\,.
\end{aligned}
$$

iii) Case $k_0^c = 0$. We have $P(\hat{k}^c = k_0^c) = P\left(\bigcap_{r=1}^{\underline{k}} \Omega_{r,\alpha_{N,T}}\right) \to 1$, because the events $\Omega_{r,\alpha_{N,T}}$, for $r = 1, ..., \underline{k}$, have all probability tending to 1, from equation (C.4).

C.2.1 Proofs of (C.4) and (C.5)

Let $r > k_0^c$. Then, from the arguments in the proof of Theorem 2 *(ii)* (see Section B.2.2), we have $\dfrac{\tilde{\xi}(r)}{N\sqrt{T}} \le -c_1$, w.p.a. 1, for a constant $c_1 > 0$. By Condition *(ii)* of Proposition 2, we have $\dfrac{z_{\alpha_{N,T}}}{N\sqrt{T}} \to 0$. Then, $P(\Omega_{r,\alpha_{N,T}}) = P\left(\frac{\tilde{\xi}(r)}{N\sqrt{T}} < \frac{z_{\alpha_{N,T}}}{N\sqrt{T}}\right) \to 1$ follows, which yields (C.4).

Now, let $r = k_0^c$. Then, from Theorem 2 *(ii)* we have $\tilde{\xi}(r) \xrightarrow{d} N(0,1)$. Moreover, since $\alpha_{N,T} \to 0$ by Condition *(i)* of Proposition 2, we have $z_{\alpha_{N,T}} \le z_{\alpha^*}$ for large N, T, for any given $\alpha^* \in (0,1)$. therefore:

$$
P(\Omega_{r,\alpha_{N,T}}) = P[\tilde{\xi}(r) < z_{\alpha_{N,T}}] \le P[\tilde{\xi}(r) < z_{\alpha^*}] \to \alpha^*.
$$

Therefore, we have $\liminf\limits_{N,T\to\infty} P(\Omega_{r,\alpha_{N,T}}) \le \alpha^*$, for any $\alpha^* \in (0,1)$. It follows $P(\Omega_{r,\alpha_{N,T}}) \to 0$, which yields (C.5). ■

C.3 Proof of Proposition 3

We omit the subpanel index j since it is immaterial for the proof's arguments. We write the factor models in each subpanel as:

$$
y_{i,t} = \lambda_i' h_t + \varepsilon_{i,t}, \quad i = 1, ..., N, \quad t = 1, ..., T, \tag{C.6}
$$

where h_t is the $(k, 1)$ vector of unobservable factors. In matrix notation, the model becomes:

$$Y = H\Lambda' + \varepsilon, \tag{C.7}$$

where Y is the (T, N) matrix of observations and H is the (T, k) matrix of factor values. We introduce a set of high-level assumptions (Assumptions C.1-C.3 below) and show in Section C.3.7 that they are implied by Assumptions A.2-A.4, A.5 *b)-c)*, A.6 *a)* and A.7.

Assumption C.1. *The factors are such that $H'H/T = I_k + o_p(1)$ as $T \to \infty$. The loadings are such that $\Lambda'\Lambda/N = \Sigma_\lambda + o(1)$ as $N \to \infty$, where matrix Σ_λ is positive definite.*

The matrix of factor estimates $\hat{H} = [\hat{h}_1, ..., \hat{h}_T]'$ corresponds to the estimator obtained by Principal Component Analysis (PCA), and satisfies the eigenvector-eigenvalue equation:

$$\frac{1}{NT} YY'\hat{H} = \hat{H}\hat{V}, \tag{C.8}$$

where $\hat{V}$ is the (k, k) diagonal matrix of the k largest eigenvalues of matrix $YY'/(NT)$, and the columns of matrix $\hat{H}$ are the associated normalized eigenvectors such that $\hat{H}'\hat{H}/T = I_k$.
We start by establishing an asymptotic expansion of the factor estimate with explicit characterization of the remainder term. It is obtained by manipulating equation (C.8) using the next assumption.

Assumption C.2. *We have (i) $\frac{1}{\sqrt{NT}} H'\varepsilon\Lambda = \frac{1}{\sqrt{T}} \sum_{t=1}^T h_t \xi_t' = O_p(1)$ and $E[\|\xi_t\|^2] = O(1)$, where $\xi_t := \frac{1}{\sqrt{N}} \sum_{i=1}^N \lambda_i \varepsilon_{i,t}$, (ii) $\|\frac{1}{NT}\varepsilon\varepsilon' H\| = O_p\left(\frac{1}{\sqrt{m}}\right)$, (iii) $\|\frac{1}{NT}\varepsilon\varepsilon'\| = O_p\left(\frac{1}{\sqrt{m}}\right)$, where $m := \min\{N, T\}$.*

PROPOSITION C.1. *Under Assumptions C.1-C.2 we have:*

$$(\mathscr{H}')^{-1}\hat{h}_t - h_t = \frac{1}{\sqrt{N}} u_t + \frac{1}{T} b_t + \frac{1}{\sqrt{NT}} d_t + \vartheta_t, \qquad t = 1, ..., T, \tag{C.9}$$

where matrix $\mathscr{H} = (\Lambda'\Lambda/N)^{-1}(\hat{H}'H/T)\hat{V}^{-1}$ is invertible w.p.a. 1, and:

$$\begin{aligned} u_t &= (\Lambda'\Lambda/N)^{-1}\xi_t, \qquad\qquad b_t = S\eta_t^2 h_t, \\ d_t &= S\Pi_1 h_t, \qquad\qquad \vartheta_t = \frac{1}{\sqrt{NT}} S\alpha_t + \frac{1}{N} D_2 h_t + r_t + \hat{\mathcal{R}}_t, \end{aligned}$$

with $\eta_t^2 = \underset{N\to\infty}{plim} \frac{1}{N} \sum_{i=1}^N E[\varepsilon_{i,t}^2 | \mathcal{F}_t]$ and $\mathcal{F}_t$ is the sigma-field generated by the h_s for $s \leq t$,

$$\begin{aligned} r_t = {} & \hat{B}'(\frac{1}{T} b_t + \frac{1}{\sqrt{NT}} S\alpha_t) + (I_k + \hat{B})' \frac{1}{T\sqrt{N}} S\kappa_t h_t + (\frac{1}{\sqrt{NT}} D_1 + \frac{1}{NT} D_3) h_t \\ & + (I_k + \hat{B}') S (\Lambda'\Lambda/N)^{-1} \left(\frac{1}{\sqrt{NT}} \eta_t^2 \xi_t + \frac{1}{NT} \kappa_t \xi_t + \frac{1}{N\sqrt{N}} \varphi_t + \frac{1}{N\sqrt{T}} \gamma_t \right) \\ & + [(I_k + \hat{B}')S]^2 \Pi_1 \frac{1}{\sqrt{NT}} \left(\frac{1}{T} \eta_t^2 h_t + \frac{1}{T\sqrt{N}} \kappa_t h_t + \frac{1}{\sqrt{NT}} \alpha_t \right) \\ & + [(I_k + \hat{B}')S]^2 \left(\frac{1}{T^2} \eta_t^4 h_t + \frac{1}{\sqrt{N}T^2} \kappa_t \eta_t^2 h_t + \frac{1}{\sqrt{NT}T} \bar{\alpha}_t \right. \\ & \qquad + \frac{1}{T^2\sqrt{N}} \kappa_t \eta_t^2 h_t + \frac{1}{NT^2} \kappa_t^2 h_t + \frac{1}{NT\sqrt{N}} \bar{\varphi}_t + \frac{1}{NT\sqrt{T}} \bar{\gamma}_t \\ & \qquad \left. + \frac{1}{T\sqrt{NT}} \eta_t^2 \alpha_t + \frac{1}{NT\sqrt{T}} \kappa_t \alpha_t + \frac{1}{N\sqrt{NT}} \delta_t + \frac{1}{NT} \chi_t \right), \end{aligned} \tag{C.10}$$

$\hat{\mathcal{R}}_t'$ is the t-th row of matrix $\hat{\mathcal{R}}$ defined by:

$$\hat{\mathcal{R}} = \left[\frac{1}{NT}\left(\varepsilon\varepsilon' + H\Lambda'\varepsilon'\right)\right]^2 \left(\hat{H}\mathscr{H}^{-1} - H\right)\left[\mathscr{H}(H'\hat{H}/T)^{-1}(\Lambda'\Lambda/N)^{-1}\right]^2, \tag{C.11}$$

and:

$$\begin{aligned}
S &= (\Lambda'\Lambda/N)^{-1}(H'H/T)^{-1}, \qquad D_1 = \hat{B}'S\Pi_1 + [(I_k + \hat{B}')S]^2\Pi_3, \\
D_2 &= (I_k + \hat{B}')S(\Lambda'\Lambda/N)^{-1}\Pi_2, \qquad D_3 = [(I_k + \hat{B}')S]^2\Pi_1^2, \\
\Pi_1 &= \frac{1}{\sqrt{NT}}H'\varepsilon\Lambda = \frac{1}{\sqrt{T}}\sum_{t=1}^T h_t\xi_t', \\
\Pi_2 &= \frac{1}{NT}\Lambda'\varepsilon'\varepsilon\Lambda = \frac{1}{T}\sum_{t=1}^T \xi_t\xi_t', \\
\Pi_3 &= \frac{1}{NT\sqrt{NT}}H'\varepsilon\varepsilon'\varepsilon\Lambda = \frac{1}{\sqrt{N}}(\frac{1}{T}\sum_{t=1}^T \alpha_t\xi_t') + \frac{1}{T}(\frac{1}{\sqrt{T}}\sum_{t=1}^T \eta_t^2 h_t\xi_t') + \frac{1}{\sqrt{NT}}(\frac{1}{T}\sum_{t=1}^T h_t\xi_t'\kappa_t), \\
\hat{B} &= (\Lambda'\Lambda/N)(H'H/T)\left[\left(I_k + \hat{A}\right)^{-1} - I_k\right](H'H/T)^{-1}(\Lambda'\Lambda/N)^{-1}, \\
\hat{A} &= (H'H/T)^{-1}H'(\hat{H}\mathscr{H}^{-1} - H)/T,
\end{aligned} \tag{C.12}$$

and we define:

$$\begin{aligned}
\kappa_t &= \frac{1}{\sqrt{N}}\sum_{i=1}^N(\varepsilon_{i,t}^2 - \eta_t^2), \qquad \alpha_t = \frac{1}{\sqrt{NT}}\sum_{i=1}^N\sum_{s=1,s\neq t}^T \varepsilon_{i,t}\varepsilon_{i,s}h_s, \\
\varphi_t &= \frac{1}{T}\sum_{i=1}^N\sum_{s=1,s\neq t}^T \varepsilon_{i,t}E[\varepsilon_{i,s}\xi_s], \qquad \gamma_t = \frac{1}{\sqrt{NT}}\sum_{i=1}^N\sum_{s=1,s\neq t}^T \varepsilon_{i,t}(\varepsilon_{i,s}\xi_s - E[\varepsilon_{i,s}\xi_s]), \\
\bar{\alpha}_t &= \frac{1}{\sqrt{NT}}\sum_{i=1}^N\sum_{s=1,s\neq t}^T \varepsilon_{i,t}\varepsilon_{i,s}\eta_s^2 h_s, \qquad \bar{\varphi}_t = \frac{1}{T}\sum_{i=1}^N\sum_{s=1,s\neq t}^T \varepsilon_{i,t}E[\varepsilon_{i,s}\kappa_s h_s], \\
\bar{\gamma}_t &= \frac{1}{\sqrt{NT}}\sum_{i=1}^N\sum_{s=1,s\neq t}^T \varepsilon_{i,t}(\varepsilon_{i,s}\kappa_s h_s - E[\varepsilon_{i,s}\kappa_s h_s]), \\
\delta_t &= \frac{1}{T}\sum_{i=1}^N\sum_{s=1,s\neq t}^T \varepsilon_{i,t}E[\varepsilon_{i,s}\alpha_s], \\
\chi_t &= \frac{1}{\sqrt{NT}}\sum_{i=1}^N\sum_{s=1,s\neq t}^T \varepsilon_{i,t}(\varepsilon_{i,s}\alpha_s - E[\varepsilon_{i,s}\alpha_s]).
\end{aligned} \tag{C.13}$$

Moreover, if the eigenvalues of matrix Σ_λ in Assumption C.1 are distinct, then, for a suitable ordering and choice of the signs of the factor estimates, we have $\mathscr{H} \xrightarrow{p} \mathscr{H}^*$, where the columns of the orthogonal matrix $\mathscr{H}^*$ are the normalized eigenvectors of Σ_λ.

In equation (C.9), the difference $(\mathscr{H}')^{-1}\hat{h}_t - h_t$ is written as a sum of a zero-mean term at stochastic order $1/\sqrt{N}$, terms at orders $1/T$, $1/\sqrt{NT}$ and $1/N$, plus remainder terms r_t and $\hat{\mathcal{R}}_t$. The remainder terms are either scaled by factors that converge to zero faster than $\max\{\frac{1}{T}, \frac{1}{\sqrt{NT}}, \frac{1}{N}\} = O(\frac{1}{m})$, where $m = \min\{N, T\}$, or are

of higher order in the sense that involve $\hat{H}\hat{\mathscr{H}}^{-1} - H$ multiplied by matrices whose elements converge to zero in probability. The result on the converge of matrix $\hat{\mathscr{H}}$ corresponds to Proposition 1 in Bai (2003).
Equation (C.9) corresponds to the expansion in (B.1). We now control for the magnitude of the remainder terms r_t and $\hat{\mathcal{R}}_t$ in ϑ_t to show the bounds in Proposition 3. The next Proposition C.2 provides an upper bound for $T^{-1/2}\|\hat{H}\hat{\mathscr{H}}^{-1} - H\| = \left(\frac{1}{T}\sum_{t=1}^T \|(\hat{\mathscr{H}}^{-1})'\hat{h}_t - h_t\|^2\right)^{1/2}$, namely the root MSE of the factor estimates. It is similar to Lemma A.1 in Bai (2003) but it yields a sharper upper bound. This result is used to derive a bound on the remainder term $\hat{\mathcal{R}}$, which is also provided in Proposition C.2.

PROPOSITION C.2. *Under Assumptions C.1-C.2, we have*

$$T^{-1/2}\|\hat{H}\hat{\mathscr{H}}^{-1} - H\| = O_p(\frac{1}{\sqrt{N}} + \frac{1}{T}).$$

Moreover:

$$\begin{aligned} T^{-1/2}\|\hat{\mathcal{R}}\| &= O_p\left[(\frac{1}{m} + \frac{1}{N})(\frac{1}{\sqrt{N}} + \frac{1}{T})\right] && \text{(C.14)} \\ &= O_p\left(\frac{1}{N} + \frac{1}{T^2}\right). && \text{(C.15)} \end{aligned}$$

From Proposition C.2 and Assumption C.1, we have term $\hat{A}$ defined in (C.12) is such that $\|\hat{A}\| = O_p(\frac{1}{\sqrt{N}} + \frac{1}{T})$. By the series representation of the inverse matrix function in a neighborhood of the identity, we deduce that $\|(I_k + \hat{A})^{-1} - I_k\| = O_p(\frac{1}{\sqrt{N}} + \frac{1}{T})$. Thus, from Proposition C.2 and Assumption C.1 we get that term $\hat{B}$ appearing in the remainder term r_t in the expansion of Proposition C.1 is such that:

$$\hat{B} = O_p(\frac{1}{\sqrt{N}} + \frac{1}{T}). \tag{C.16}$$

To control for the remainder term r_t we use the next assumption.

Assumption C.3. *We have: (i) $E[\varepsilon_{i,t}^2|\mathcal{F}_t] \leq M$ for all $i \geq 1$ and $t \geq 1$, and a constant $M > 0$, (ii) $\frac{1}{\sqrt{T}}\sum_{t=1}^T h_t\alpha_t' = O_p(1)$, (iii) $\frac{1}{\sqrt{T}}\sum_{t=1}^T \eta_t^2 h_t\alpha_t' = O_p(1)$, (iv) $\frac{1}{T}\sum_{t=1}^T \xi_t\alpha_t' = o_p(1)$, and (v) $E[\|a_t\|^2] = O(1)$, where a_t is any of the following processes: $\kappa_t h_t$, α_t, $\kappa_t\xi_t$, φ_t, γ_t, $\bar{\alpha}_t$, $\kappa_t^2 h_t$, $\bar{\varphi}_t$, $\bar{\gamma}_t$, $\kappa_t\alpha_t$, δ_t, χ_t.*

PROPOSITION C.3. *Under Assumptions A.1 and C.1-C.3, we have:*

$$\frac{1}{T}\sum_{t=1}^T \|r_t\|^2 = O_p(\frac{1}{N} + \frac{1}{T^2}). \tag{C.17}$$

Moreover, ϑ_t satisfies $\frac{1}{T}\sum_{t=1}^T \vartheta_t h_t' = O_p(\frac{1}{N} + \frac{1}{T^2})$ and $\frac{1}{T}\sum_{t=1}^T(\frac{1}{\sqrt{N}}u_t + \frac{1}{T}b_t + \frac{1}{\sqrt{NT}}d_t + \vartheta_t)\vartheta_t' = o_p(\frac{1}{N\sqrt{T}})$.

Propositions C.1 and C.3 yield Proposition 3 (with $\mathcal{H} = \mathscr{H}'$ in each group).
In the rest of this appendix we provide the proofs of Propositions C.1-C.3 and show that Assumptions C.1-C.3 are implied by the Assumptions in Appendix A.

C.3.1 Proof of Proposition C.1

From equation (C.7) we have $YY' = H\Lambda'\Lambda H' + H\Lambda'\varepsilon' + \varepsilon\Lambda H' + \varepsilon\varepsilon'$. By plugging this equation into (C.8), and rearranging the terms, we get:

$$\hat{H}\hat{V} - H\left(\Lambda'\Lambda/N\right)\left(H'\hat{H}/T\right) = \frac{1}{NT}(\varepsilon\varepsilon'\hat{H} + H\Lambda'\varepsilon'\hat{H} + \varepsilon\Lambda H'\hat{H}). \tag{C.18}$$

The large sample behaviours of the matrices $H'\hat{H}/T$ and $\hat{V}$ are given in the next Lemmas C.4 and C.5, respectively. These lemmas are similar to the results derived in e.g. Bai and Ng (2002), Bai (2003) - see in particular their Proposition 1 - and Bai (2009) - see in particular his Proposition 1.

LEMMA C.4. *Under Assumptions C.1-C.2, the matrix $H'\hat{H}/T$ is invertible w.p.a. 1, and the inverse is such that $\|(H'\hat{H}/T)^{-1}\| = O_p(1)$.*

LEMMA C.5. *Under Assumptions C.1-C.2, we have $\hat{V} \xrightarrow{p} V$, where V is the (k,k) diagonal matrix with diagonal elements corresponding to the eigenvalues of matrix Σ_λ.*

From Lemma C.5 and Assumption C.1, the matrix $\hat{V}$ is invertible w.p.a. 1. Define the matrix:

$$\hat{\mathscr{H}} = (\Lambda'\Lambda/N)(H'\hat{H}/T)\hat{V}^{-1}. \tag{C.19}$$

From Assumption C.1 and Lemmas C.4 and C.5, matrix $\hat{\mathscr{H}}$ is invertible w.p.a. 1. By post-multiplication of equation (C.18) times the matrix $(H'\hat{H}/T)^{-1}(\Lambda'\Lambda/N)^{-1}$, and using the definition of matrix $\hat{\mathscr{H}}$ given in (C.19), we get:

$$\hat{H}\hat{\mathscr{H}}^{-1} - H = \frac{1}{NT}\left(\varepsilon\varepsilon' + H\Lambda'\varepsilon'\right)\hat{H}(H'\hat{H}/T)^{-1}(\Lambda'\Lambda/N)^{-1} + \frac{1}{N}\varepsilon\Lambda(\Lambda'\Lambda/N)^{-1}. \tag{C.20}$$

This equation can be rewritten as:

$$\begin{aligned}\hat{H}\hat{\mathscr{H}}^{-1} - H &= \frac{1}{NT}\left(\varepsilon\varepsilon' H + H\Lambda'\varepsilon' H\right)\hat{\mathscr{H}}(H'\hat{H}/T)^{-1}(\Lambda'\Lambda/N)^{-1} + \frac{1}{N}\varepsilon\Lambda(\Lambda'\Lambda/N)^{-1} \\ &\quad + \frac{1}{NT}\left(\varepsilon\varepsilon' + H\Lambda'\varepsilon'\right)\left(\hat{H}\hat{\mathscr{H}}^{-1} - H\right)\hat{\mathscr{H}}(H'\hat{H}/T)^{-1}(\Lambda'\Lambda/N)^{-1}. \end{aligned}\tag{C.21}$$

By using $\hat{H} = [H + (\hat{H}\hat{\mathscr{H}}^{-1} - H)]\hat{\mathscr{H}}$, we have:

$$\begin{aligned}(H'\hat{H}/T)^{-1} &= \left[(H'H/T)\left(I_k + (H'H/T)^{-1}H'(\hat{H}\hat{\mathscr{H}}^{-1} - H)/T\right)\hat{\mathscr{H}}\right]^{-1} \\ &= \hat{\mathscr{H}}^{-1}\left(I_k + \hat{A}\right)^{-1}(H'H/T)^{-1}, \end{aligned}\tag{C.22}$$

where $\hat{A} = (H'H/T)^{-1}H'(\hat{H}\hat{\mathscr{H}}^{-1} - H)/T$. By substituting (C.22) in the first term in the RHS of (C.21), and rearranging terms, we get:

$$\begin{aligned}\hat{H}\hat{\mathscr{H}}^{-1} - H &= \frac{1}{N}\varepsilon\Lambda(\Lambda'\Lambda/N)^{-1} + \frac{1}{NT}\left(\varepsilon\varepsilon' H + H\Lambda'\varepsilon' H\right)(H'H/T)^{-1}(\Lambda'\Lambda/N)^{-1}(I_k + \hat{B}) \\ &\quad + \frac{1}{NT}\left(\varepsilon\varepsilon' + H\Lambda'\varepsilon'\right)\left(\hat{H}\hat{\mathscr{H}}^{-1} - H\right)\hat{\mathscr{H}}(H'\hat{H}/T)^{-1}(\Lambda'\Lambda/N)^{-1}, \end{aligned}\tag{C.23}$$

where

$$\hat{B} = (\Lambda'\Lambda/N)(H'H/T)\left[\left(I_k + \hat{A}\right)^{-1} - I_k\right](H'H/T)^{-1}(\Lambda'\Lambda/N)^{-1}. \tag{C.24}$$

Equation (C.23) is a recursive equation for $\hat{H}\hat{\mathscr{H}}^{-1} - H$, since this quantity appears in the third term in the r.h.s. By iterating this equation, we get:

$$
\begin{aligned}
& \hat{H}\hat{\mathscr{H}}^{-1} - H \\
= \quad & \frac{1}{N}\varepsilon\Lambda(\Lambda'\Lambda/N)^{-1} + \frac{1}{NT}\left(\varepsilon\varepsilon' H + H\Lambda'\varepsilon' H\right)(H'H/T)^{-1}(\Lambda'\Lambda/N)^{-1}(I_k + \hat{B}) \\
& + \frac{1}{NT}\left(\varepsilon\varepsilon' + H\Lambda'\varepsilon'\right)\left(\frac{1}{N}\varepsilon\Lambda(\Lambda'\Lambda/N)^{-1}\right)\hat{\mathscr{H}}(H'\hat{H}/T)^{-1}(\Lambda'\Lambda/N)^{-1} \\
& + \frac{1}{NT}\left(\varepsilon\varepsilon' + H\Lambda'\varepsilon'\right)\left[\frac{1}{NT}\left(\varepsilon\varepsilon' H + H\Lambda'\varepsilon' H\right)(H'H/T)^{-1}(\Lambda'\Lambda/N)^{-1}(I_k + \hat{B})\right] \\
& \times\hat{\mathscr{H}}(H'\hat{H}/T)^{-1}(\Lambda'\Lambda/N)^{-1} \\
& + \frac{1}{NT}\left(\varepsilon\varepsilon' + H\Lambda'\varepsilon'\right)\left[\frac{1}{NT}\left(\varepsilon\varepsilon' + H\Lambda'\varepsilon'\right)\left(\hat{H}\hat{\mathscr{H}}^{-1} - H\right)\hat{\mathscr{H}}(H'\hat{H}/T)^{-1}(\Lambda'\Lambda/N)^{-1}\right] \\
& \times\hat{\mathscr{H}}(H'\hat{H}/T)^{-1}(\Lambda'\Lambda/N)^{-1}.
\end{aligned}
$$

By using that $\hat{\mathscr{H}}(H'\hat{H}/T)^{-1}(\Lambda'\Lambda/N)^{-1} = (H'H/T)^{-1}(\Lambda'\Lambda/N)^{-1}(I_k + \hat{B})$ from (C.22) and (C.24), we get the expansion:

$$
\begin{aligned}
\hat{H}\hat{\mathscr{H}}^{-1} - H \quad = \quad & \frac{1}{N}\varepsilon\Lambda(\Lambda'\Lambda/N)^{-1} + \frac{1}{NT}\left(\varepsilon\varepsilon' H + H\Lambda'\varepsilon' H\right)(H'H/T)^{-1}(\Lambda'\Lambda/N)^{-1}(I_k + \hat{B}) \\
& + \frac{1}{NT}\left(\varepsilon\varepsilon' + H\Lambda'\varepsilon'\right)\left(\frac{1}{N}\varepsilon\Lambda\right)(\Lambda'\Lambda/N)^{-1}(H'H/T)^{-1}(\Lambda'\Lambda/N)^{-1}(I_k + \hat{B}) \\
& + \frac{1}{NT}\left(\varepsilon\varepsilon' + H\Lambda'\varepsilon'\right)\frac{1}{NT}\left(\varepsilon\varepsilon' H + H\Lambda'\varepsilon' H\right)\left[(H'H/T)^{-1}(\Lambda'\Lambda/N)^{-1}(I_k + \hat{B})\right]^2 \\
& + \hat{\mathcal{R}},
\end{aligned}
$$

where the higher-order remainder term $\hat{\mathcal{R}}$ is defined in (C.11).
Let us rewrite the expansion as:

$$
\begin{aligned}
\hat{H}\hat{\mathscr{H}}^{-1} - H \quad = \quad & \frac{1}{N}\varepsilon\Lambda(\Lambda'\Lambda/N)^{-1} + \frac{1}{NT}\left[\varepsilon\varepsilon' H + H(\Lambda'\varepsilon' H)\right](H'H/T)^{-1}(\Lambda'\Lambda/N)^{-1}(I_k + \hat{B}) \\
& + \frac{1}{N^2T}\left[\varepsilon\varepsilon'\varepsilon\Lambda + H(\Lambda'\varepsilon'\varepsilon\Lambda)\right](\Lambda'\Lambda/N)^{-1}(H'H/T)^{-1}(\Lambda'\Lambda/N)^{-1}(I_k + \hat{B}) \\
& + \frac{1}{N^2T^2}\left\{(\varepsilon\varepsilon')(\varepsilon\varepsilon' H) + (\varepsilon\varepsilon' H)(\Lambda'\varepsilon' H) + H[(\Lambda'\varepsilon'\varepsilon\varepsilon' H) + (\Lambda'\varepsilon' H)^2]\right\} \\
& \times\left[(H'H/T)^{-1}(\Lambda'\Lambda/N)^{-1}(I_k + \hat{B})\right]^2 + \hat{\mathcal{R}}. \qquad \text{(C.25)}
\end{aligned}
$$

We have:

$$
\begin{aligned}
\frac{1}{\sqrt{NT}}\Lambda'\varepsilon' H \quad & = \quad \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{t=1}^{T}\lambda_i\varepsilon_{i,t}h_t' = \Pi_1', \\
\frac{1}{NT}\Lambda'\varepsilon'\varepsilon\Lambda \quad & = \quad \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{\ell=1}^{N}\lambda_i\varepsilon_{i,t}\varepsilon_{\ell,t}\lambda_\ell' = \frac{1}{T}\sum_{t=1}^{T}\xi_t\xi_t' = \Pi_2',
\end{aligned}
$$

and

$$\begin{aligned}
\frac{1}{NT\sqrt{NT}}\Lambda'\varepsilon'\varepsilon\varepsilon' H &= \frac{1}{NT\sqrt{NT}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{\ell=1}^{N}\sum_{s=1}^{T}\lambda_i\varepsilon_{i,t}\varepsilon_{\ell,t}\varepsilon_{\ell,s}h_s' \\
&= \frac{1}{NT\sqrt{T}}\sum_{t=1}^{T}\sum_{\ell=1}^{N}\sum_{s=1}^{T}\xi_t\varepsilon_{\ell,t}\varepsilon_{\ell,s}h_s' \\
&= \frac{1}{NT\sqrt{T}}\sum_{t=1}^{T}\sum_{\ell=1}^{N}\sum_{s=1,s\neq t}^{T}\xi_t\varepsilon_{\ell,t}\varepsilon_{\ell,s}h_s' + \frac{1}{NT\sqrt{T}}\sum_{t=1}^{T}\sum_{\ell=1}^{N}\xi_t\varepsilon_{\ell,t}^2 h_t' \\
&= \frac{1}{\sqrt{N}}(\frac{1}{T}\sum_{t=1}^{T}\xi_t\alpha_t') + \frac{1}{T}(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\xi_t\eta_t^2 h_t') + \frac{1}{\sqrt{NT}}(\frac{1}{T}\sum_{t=1}^{T}\xi_t\kappa_t h_t') = \Pi_3',
\end{aligned}$$

where matrices Π_1, Π_2 and Π_3 are defined in (C.12). Let us now write expansion (C.25) for each date t. We denote by $a_t = [A]_t$ the column vector corresponding to the t-th row a_t' of matrix A. We have:

$$\begin{aligned}
\frac{1}{NT}[\varepsilon\varepsilon' H]_t &= \frac{1}{NT}\sum_{i=1}^{N}\sum_{s=1}^{T}\varepsilon_{i,t}\varepsilon_{i,s}h_s = \frac{1}{NT}\sum_{i=1}^{N}\varepsilon_{i,t}^2 h_t + \frac{1}{NT}\sum_{i=1}^{N}\sum_{s=1,s\neq t}^{T}\varepsilon_{i,t}\varepsilon_{i,s}h_s \\
&= \frac{1}{T}\eta_t^2 h_t + \frac{1}{T\sqrt{N}}\kappa_t h_t + \frac{1}{\sqrt{NT}}\alpha_t, \qquad \text{(C.26)} \\
\frac{1}{N^2T}[\varepsilon\varepsilon'\varepsilon\Lambda]_t &= \frac{1}{N^2T}\sum_{i=1}^{N}\sum_{s=1}^{T}\sum_{\ell=1}^{N}\varepsilon_{i,t}\varepsilon_{i,s}\varepsilon_{\ell,s}\lambda_\ell = \frac{1}{N\sqrt{NT}}\sum_{i=1}^{N}\sum_{s=1}^{T}\varepsilon_{i,t}\varepsilon_{i,s}\xi_s \\
&= \frac{1}{N\sqrt{NT}}\sum_{i=1}^{N}\varepsilon_{i,t}^2\xi_t + \frac{1}{N\sqrt{NT}}\sum_{i=1}^{N}\sum_{s=1,s\neq t}^{T}\varepsilon_{i,t}\varepsilon_{i,s}\xi_s \\
&= \frac{1}{\sqrt{NT}}\eta_t^2\xi_t + \frac{1}{NT}\kappa_t\xi_t + \frac{1}{N\sqrt{N}}\varphi_t + \frac{1}{N\sqrt{T}}\gamma_t,
\end{aligned}$$

and:

$$\begin{aligned}
\frac{1}{N^2T^2}[\varepsilon\varepsilon'\varepsilon\varepsilon' H]_t &= \frac{1}{N^2T^2}\sum_{s=1}^{T}\sum_{i=1}^{N}\varepsilon_{i,t}\varepsilon_{i,s}[\varepsilon\varepsilon' H]_s \\
&= \frac{1}{NT}\sum_{s=1}^{T}\sum_{i=1}^{N}\varepsilon_{i,t}\varepsilon_{i,s}\left(\frac{1}{T}\eta_s^2 h_s + \frac{1}{T\sqrt{N}}\kappa_s h_s + \frac{1}{\sqrt{NT}}\alpha_s\right) \\
&= \frac{1}{T^2}\eta_t^4 h_t + \frac{1}{\sqrt{N}T^2}\kappa_t\eta_t^2 h_t + \frac{1}{\sqrt{NT}T}\bar{\alpha}_t \\
&\quad + \frac{1}{T^2\sqrt{N}}\kappa_t\eta_t^2 h_t + \frac{1}{NT^2}\kappa_t^2 h_t + \frac{1}{NT\sqrt{N}}\bar{\varphi}_t + \frac{1}{NT\sqrt{T}}\bar{\gamma}_t \\
&\quad + \frac{1}{T\sqrt{NT}}\eta_t^2\alpha_t + \frac{1}{NT\sqrt{T}}\kappa_t\alpha_t + \frac{1}{N\sqrt{NT}}\delta_t + \frac{1}{NT}\chi_t,
\end{aligned}$$

using the definition of the processes in (C.13).

The expansion in (C.25) for date t reads:

$$\begin{aligned}(\hat{\mathscr{H}}')^{-1}\hat{h}_t - h_t &= \frac{1}{\sqrt{N}}(\Lambda'\Lambda/N)^{-1}\xi_t + (I_k+\hat{B}')(\Lambda'\Lambda/N)^{-1}(H'H/T)^{-1}\frac{1}{NT}[\varepsilon\varepsilon' H]_t \\ &+\frac{1}{\sqrt{NT}}(I_k+\hat{B}')(\Lambda'\Lambda/N)^{-1}(H'H/T)^{-1}\Pi_1 h_t \\ &+(I_k+\hat{B}')(\Lambda'\Lambda/N)^{-1}(H'H/T)^{-1}(\Lambda'\Lambda/N)^{-1}\frac{1}{N^2T}[\varepsilon\varepsilon'\varepsilon\Lambda]_t \\ &+\frac{1}{N}(I_k+\hat{B}')(\Lambda'\Lambda/N)^{-1}(H'H/T)^{-1}(\Lambda'\Lambda/N)^{-1}\Pi_2 h_t \\ &+[(I_k+\hat{B}')(\Lambda'\Lambda/N)^{-1}(H'H/T)^{-1}]^2\Pi_1\frac{1}{NT\sqrt{NT}}[\varepsilon\varepsilon' H]_t \\ &+[(I_k+\hat{B}')(\Lambda'\Lambda/N)^{-1}(H'H/T)^{-1}]^2\frac{1}{N^2T^2}[\varepsilon\varepsilon'\varepsilon\varepsilon' H]_t \\ &+[(I_k+\hat{B}')(\Lambda'\Lambda/N)^{-1}(H'H/T)^{-1}]^2[\frac{1}{\sqrt{NT}}\Pi_3+\frac{1}{NT}\Pi_1^2]h_t+\hat{\mathcal{R}}_t.\end{aligned}$$

By plugging the expressions for $\frac{1}{NT}[\varepsilon\varepsilon' H]_t$, $\frac{1}{N^2T}[\varepsilon\varepsilon'\varepsilon\Lambda]_t$ and $\frac{1}{N^2T^2}[\varepsilon\varepsilon'\varepsilon\varepsilon' H]_t$, and rearranging terms, the expansion in (C.9) follows.

Let us now prove the convergence of matrix $\hat{\mathscr{H}}$. From Proposition C.2 and Assumption C.1, we have $o_p(1) = H'(\hat{H}-H\hat{\mathscr{H}})/T = (H'\hat{H}/T) - \hat{\mathscr{H}} + o_p(1)$, which implies:

$$(H'\hat{H}/T) = \hat{\mathscr{H}} + o_p(1). \tag{C.27}$$

By combining equations (C.19) and (C.27), and using Lemma C.5 and Assumption C.1, we get:

$$\Sigma_\lambda\hat{\mathscr{H}} = \hat{\mathscr{H}}V + o_p(1). \tag{C.28}$$

Moreover, from $\hat{H}'\hat{H}/T = I_k$, Assumption C.1, Proposition C.2 and equation (C.27), we get:

$$\hat{\mathscr{H}}'\hat{\mathscr{H}} = I_k + o_p(1). \tag{C.29}$$

Recall that V is the diagonal matrix with diagonal elements corresponding to the eigenvalues of the symmetric matrix Σ_λ. Then, if these eigenvalues are distinct, equations (C.28) and (C.29) imply that the columns of matrix $\hat{\mathscr{H}}$ converge in probability to the orthonormal eigenvectors of matrix Σ_λ. The conclusion follows.

C.3.2 Proof of Proposition C.2

By computing the norms of both sides of equation (C.21), using the triangular inequality and the Cauchy-Schwarz inequality, Lemmas C.4 and C.5, and Assumption C.1, we get:

$$\begin{aligned}\|\hat{H}\hat{\mathscr{H}}^{-1}-H\| &= O_p\left(\|\frac{1}{NT}\varepsilon\varepsilon' H\| + \|\frac{1}{NT}H\Lambda'\varepsilon' H\| + \|\frac{1}{N}\varepsilon\Lambda\|\right) \\ &+O_p\left[\left(\|\frac{1}{NT}\varepsilon\varepsilon'\| + \|\frac{1}{NT}H\Lambda'\varepsilon'\|\right)\|\hat{H}\hat{\mathscr{H}}^{-1}-H\|\right].\end{aligned} \tag{C.30}$$

To control the term in the r.h.s. we use the next lemma.

LEMMA C.6. *Under Assumptions C.1 and C.2, we have: (i)* $\|\frac{1}{N}\varepsilon\Lambda\| = O_p\left(\sqrt{\frac{T}{N}}\right)$, *(ii)* $\|\frac{1}{NT}H\Lambda'\varepsilon'\| =$

$O_p\left(\frac{1}{\sqrt{N}}\right)$, *(iii)* $\|\frac{1}{NT}H\Lambda'\varepsilon' H\| = O_p\left(\frac{1}{\sqrt{N}}\right)$.

By multiplying both sides of equation (C.30) times $T^{-1/2}$, and using Assumption C.2 ii)-iii) and Lemma C.6, we get:

$$T^{-1/2}\|\hat{H}\hat{\mathscr{H}}^{-1} - H\| \;=\; O_p\left(\frac{1}{\sqrt{N}} + \frac{1}{\sqrt{NT}} + \frac{1}{\sqrt{Tm}}\right) + o_p(T^{-1/2}\|\hat{H}\hat{\mathscr{H}}^{-1} - H\|),$$

where $m = \min\{N, T\}$, that is:

$$T^{-1/2}\|\hat{H}\hat{\mathscr{H}}^{-1} - H\| \;=\; O_p\left(\frac{1}{\sqrt{N}} + \frac{1}{T}\right). \tag{C.31}$$

From equations (C.11) and (C.31), Assumptions C.1 and C.2 ii), and Lemmas C.4, C.5 and C.6 ii), and the Cauchy-Schwarz inequality we have:

$$\begin{aligned} T^{-1/2}\|\hat{\mathcal{R}}\| &= O_p\left[(\|\frac{1}{NT}\varepsilon\varepsilon'\|^2 + \|\frac{1}{NT}H\Lambda'\varepsilon'\|^2)(T^{-1/2}\|\hat{H}\hat{\mathscr{H}}^{-1} - H\|)\right] \\ &= O_p\left[(\frac{1}{m} + \frac{1}{N})(\frac{1}{\sqrt{N}} + \frac{1}{T})\right] \\ &= O_p\left(\frac{1}{N} + \frac{1}{T^2}\right), \end{aligned}$$

where $m = \min\{N, T\}$ and we use $\frac{1}{m\sqrt{N}} = O(\frac{1}{N} + \frac{1}{T^2})$ and $\frac{1}{mT} = O(\frac{1}{N} + \frac{1}{T^2})$.

C.3.3 Proof of Proposition C.3

Let us first establish the MSE bound for remainder term r_t. From its definition in (C.10) and Assumption C.3 we have

$$\begin{aligned} (\frac{1}{T}\sum_{t=1}^{T}\|r_t\|^2)^{1/2} &= O_p\left[\|\hat{B}\|(\frac{1}{T} + \frac{1}{\sqrt{NT}})\right] + O_p(\frac{1}{N} + \frac{1}{T^2}) \\ &= O_p\left[(\frac{1}{\sqrt{N}} + \frac{1}{T})(\frac{1}{T} + \frac{1}{\sqrt{NT}})\right] + O_p(\frac{1}{N} + \frac{1}{T^2}) \\ &= O_p(\frac{1}{N} + \frac{1}{T^2}). \end{aligned} \tag{C.32}$$

Let us now show that $\frac{1}{T}\sum_{t=1}^{T}\vartheta_t h_t' = O_p(\frac{1}{N} + \frac{1}{T^2})$. We use $\vartheta_t = \tilde{\vartheta}_t + \hat{\mathcal{R}}_t$ where $\tilde{\vartheta}_t = \frac{1}{\sqrt{NT}}S\alpha_t + \frac{1}{N}D_2 h_t + r_t$. From the Cauchy-Schwarz inequality and the bound in (C.15), we have:

$$\frac{1}{T}\sum_{t=1}^{T}\vartheta_t h_t' \;=\; \frac{1}{T}\sum_{t=1}^{T}\tilde{\vartheta}_t h_t' + \frac{1}{T}\sum_{t=1}^{T}\hat{\mathcal{R}}_t h_t' = \frac{1}{T}\sum_{t=1}^{T}\tilde{\vartheta}_t h_t' + O_p\left(\frac{1}{N} + \frac{1}{T^2}\right).$$

Moreover, by using Assumption C.3, bound (C.32) and $\frac{1}{\sqrt{NT}} = O(\frac{1}{N} + \frac{1}{T^2})$, we have:

$$\frac{1}{T}\sum_{t=1}^{T}\tilde{\vartheta}_t h_t' \;=\; \frac{1}{\sqrt{NT}}S(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\alpha_t h_t') + O_p(\frac{1}{N} + \frac{1}{T^2}) = O_p(\frac{1}{N} + \frac{1}{T^2}).$$

Then, $\frac{1}{T}\sum_{t=1}^{T}\vartheta_t h_t' = O_p(\frac{1}{N}+\frac{1}{T^2})$ follows.
Let us finally show that $\frac{1}{T}\sum_{t=1}^{T}(\frac{1}{\sqrt{N}}u_t + \frac{1}{T}b_t + \frac{1}{\sqrt{NT}}d_t + \vartheta_t)\vartheta_t' = o_p(\frac{1}{N\sqrt{T}})$. We have:

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T}(\frac{1}{\sqrt{N}}u_t + \frac{1}{T}b_t + \frac{1}{\sqrt{NT}}d_t + \vartheta_t)\vartheta_t' &= \frac{1}{T}\sum_{t=1}^{T}(\frac{1}{\sqrt{N}}u_t + \frac{1}{T}b_t + \frac{1}{\sqrt{NT}}d_t + \tilde{\vartheta}_t)\tilde{\vartheta}_t' \\ &+\frac{1}{T}\sum_{t=1}^{T}\hat{\mathcal{R}}_t\tilde{\vartheta}_t' + \frac{1}{T}\sum_{t=1}^{T}[(\hat{\mathscr{H}}')^{-1}\hat{h}_t - h_t]\hat{\mathcal{R}}_t'.\end{aligned}$$

Moreover, by using bound (C.32) and Assumption C.3:

$$(\frac{1}{T}\sum_{t=1}^{T}\|\tilde{\vartheta}_t\|^2)^{1/2} = O_p(\frac{1}{N}+\frac{1}{T^2}+\frac{1}{\sqrt{NT}}), \tag{C.33}$$

and thus, from (C.15) and $\sqrt{T} \ll N \ll T^{5/2}$ (Assumption A.1), we get:

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T}\hat{\mathcal{R}}_t\tilde{\vartheta}_t' &= O_p\left[T^{-1/2}\|\hat{\mathcal{R}}\|(\frac{1}{T}\sum_{t=1}^{T}\|\tilde{\vartheta}_t\|^2)^{1/2}\right] \\ &= O_p\left[(\frac{1}{N}+\frac{1}{T^2})(\frac{1}{N}+\frac{1}{T^2}+\frac{1}{\sqrt{NT}})\right] = o_p(\frac{1}{N\sqrt{T}}).\end{aligned}$$

Further, from Proposition C.2 and (C.14):

$$\frac{1}{T}\sum_{t=1}^{T}[(\hat{\mathscr{H}}')^{-1}\hat{h}_t - h_t]\hat{\mathcal{R}}_t' = O_p\left(T^{-1/2}\|\hat{H}\hat{\mathscr{H}}^{-1} - H\|T^{-1/2}\|\hat{\mathcal{R}}\|\right) = O_p\left[(\frac{1}{m}+\frac{1}{N})(\frac{1}{N}+\frac{1}{T^2})\right] = o_p(\frac{1}{N\sqrt{T}}),$$

since $\sqrt{T} \ll N \ll T^{5/2}$. Finally, from Assumption C.3 and the bound in (C.33), we have:

$$\begin{aligned}&\frac{1}{T}\sum_{t=1}^{T}(\frac{1}{\sqrt{N}}u_t + \frac{1}{T}b_t + \frac{1}{\sqrt{NT}}d_t + \tilde{\vartheta}_t)\tilde{\vartheta}_t' \\ =& \frac{1}{T}\sum_{t=1}^{T}\frac{1}{\sqrt{N}}u_t\tilde{\vartheta}_t' + \frac{1}{T}\sum_{t=1}^{T}\frac{1}{T}b_t\tilde{\vartheta}_t' + O_p\left[(\frac{1}{\sqrt{NT}}+\frac{1}{N}+\frac{1}{T^2})^2\right] \\ =& \frac{1}{T\sqrt{N}}\sum_{t=1}^{T}u_t\tilde{\vartheta}_t' + \frac{1}{T^2}\sum_{t=1}^{T}b_t\tilde{\vartheta}_t' + o_p(\frac{1}{N\sqrt{T}}),\end{aligned}$$

and:

$$\begin{aligned}\frac{1}{T\sqrt{N}}\sum_{t=1}^{T}u_t\tilde{\vartheta}_t' &= \frac{1}{N\sqrt{T}}(\frac{1}{T}\sum_{t=1}^{T}u_t\alpha_t')S' + \frac{1}{\sqrt{NT}}(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}u_t h_t')(\frac{1}{\sqrt{NT}}D_1 + \frac{1}{N}D_2 + \frac{1}{NT}D_3)' \\ &+O_p\left[\frac{1}{\sqrt{N}}\left((\frac{1}{T}+\frac{1}{\sqrt{NT}})\|\hat{B}\| + \frac{1}{\sqrt{NT}} + \frac{1}{N\sqrt{N}} + \frac{1}{N\sqrt{T}} + \frac{1}{T^2}\right)\right] + o_p(\frac{1}{N\sqrt{T}}) \\ &= \frac{1}{N\sqrt{T}}(\Lambda'\Lambda/N)^{-1}(\frac{1}{T}\sum_{t=1}^{T}\xi_t\alpha_t')S' + o_p(\frac{1}{N\sqrt{T}}) = o_p(\frac{1}{N\sqrt{T}}),\end{aligned}$$

from Assumption C.3 (iv), and:

$$\begin{aligned}\frac{1}{T^2}\sum_{t=1}^T b_t\tilde{\vartheta}_t' &= \frac{1}{\sqrt{N}T^2}(\frac{1}{\sqrt{T}}\sum_{t=1}^T b_t\alpha_t')S' \\ &\quad +O_p\left[\frac{1}{T}\left((\frac{1}{T}+\frac{1}{\sqrt{NT}})\|\hat{B}\|+\frac{1}{\sqrt{NT}}+\frac{1}{N}+\frac{1}{N\sqrt{T}}+\frac{1}{T^2}\right)\right]+o_p(\frac{1}{N\sqrt{T}}) \\ &= \frac{1}{\sqrt{N}T^2}S(\frac{1}{\sqrt{T}}\sum_{t=1}^T \eta_t^2 h_t\alpha_t')S'+o_p(\frac{1}{N\sqrt{T}})=o_p(\frac{1}{N\sqrt{T}}),\end{aligned}$$

since $\sqrt{T}\ll N\ll T^{5/2}$. Hence, $\frac{1}{T}\sum_{t=1}^T(\frac{1}{\sqrt{N}}u_t+\frac{1}{T}b_t+\frac{1}{\sqrt{NT}}d_t+\vartheta_t)\vartheta_t'=o_p(\frac{1}{N\sqrt{T}})$ follows.

C.3.4 Proof of Lemma C.4

The proof follows closely the proof of Proposition 1 (ii) in Bai (2009). Let us denote by H^0 the matrix of true factor values, in order to distinguish it from a matrix H of generic factor values. The estimator $\hat{H}$ is obtained from minimization of the LS criterion:

$$\min_{H,\Lambda:H'H/T=I_k} tr[(Y-H\Lambda')(Y-H\Lambda')']. \tag{C.34}$$

The criterium in (C.34), after concentration w.r.t. Λ, becomes $tr(Y'M_HY)$, where $M_H=I_T-P_H$ and $P_H=H(H'H)^{-1}H'$. Let us divide the criterium by NT, and subtract its value at H^0, to get:

$$S_{NT}(H)=\frac{1}{NT}tr(Y'M_HY)-\frac{1}{NT}tr(\varepsilon'M_{H^0}\varepsilon).$$

The matrix of factor estimates $\hat{H}$ is the minimizer of function $S_{NT}(H)$ w.r.t. H such that $H'H/T=I_k$. By using $Y=H^0\Lambda'+\varepsilon$, we get:

$$S_{NT}(H)=\frac{1}{NT}tr(\Lambda H^{0\prime}M_HH^0\Lambda')+2\frac{1}{NT}tr(\Lambda H^{0\prime}M_H\varepsilon)+\frac{1}{NT}tr(\varepsilon'(P_H-P_{H^0})\varepsilon). \tag{C.35}$$

Now, let us show that the second and third terms in the RHS are $o_p(1)$ uniformly w.r.t. the (T,k) matrix H such that $H'H/T=I_k$. We follow here different arguments compared to the ones in the proof of Lemma A.1 in Bai (2009), since we deploy slightly different assumptions. We have:

$$\begin{aligned}\frac{1}{NT}tr(\Lambda H^{0\prime}M_H\varepsilon) &= \frac{1}{NT}tr(H^{0\prime}\varepsilon\Lambda)-tr\left[\frac{1}{T}H^{0\prime}H(\frac{1}{T}H'H)^{-1}\frac{1}{NT}H'\varepsilon\Lambda\right] \\ &= O_p(\|\frac{1}{NT}H^{0\prime}\varepsilon\Lambda\|)+O_p(\|\frac{1}{NT}H'\varepsilon\Lambda\|)=O_p(\|\frac{1}{\sqrt{T}N}\varepsilon\Lambda\|)=O_p(\frac{1}{\sqrt{N}}),\end{aligned}$$

and:

$$\begin{aligned}\frac{1}{NT}tr(\varepsilon'(P_H-P_{H^0})\varepsilon) &= \frac{1}{N}tr\left[\frac{1}{T}\varepsilon'H(\frac{1}{T}H'H)^{-1}\frac{1}{T}H'\varepsilon\right]-tr\left[\frac{1}{T}\varepsilon'H^0(\frac{1}{T}H^{0\prime}H^0)^{-1}\frac{1}{T}H^{0\prime}\varepsilon\right] \\ &= \frac{1}{T}tr\left[H'(\frac{1}{NT}\varepsilon\varepsilon')H\right]-\frac{1}{T}tr\left[(\frac{1}{T}H^{0\prime}H^0)^{-1}H^{0\prime}(\frac{1}{NT}\varepsilon\varepsilon')H^0\right] \\ &= O_p(\|\frac{1}{NT}\varepsilon\varepsilon'\|)=o_p(1),\end{aligned}$$

uniformly w.r.t. the (T, k) matrix H such that $H'H/T = I_k$, using Assumptions C.1 and C.2 i) and iii), Lemma C.6 i), and the invariance of the trace under cyclical permutations.
Thus, from (C.35) we get $S_{NT}(H) = \tilde{S}_{NT}(H) + o_p(1)$, where:

$$\tilde{S}_{NT}(H) = \frac{1}{NT} tr(\Lambda H^{0\prime} M_H H^0 \Lambda') = tr[(H^{0\prime} M_H H^0/T)(\Lambda'\Lambda/N)], \tag{C.36}$$

and the $o_p(1)$ term is uniform w.r.t. H such that $H'H/T = I_k$. We have:

$$\begin{aligned} \tilde{S}_{NT}(\hat{H}) &\geq 0, \\ 0 = S_{NT}(H^0) &\geq S_{NT}(\hat{H}) = \tilde{S}_{NT}(\hat{H}) + o_p(1), \end{aligned}$$

which imply $\tilde{S}_{NT}(\hat{H}) = o_p(1)$. Then, from equation (C.36), Assumption C.1 and $\hat{H}'\hat{H}/T = I_k$, it follows:

$$H^{0\prime} H^0/T - (H^{0\prime}\hat{H}/T)(\hat{H}'H^0/T) = o_p(1).$$

Thus, from Assumption C.1, we have $(H^{0\prime}\hat{H}/T)(\hat{H}'H^0/T) = I_k + o_p(1)$. Lemma C.4 follows.

C.3.5 Proof of Lemma C.5

Let us multiply both sides of equation (C.18) by $T^{-1}H'$ to get:

$$(H'\hat{H}/T)\hat{V} - (H'H/T)(\Lambda'\Lambda/N)(H'\hat{H}/T) = \frac{1}{NT^2} H'(\varepsilon\varepsilon'\hat{H} + H\Lambda'\varepsilon'\hat{H} + \varepsilon\Lambda H'\hat{H}).$$

By applying the Cauchy-Schwarz inequality, Assumption C.2 ii), Lemmas C.4 and C.6 *(i)*, and $T^{-1/2}\|\hat{H}\| = \sqrt{k}$, we get:

$$(H'\hat{H}/T)\hat{V} - (H'H/T)(\Lambda'\Lambda/N)(H'\hat{H}/T) = o_p(1).$$

Then, from Lemma C.4 and Assumption C.1, we get:

$$\begin{aligned} \hat{V} &= (H'\hat{H}/T)^{-1}(H'H/T)(\Lambda'\Lambda/N)(H'\hat{H}/T) + o_p(1) \\ &= (H'\hat{H}/T)^{-1}\Sigma_\lambda(H'\hat{H}/T) + o_p(1). \end{aligned}$$

We deduce that the eigenvalues of matrix $\hat{V}$ converge in probability to the eigenvalues of matrix Σ_λ. Since matrix $\hat{V}$ is diagonal, the conclusion follows.

C.3.6 Proof of Lemma C.6

(i) Using $\frac{1}{\sqrt{N}}[\varepsilon\Lambda]_t = \frac{1}{\sqrt{N}}\sum_{i=1}^N \lambda_i\varepsilon_{i,t} = \xi_t$ and Assumption C.2 i), we have:

$$\|\frac{1}{N}\varepsilon\Lambda\| = \frac{1}{\sqrt{N}}\left[tr\left(\sum_{t=1}^T \xi_t\xi_t'\right)\right]^{1/2} = \sqrt{\frac{T}{N}}\left[tr\left(\frac{1}{T}\sum_{t=1}^T \xi_t\xi_t'\right)\right]^{1/2} = O_p\left(\sqrt{\frac{T}{N}}\right). \tag{C.37}$$

(ii) By using (C.37) and $T^{-1/2}\|H\| = O_p(1)$, we have:

$$\|\frac{1}{NT}H\Lambda'\varepsilon'\| \leq \frac{1}{T}\|H\|\|\frac{1}{N}\varepsilon\Lambda\| = O_p\left(\frac{1}{\sqrt{N}}\right).$$

(iii) We have:

$$\|\frac{1}{NT}H\Lambda'\varepsilon'H\| \leq \frac{1}{\sqrt{N}}T^{-1/2}\|H\|\|\frac{1}{\sqrt{NT}}\Lambda'\varepsilon'H\| = O_p(\frac{1}{\sqrt{N}}),$$

by using $\frac{1}{\sqrt{NT}}\Lambda'\varepsilon'H = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\xi_t h_t' = O_p(1)$ from Assumption C.2 i).

C.3.7 Check of the conditions in Assumptions C.1-C.3

Assumption C.1 is standard in the factor literature, see e.g. Bai and Ng (2002), Stock and Watson (2002), Bai (2003). It is implied by Assumptions A.2 and A.3.

a) Check of Assumption C.2

Assumption C.2 i) is implied by Assumptions A.4 b), A.5 b) and A.6 a). Assumption C.2 ii) is implied by Assumptions A.4 a) and Assumption C.3 (which is checked below). Indeed, from (C.26) we have:

$$\begin{aligned}
\|\frac{1}{NT}\varepsilon\varepsilon'H\|^2 &= \sum_{t=1}^{T}(\frac{1}{NT}[\varepsilon\varepsilon'H]_t)'(\frac{1}{NT}[\varepsilon\varepsilon'H]_t) = \sum_{t=1}^{T}\|\frac{1}{NT}[\varepsilon\varepsilon'H]_t\|^2 \\
&\leq 2\sum_{t=1}^{T}(\frac{1}{T^2}\|\eta_t^2 h_t\|^2 + \frac{1}{T^2N}\|\kappa_t h_t\|^2 + \frac{1}{NT}\|\alpha_t\|^2) = O_p(\frac{1}{T}+\frac{1}{N}),
\end{aligned}$$

under Assumption C.3 v), since $\eta_t^2 \leq M$ (Assumption A.4 a)).
Let us now show the validity of Assumption C.2 iii). We have:

$$\begin{aligned}
\|\frac{1}{NT}\varepsilon\varepsilon'\|^2 &= \frac{1}{N^2T^2}Tr[\varepsilon\varepsilon'\varepsilon\varepsilon'] = \frac{1}{N^2T^2}\sum_{t=1}^{T}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{s=1}^{T}\varepsilon_{i,t}\varepsilon_{i,s}\varepsilon_{j,s}\varepsilon_{j,t} \\
&= \frac{1}{N^2T^2}\sum_{t=1}^{T}\sum_{i=1}^{N}\sum_{j=1}^{N}\varepsilon_{i,t}^2\varepsilon_{j,t}^2 + \frac{1}{N^2T^2}\sum_{t=1}^{T}\sum_{s=1,s\neq t}^{T}\sum_{i=1}^{N}\sum_{j=1}^{N}\varepsilon_{i,t}\varepsilon_{i,s}\varepsilon_{j,t}\varepsilon_{j,s}.
\end{aligned}$$

The first term in the RHS is $O_p(T^{-1})$ from Assumption A.4 b). Let us now consider the second term in the RHS. We have:

$$\begin{aligned}
&\frac{1}{N^2T^2}\sum_{t=1}^{T}\sum_{s=1,s\neq t}^{T}\sum_{i=1}^{N}\sum_{j=1}^{N}\varepsilon_{i,t}\varepsilon_{i,s}\varepsilon_{j,t}\varepsilon_{j,s} \\
=\ &\frac{2}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{t-1}\left(\frac{1}{N}\sum_{i=1}^{N}\varepsilon_{i,t}\varepsilon_{i,s}\right)^2 \\
=\ &\frac{2}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{t-1}\eta_{ts}^4 + \frac{4}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{t-1}\left(\frac{1}{N}\sum_{i=1}^{N}(\varepsilon_{i,t}\varepsilon_{i,s}-\eta_{ts}^2)\right)\eta_{ts}^2 + \frac{2}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{t-1}\left(\frac{1}{N}\sum_{i=1}^{N}(\varepsilon_{i,t}\varepsilon_{i,s}-\eta_{ts}^2)\right)^2,
\end{aligned}$$

where $\eta_{ts}^2 := \underset{N\to\infty}{\text{plim}}\ \frac{1}{N}\sum_{i=1}^{N}\varepsilon_{i,t}\varepsilon_{i,s}$. By taking expectations, and using the Cauchy-Schwarz inequality and Assumption A.7 a), we get that $\frac{1}{N^2T^2}\sum_{t=1}^{T}\sum_{s=1,s\neq t}^{T}\sum_{i=1}^{N}\sum_{j=1}^{N}\varepsilon_{i,t}\varepsilon_{i,s}\varepsilon_{j,t}\varepsilon_{j,s} = O_p(\frac{1}{T}+\frac{1}{N})$. Assumption C.2 iii) follows.

b) Check of the conditions in Assumption C.3

Assumption C.3 i) corresponds to Assumption A.4 a). Assumptions C.3 (ii)-(iv) are implied by Assumption A.7 b). Assumption C.3 (v) is implied by Assumptions A.5 b), c), and A.7 c).

C.4 Proof of Lemma B.1

We prove the bound for $\hat{X}_{1,2}$; the bounds for the other terms are obtained similarly. We substitute the definition $\psi_{j,t} = \frac{1}{\sqrt{N_j}} u_{j,t} + \frac{1}{T} b_{j,t} + \frac{1}{\sqrt{N_j T}} d_{j,t} + \vartheta_{j,t}$ into (B.3) and use $N_2 = N$, $N_1 = N/\mu_N^2$. We get:

$$\begin{aligned}
\hat{X}_{12} \quad = \quad & \frac{1}{T\sqrt{N}} \sum_{t=1}^{T} (h_{1,t} u'_{2,t} + \mu_N u_{1,t} h'_{2,t}) + \frac{\mu_N}{TN} \sum_{t=1}^{T} u_{1,t} u'_{2,t} \\
& + \frac{1}{T^2} \sum_{t=1}^{T} \left(h_{1,t} b'_{2,t} + b_{1,t} h'_{2,t}\right) + \frac{1}{T^2\sqrt{N}} \sum_{t=1}^{T} \left(b_{1,t} u'_{2,t} + \mu_N u_{1,t} b'_{2,t}\right) + \frac{1}{T^3} \sum_{t=1}^{T} b_{1,t} b'_{2,t} \\
& + \frac{1}{T\sqrt{NT}} \sum_{t=1}^{T} \left(h_{1,t} d'_{2,t} + \mu_N d_{1,t} h'_{2,t}\right) + \frac{\mu_N}{TN\sqrt{T}} \sum_{t=1}^{T} \left(u_{1,t} d'_{2,t} + d_{1,t} u'_{2,t}\right) \\
& + \frac{1}{T^2\sqrt{NT}} \sum_{t=1}^{T} \left(b_{1,t} d'_{2,t} + \mu_N d_{1,t} b'_{2,t}\right) + \frac{\mu_N}{NT^2} \sum_{t=1}^{T} d_{1,t} d'_{2,t} + \frac{1}{T} \sum_{t=1}^{T} (h_{1,t} \vartheta'_{2,t} + \vartheta_{1,t} h'_{2,t}) \\
& + \frac{1}{T} \sum_{t=1}^{T} \left[\left(\frac{\mu_N}{\sqrt{N}} u_{1,t} + \frac{1}{T} b_{1,t} + \frac{\mu_N}{\sqrt{NT}} d_{1,t} + \vartheta_{1,t} \right) \vartheta'_{2,t} + \vartheta_{1,t} \left(\frac{1}{\sqrt{N}} u_{2,t} + \frac{1}{T} b_{2,t} + \frac{1}{\sqrt{NT}} d_{2,t} \right)' \right].
\end{aligned} \tag{C.38}$$

To bound the terms in the r.h.s. of (C.38), we use that under Assumptions A.2-A.4, A.5 *b)-c)* and A.6 *a)* we have:

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T} h_{j,t}u'_{k,t} = O_p(1), \qquad \frac{1}{T}\sum_{t=1}^{T} u_{j,t}u'_{k,t} = O_p(1), \tag{C.39}$$

$$\frac{1}{T}\sum_{t=1}^{T} h_{j,t}b'_{k,t} = O_p(1), \tag{C.40}$$

$$\frac{1}{T}\sum_{t=1}^{T} b_{j,t}u'_{k,t} = O_p\left(\frac{1}{\sqrt{T}}\right), \tag{C.41}$$

$$\frac{1}{T}\sum_{t=1}^{T} b_{j,t}b'_{k,t} = O_p(1), \tag{C.42}$$

$$\frac{1}{T}\sum_{t=1}^{T} h_{j,t}d'_{k,t} = O_p(1), \tag{C.43}$$

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T} u_{j,t}d'_{k,t} = O_p(1), \tag{C.44}$$

$$\frac{1}{T}\sum_{t=1}^{T} b_{j,t}d'_{k,t} = O_p(1), \tag{C.45}$$

$$\frac{1}{T}\sum_{t=1}^{T} d_{j,t}d'_{k,t} = O_p(1), \tag{C.46}$$

for $j,k = 1,2$. These bounds are shown below by using the definitions of $u_{j,t}$, $b_{j,t}$, $d_{j,t}$ in Proposition 3. Therefore, the first nine summation terms in the r.h.s. of (C.38) are of order $O_p(\frac{1}{\sqrt{NT}})$, $O_p(\frac{1}{N})$, $O_p(\frac{1}{T})$, $O_p(\frac{1}{T\sqrt{NT}})$, $O_p(\frac{1}{T^2})$, $O_p(\frac{1}{\sqrt{NT}})$, $O_p(\frac{1}{NT})$, $O_p(\frac{1}{T\sqrt{NT}})$ and $O_p(\frac{1}{NT})$, respectively. From Proposition 3, the last two summation terms in the r.h.s. of (C.38) are of order $O_p(\frac{1}{N}+\frac{1}{T^2})$ and $o_p(\frac{1}{N\sqrt{T}})$, respectively. Therefore, we get $\hat{X}_{1,2} = O_p(\delta_{N,T})$, where $\delta_{N,T} = \max\{\frac{1}{N},\frac{1}{T}\} = (\min\{N,T\})^{-1}$.
Proof of (C.40). We have:

$$\frac{1}{T}\sum_{t=1}^{T} h_{j,t}b'_{k,t} = \left(\frac{1}{T}\sum_{t=1}^{T} h_{j,t}h'_{k,t}\eta^2_{k,t}\right)\left(\frac{1}{T}\sum_{t=1}^{T} h_{k,t}h'_{k,t}\right)^{-1}\left(\frac{1}{N_k}\sum_{i=1}^{N_k}\lambda_{k,i}\lambda'_{k,i}\right)^{-1}.$$

The first and second terms in the r.h.s. are $O_p(1)$ by Assumptions A.2, A.4 *b)* and A.6 *a)* and an application of a LLN for mixing processes. The third term in the r.h.s. is $O_p(1)$ by Assumption A.3. Then, (C.40) follows.
Proof of (C.41). We have:

$$b_{j,t} = \left(\frac{1}{N_j}\sum_{i=1}^{N_j}\lambda_{j,i}\lambda'_{j,i}\right)^{-1}\left(\frac{1}{T}\sum_{t=1}^{T} h_{j,t}h'_{j,t}\right)^{-1} h_{j,t}\eta^2_{j,t}, \tag{C.47}$$

and:

$$u_{j,t} = \left(\frac{1}{N_j}\sum_{i=1}^{N_j}\lambda_{j,i}\lambda'_{j,i}\right)^{-1}\xi_{j,t},$$

where $\eta_{j,t}^2$ and $\xi_{j,t}$ are defined as in Assumption A.5. Then, we have:

$$\frac{1}{T}\sum_{t=1}^{T} b_{j,t}u_{k,t}' = \frac{1}{\sqrt{T}}\left(\frac{1}{N_j}\sum_{i=1}^{N_j}\lambda_{j,i}\lambda_{j,i}'\right)^{-1}\left(\frac{1}{T}\sum_{t=1}^{T}h_{j,t}h_{j,t}'\right)^{-1}\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\eta_{j,t}^2 h_{j,t}\xi_{k,t}'\right]\left(\frac{1}{N_k}\sum_{i=1}^{N_k}\lambda_{k,i}\lambda_{k,i}'\right)^{-1}.$$

Now, $\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\eta_{j,t}^2 h_{j,t}\xi_{k,t}' = O_p(1)$ follows from the bound $\|\eta_{j,t}^2 h_{j,t}\xi_{k,t}'\|_r \leq M$ with $r > 2$ (implied by Assumptions A.4 *a)-b)* and A.5 *b)* and Cauchy-Schwarz inequality), the mixing property with size $r/(r-2)$ in A.6 *a)*, and an application of Corollary 14.3 in Davidson (1994). Then, (C.41) follows.
The proofs of the other bounds are established by similar arguments and are omitted. ■

C.5 Proof of Lemma B.2

By using $(I - X)^{-1} = I + X + X^2 + O_p(\delta_{N,T}^3)$ for $X = O_p(\delta_{N,T})$, from (B.4) and Lemma B.1 we have $\hat{R} = \left(I_{k_1} - \tilde{V}_{11}^{-1}\hat{X}_{11} + \tilde{V}_{11}^{-1}\hat{X}_{11}\tilde{V}_{11}^{-1}\hat{X}_{11}\right)\tilde{V}_{11}^{-1}\left(\tilde{V}_{12} + \hat{X}_{12}\right)\left(I_{k_2} - \tilde{V}_{22}^{-1}\hat{X}_{22} + \tilde{V}_{22}^{-1}\hat{X}_{22}\tilde{V}_{22}^{-1}\hat{X}_{22}\right)\tilde{V}_{22}^{-1}$ $\times\left(\tilde{V}_{21} + \hat{X}_{21}\right) + O_p(\delta_{N,T}^3)$ and therefore:

$$\begin{aligned}\hat{R} = {} & \tilde{V}_{11}^{-1}\tilde{V}_{12}\tilde{V}_{22}^{-1}\tilde{V}_{21} \\ & -\tilde{V}_{11}^{-1}\hat{X}_{11}\tilde{V}_{11}^{-1}\tilde{V}_{12}\tilde{V}_{22}^{-1}\tilde{V}_{21} + \tilde{V}_{11}^{-1}\hat{X}_{12}\tilde{V}_{22}^{-1}\tilde{V}_{21} - \tilde{V}_{11}^{-1}\tilde{V}_{12}\tilde{V}_{22}^{-1}\hat{X}_{22}\tilde{V}_{22}^{-1}\tilde{V}_{21} + \tilde{V}_{11}^{-1}\tilde{V}_{12}\tilde{V}_{22}^{-1}\hat{X}_{21} \\ & -\tilde{V}_{11}^{-1}\hat{X}_{11}\tilde{V}_{11}^{-1}\hat{X}_{12}\tilde{V}_{22}^{-1}\tilde{V}_{21} + \tilde{V}_{11}^{-1}\hat{X}_{11}\tilde{V}_{12}\tilde{V}_{22}^{-1}\hat{X}_{22}\tilde{V}_{22}^{-1}\tilde{V}_{21} - \tilde{V}_{11}^{-1}\hat{X}_{11}\tilde{V}_{11}^{-1}\tilde{V}_{12}\tilde{V}_{22}^{-1}\hat{X}_{21} \\ & -\tilde{V}_{11}^{-1}\hat{X}_{12}\tilde{V}_{22}^{-1}\hat{X}_{22}\tilde{V}_{22}^{-1}\tilde{V}_{21} + \tilde{V}_{11}^{-1}\hat{X}_{12}\tilde{V}_{22}^{-1}\hat{X}_{21} - \tilde{V}_{11}^{-1}\tilde{V}_{12}\tilde{V}_{22}^{-1}\hat{X}_{22}\tilde{V}_{22}^{-1}\hat{X}_{21} \\ & +\tilde{V}_{11}^{-1}\hat{X}_{11}\tilde{V}_{11}^{-1}\hat{X}_{11}\tilde{V}_{11}^{-1}\tilde{V}_{12}\tilde{V}_{22}^{-1}\tilde{V}_{21} + \tilde{V}_{11}^{-1}\tilde{V}_{12}\tilde{V}_{22}^{-1}\hat{X}_{22}\tilde{V}_{22}^{-1}\hat{X}_{22}\tilde{V}_{22}^{-1}\tilde{V}_{21} + O_p(\delta_{N,T}^3).\end{aligned}$$

By rearranging the terms, and using the definitions of $\hat{\Psi}$, $\hat{\Psi}^*$, $\tilde{B}$, $\tilde{R}$, the conclusion follows. ■

C.6 Proof of Lemma B.3

We define $\tilde{A} = \tilde{V}_{11}^{-1}\tilde{V}_{12}$ and $\tilde{B} = \tilde{V}_{22}^{-1}\tilde{V}_{21}$. Then, $\tilde{R} = \tilde{A}\tilde{B}$. Let us first derive the block form of matrix $\tilde{A}$. From the formula for the inverse of the symmetric block matrix $\tilde{V}_{11} = \begin{bmatrix}\tilde{\Sigma}_{cc} & \tilde{\Sigma}_{c1}\\ \tilde{\Sigma}_{1c} & \tilde{\Sigma}_{11}\end{bmatrix}$, we have:

$$\tilde{V}_{11}^{-1} = \begin{bmatrix}\Omega_{cc} & \Omega_{cs}\\ \Omega_{sc} & \Omega_{ss}\end{bmatrix} \tag{C.48}$$

where:

$$\Omega_{cc} = \left(\tilde{\Sigma}_{cc} - \tilde{\Sigma}_{c1}\tilde{\Sigma}_{11}^{-1}\tilde{\Sigma}_{1c}\right)^{-1} \tag{C.49}$$

$$= \tilde{\Sigma}_{cc}^{-1} + \tilde{\Sigma}_{cc}^{-1}\tilde{\Sigma}_{c1}\left(\tilde{\Sigma}_{11} - \tilde{\Sigma}_{1c}\tilde{\Sigma}_{cc}^{-1}\tilde{\Sigma}_{c1}\right)^{-1}\tilde{\Sigma}_{1c}\tilde{\Sigma}_{cc}^{-1}, \tag{C.50}$$

$$\Omega_{ss} = \left(\tilde{\Sigma}_{11} - \tilde{\Sigma}_{1c}\tilde{\Sigma}_{cc}^{-1}\tilde{\Sigma}_{c1}\right)^{-1} \tag{C.51}$$

$$= \tilde{\Sigma}_{11}^{-1} + \tilde{\Sigma}_{11}^{-1}\tilde{\Sigma}_{1c}\left(\tilde{\Sigma}_{cc} - \tilde{\Sigma}_{c1}\tilde{\Sigma}_{11}^{-1}\tilde{\Sigma}_{1c}\right)^{-1}\tilde{\Sigma}_{c1}\tilde{\Sigma}_{11}^{-1}, \tag{C.52}$$

and:

$$\Omega_{cs} = -\left(\tilde{\Sigma}_{cc} - \tilde{\Sigma}_{c1}\tilde{\Sigma}_{11}^{-1}\tilde{\Sigma}_{1c}\right)^{-1}\tilde{\Sigma}_{c1}\tilde{\Sigma}_{11}^{-1} = \Omega_{sc}', \tag{C.53}$$

$$\Omega_{sc} = -\left(\tilde{\Sigma}_{11} - \tilde{\Sigma}_{1c}\tilde{\Sigma}_{cc}^{-1}\tilde{\Sigma}_{c1}\right)^{-1}\tilde{\Sigma}_{1c}\tilde{\Sigma}_{cc}^{-1} = \Omega_{cs}'. \tag{C.54}$$

Then, by matrix multiplication we get:

$$A = \tilde{V}_{11}^{-1}\tilde{V}_{12} = \left[\begin{array}{cc} \Omega_{cc}\tilde{\Sigma}_{cc} + \Omega_{cs}\tilde{\Sigma}_{1c} & \Omega_{cc}\tilde{\Sigma}_{c2} + \Omega_{cs}\tilde{\Sigma}_{12} \\ \Omega_{sc}\tilde{\Sigma}_{cc} + \Omega_{ss}\tilde{\Sigma}_{1c} & \Omega_{sc}\tilde{\Sigma}_{c2} + \Omega_{ss}\tilde{\Sigma}_{12} \end{array}\right],$$

and from (C.49) and (C.53) we have $\Omega_{cc}\tilde{\Sigma}_{cc}+\Omega_{cs}\tilde{\Sigma}_{1c} = I_{k^c}$, from (C.51) and (C.54) we have $\Omega_{sc}\tilde{\Sigma}_{cc}+\Omega_{ss}\tilde{\Sigma}_{1c} = 0$, from (C.51) and (C.54) we have

$$\Omega_{sc}\tilde{\Sigma}_{c2} + \Omega_{ss}\tilde{\Sigma}_{12} = \left(\tilde{\Sigma}_{11} - \tilde{\Sigma}_{1c}\tilde{\Sigma}_{cc}^{-1}\tilde{\Sigma}_{c1}\right)^{-1}\left(\tilde{\Sigma}_{12} - \tilde{\Sigma}_{1c}\tilde{\Sigma}_{cc}^{-1}\tilde{\Sigma}_{c2}\right) = \tilde{\Sigma}_{11|c}^{-1}\tilde{\Sigma}_{12|c},$$

where we use the notation $\tilde{\Sigma}_{jk|\ell} := \tilde{\Sigma}_{jk} - \tilde{\Sigma}_{j\ell}\tilde{\Sigma}_{\ell\ell}^{-1}\tilde{\Sigma}_{\ell k}$ for $j, k, \ell = 1, 2, c$. Moreover, from (C.49) and (C.53) we have:

$$\Omega_{cc}\tilde{\Sigma}_{c2} + \Omega_{cs}\tilde{\Sigma}_{12} = \left(\tilde{\Sigma}_{cc} - \tilde{\Sigma}_{c1}\tilde{\Sigma}_{11}^{-1}\tilde{\Sigma}_{1c}\right)^{-1}\left(\tilde{\Sigma}_{c2} - \tilde{\Sigma}_{c1}\tilde{\Sigma}_{11}^{-1}\tilde{\Sigma}_{12}\right) = \tilde{\Sigma}_{cc|1}^{-1}\tilde{\Sigma}_{c2|1},$$

while from (C.50) and (C.54) we have:

$$\begin{aligned} \Omega_{cc}\tilde{\Sigma}_{c2} + \Omega_{cs}\tilde{\Sigma}_{12} &= \tilde{\Sigma}_{cc}^{-1}\left[\tilde{\Sigma}_{c2} - \tilde{\Sigma}_{c1}\left(\tilde{\Sigma}_{11} - \tilde{\Sigma}_{1c}\tilde{\Sigma}_{cc}^{-1}\tilde{\Sigma}_{c1}\right)^{-1}\left(\tilde{\Sigma}_{12} - \tilde{\Sigma}_{1c}\tilde{\Sigma}_{cc}^{-1}\tilde{\Sigma}_{c2}\right)\right] \\ &= \tilde{\Sigma}_{cc}^{-1}\left(\tilde{\Sigma}_{c2} - \tilde{\Sigma}_{c1}\tilde{\Sigma}_{11|c}^{-1}\tilde{\Sigma}_{12|c}\right). \end{aligned}$$

Thus, we get:

$$\begin{aligned} A &= \tilde{V}_{11}^{-1}\tilde{V}_{12} = \left[\begin{array}{cc} I_{k^c} & \tilde{\Sigma}_{cc|1}^{-1}\tilde{\Sigma}_{c2|1} \\ 0 & \tilde{\Sigma}_{11|c}^{-1}\tilde{\Sigma}_{12|c} \end{array}\right] \\ &= \left[\begin{array}{cc} I_{k^c} & \tilde{\Sigma}_{cc}^{-1}\left(\tilde{\Sigma}_{c2} - \tilde{\Sigma}_{c1}\tilde{\Sigma}_{11|c}^{-1}\tilde{\Sigma}_{12|c}\right) \\ 0 & \tilde{\Sigma}_{11|c}^{-1}\tilde{\Sigma}_{12|c} \end{array}\right]. \end{aligned} \tag{C.55}$$

We get the expression for $B = \tilde{V}_{22}^{-1}\tilde{V}_{21}$ by interchanging the indices 1 and 2:

$$\begin{aligned} B &= \tilde{V}_{22}^{-1}\tilde{V}_{21} = \left[\begin{array}{cc} I_{k^c} & \tilde{\Sigma}_{cc|2}^{-1}\tilde{\Sigma}_{c1|2} \\ 0 & \tilde{\Sigma}_{22|c}^{-1}\tilde{\Sigma}_{21|c} \end{array}\right] \\ &= \left[\begin{array}{cc} I_{k^c} & \tilde{\Sigma}_{cc}^{-1}\left(\tilde{\Sigma}_{c1} - \tilde{\Sigma}_{c2}\tilde{\Sigma}_{22|c}^{-1}\tilde{\Sigma}_{21|c}\right) \\ 0 & \tilde{\Sigma}_{22|c}^{-1}\tilde{\Sigma}_{21|c} \end{array}\right]. \end{aligned} \tag{C.56}$$

By multiplying the matrices in (C.55) and (C.56) we get:

$$\tilde{R} = \tilde{A}\tilde{B} = \left[\begin{array}{cc} I_{k^c} & \tilde{R}_{cs} \\ 0 & \tilde{R}_{ss} \end{array}\right]$$

where

$$\tilde{R}_{ss} = \tilde{\Sigma}_{11|c}^{-1}\tilde{\Sigma}_{12|c}\tilde{\Sigma}_{22|c}^{-1}\tilde{\Sigma}_{21|c},$$

and:

$$\begin{aligned} \tilde{R}_{cs} &= \tilde{\Sigma}_{cc}^{-1}\left(\tilde{\Sigma}_{c1} - \tilde{\Sigma}_{c2}\tilde{\Sigma}_{22|c}^{-1}\tilde{\Sigma}_{21|c}\right) + \tilde{\Sigma}_{cc}^{-1}\left(\tilde{\Sigma}_{c2} - +\tilde{\Sigma}_{c1}\tilde{\Sigma}_{11|c}^{-1}\tilde{\Sigma}_{12|c}\right)\tilde{\Sigma}_{22|c}^{-1}\tilde{\Sigma}_{21|c} \\ &= \tilde{\Sigma}_{cc}^{-1}\tilde{\Sigma}_{c1}\left(I_{k_1-k^c} - \tilde{\Sigma}_{11|c}^{-1}\tilde{\Sigma}_{12|c}\tilde{\Sigma}_{22|c}^{-1}\tilde{\Sigma}_{21|c}\right) = \tilde{\Sigma}_{cc}^{-1}\tilde{\Sigma}_{c1}\left(I_{k_1-k^c} - \tilde{R}_{ss}\right). \end{aligned} \tag{C.57}$$

■

C.7 Proof of Lemma B.4

Substituting the expansions (B.5) and (B.10) into the eigenvalue-eigenvector equation (B.9), we get:

$$\left(\tilde{R} + \hat{\Psi} + O_p(\delta_{N,T}^3)\right)(E_c\hat{\mathcal{U}} + E_s\hat{\alpha}) = (E_c\hat{\mathcal{U}} + E_s\hat{\alpha})(I_{k^c} + \hat{M}).$$

By using $\tilde{R}E_c = E_c$ from Lemma B.3, and keeping only the terms up to second order in $\delta_{N,T}$, we get:

$$\tilde{R}E_s\hat{\alpha} + \hat{\Psi}E_c\hat{\mathcal{U}} + \hat{\Psi}E_s\hat{\alpha} = E_c\hat{\mathcal{U}}\hat{M} + E_s\hat{\alpha} + E_s\hat{\alpha}\hat{M} + O_p(\delta_{N,T}^3). \tag{C.58}$$

Pre-multiplying equation (C.58) by E_c', using the block notation $\hat{\Psi}_{cc} = E_c'\hat{\Psi}E_c$, $\hat{\Psi}_{cs} = E_c'\hat{\Psi}E_s$ and $\tilde{R}_{cs} = E_c'\tilde{R}E_s$ and the fact that matrix $\hat{\mathcal{U}}$ is non-singular, we get:

$$\hat{M} = \hat{\mathcal{U}}^{-1}\left(\tilde{R}_{cs}\hat{\alpha} + \hat{\Psi}_{cc}\hat{\mathcal{U}} + \hat{\Psi}_{cs}\hat{\alpha}\right) + O_p(\delta_{N,T}^3). \tag{C.59}$$

Similarly, pre-multiplying equation (C.58) by E_s', we get:

$$\tilde{R}_{ss}\hat{\alpha} + \hat{\Psi}_{sc}\hat{\mathcal{U}} + \hat{\Psi}_{ss}\hat{\alpha} = \hat{\alpha} + \hat{\alpha}\hat{M} + O_p(\delta_{N,T}^3), \tag{C.60}$$

where $\hat{\Psi}_{sc} = E_s'\hat{\Psi}E_c$ and $\hat{\Psi}_{ss} = E_s'\hat{\Psi}E_s$. To solve this equation for $\hat{\alpha}$ up to terms $O_p(\delta_{N,T}^3)$, it is instrumental to get first expansions for $\hat{M}$ and $\hat{\alpha}$ at order $O_p(\delta_{N,T}^2)$.

i) Expansions at order $O_p(\delta_{N,T}^2)$

Since $\hat{\Psi}_{ss}\hat{\alpha} = O_p(\delta_{N,T}^2)$ and $\hat{\alpha}\hat{M} = O_p(\delta_{N,T}^2)$, from (C.60) we have:

$$\hat{\alpha} = (I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\hat{\Psi}_{sc}\hat{\mathcal{U}} + O_p(\delta_{N,T}^2). \tag{C.61}$$

If we plug this into (C.59) we get:

$$\begin{aligned} \hat{M} &= \hat{\mathcal{U}}^{-1}\left(\tilde{R}_{cs}\hat{\alpha} + \hat{\Psi}_{cc}\hat{\mathcal{U}}\right) + O_p(\delta_{N,T}^2) \\ &= \hat{\mathcal{U}}^{-1}\left(\hat{\Psi}_{cc} + \tilde{R}_{cs}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\hat{\Psi}_{sc}\right)\hat{\mathcal{U}} + O_p(\delta_{N,T}^2). \end{aligned} \tag{C.62}$$

ii) Expansions at order $O_p(\delta_{N,T}^3)$

By plugging (C.61) and (C.62) into terms $\hat{\Psi}_{ss}\hat{\alpha}$ and $\hat{\alpha}\hat{M}$ in equation (C.60), and solving for $\hat{\alpha}$, we get:

$$\begin{aligned}\hat{\alpha} &= (I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\Big[\hat{\Psi}_{sc} + \hat{\Psi}_{ss}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\hat{\Psi}_{sc} \\ &\quad -(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\hat{\Psi}_{sc}\Big(\hat{\Psi}_{cc} + \tilde{R}_{cs}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\hat{\Psi}_{sc}\Big)\Big]\hat{\mathcal{U}} + O_p(\delta^3_{N,T}). \qquad \text{(C.63)}\end{aligned}$$

We replace (C.63) and (C.61) for the first and second occurrences of $\hat{\alpha}$ in the r.h.s. of (C.59), respectively, and we rearrange terms to get:

$$\begin{aligned}\hat{M} &= \hat{\mathcal{U}}^{-1}\Big\{\hat{\Psi}_{cc} + \tilde{R}_{cs}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\hat{\Psi}_{sc} \\ &\quad +\tilde{R}_{cs}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\Big[\hat{\Psi}_{ss}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\hat{\Psi}_{sc} \\ &\quad -(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\hat{\Psi}_{sc}\Big(\hat{\Psi}_{cc} + \tilde{R}_{cs}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\hat{\Psi}_{sc}\Big)\Big] \\ &\quad +\hat{\Psi}_{cs}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\hat{\Psi}_{sc}\Big\}\hat{\mathcal{U}} + O_p(\delta^3_{N,T}). \qquad \text{(C.64)}\end{aligned}$$

Substituting the second-order approximation of $\hat{\alpha}$ from equation (C.63) into the equation for $\hat{W}_1^*$ in (B.10), we get:

$$\begin{aligned}\hat{W}_1^* &= \Big(E_c + E_s(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\Big[\hat{\Psi}_{sc} + \hat{\Psi}_{ss}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\hat{\Psi}_{sc} \\ &\quad -(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\hat{\Psi}_{sc}\Big(\hat{\Psi}_{cc} + \tilde{R}_{cs}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\hat{\Psi}_{sc}\Big)\Big]\Big)\hat{\mathcal{U}} + O_p(\delta^3_{N,T}). \qquad \text{(C.65)}\end{aligned}$$

By replacing equation (C.64) into (B.10), we get the asymptotic expansion of $\hat{\Lambda}$:

$$\begin{aligned}\hat{\Lambda} &= I_{k^c} + \hat{\mathcal{U}}^{-1}\Big\{\hat{\Psi}_{cc} + \tilde{R}_{cs}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\hat{\Psi}_{sc} \\ &\quad +\tilde{R}_{cs}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\Big[\hat{\Psi}_{ss}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\hat{\Psi}_{sc} \\ &\quad -(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\hat{\Psi}_{sc}\Big(\hat{\Psi}_{cc} + \tilde{R}_{cs}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\hat{\Psi}_{sc}\Big)\Big] \\ &\quad +\hat{\Psi}_{cs}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\hat{\Psi}_{sc}\Big\}\hat{\mathcal{U}} + O_p(\delta^3_{N,T}). \qquad \text{(C.66)}\end{aligned}$$

Note that this approximation holds for the terms in the main diagonal, as matrix $\hat{\Lambda}$ has been defined to be diagonal.
The asymptotic expansions of $\hat{\Lambda}$ and $\hat{W}_1^*$ can be further simplified by using the next equation:

$$\tilde{R}_{cs}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1} = \tilde{\Sigma}_{cc}^{-1}\tilde{\Sigma}_{c,1}, \qquad \text{(C.67)}$$

which follows from (C.57). Equation (C.67) and equation $\tilde{V}_{11}\hat{\Psi} = \hat{\Psi}^*$ (Lemma B.2) imply:

$$\hat{\Psi}_{cc} + \tilde{R}_{cs}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\hat{\Psi}_{sc} = \tilde{\Sigma}_{cc}^{-1}\Big[\tilde{\Sigma}_{cc}\hat{\Psi}_{cc} + \tilde{\Sigma}_{c,1}\hat{\Psi}_{sc}\Big] = \tilde{\Sigma}_{cc}^{-1}\Big(\tilde{V}_{11}\hat{\Psi}\Big)_{(cc)} = \tilde{\Sigma}_{cc}^{-1}\hat{\Psi}^*_{cc},$$

$M_{(cc)}$ denoting the upper-left (k^c, k^c) block of matrix M. Using the latter equation as well as $\tilde{\Sigma}_{cc}\hat{\Psi}_{cs} + \tilde{\Sigma}_{c,1}\hat{\Psi}_{ss}$

= $\hat{\Psi}^*_{cs}$ in (C.66), and rearranging terms, we get:

$$\hat{\Lambda} = I_{k^c} + \hat{\mathcal{U}}^{-1}\tilde{\Sigma}_{cc}^{-1}\left\{\hat{\Psi}^*_{cc} + \hat{\Psi}^*_{cs}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\hat{\Psi}_{sc} - \tilde{\Sigma}_{c,1}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\hat{\Psi}_{sc}\tilde{\Sigma}_{cc}^{-1}\hat{\Psi}^*_{cc}\right\}\hat{\mathcal{U}} + O_p(\delta^3_{N,T}), \tag{C.68}$$

which yields equation (B.11). A similar argument yields (B.12) from (C.65). ■

C.8 Proof of Lemma B.5

The proof is based on the asymptotic expansions of the terms within the trace operator in the r.h.s. of equation (B.13). We distinguish the terms that are first-order, resp. second-order, with respect to the $\hat{X}_{j,k}$.

i) Asymptotic expansion of first-order term $\hat{\Psi}^{*(I)}_{cc}$

From equation (B.6), we have $\hat{\Psi}^{*(I)}_{cc} = \left[-\hat{X}_{11}\tilde{R} + \hat{X}_{12}\tilde{B} - \tilde{B}'\hat{X}_{22}\tilde{B} + \tilde{B}'\hat{X}_{21}\right]^{(cc)}$. As matrices $\tilde{R}$ and $\tilde{B}$ have the same structure $[\ E_c \ \vdots \ * \ ]$ (see Lemma B.3), we have:

$$\hat{\Psi}^{*(I)}_{cc} = -\hat{X}^{(cc)}_{11} + \hat{X}^{(cc)}_{12} - \hat{X}^{(cc)}_{22} + \hat{X}^{(cc)}_{21}. \tag{C.69}$$

From the expressions of the matrices $\hat{X}_{j,k}$ in (B.3), and using the fact that upper k^c-dimensional subvector of both $h_{1,t}$ and $h_{2,t}$ is f^c_t, the upper-left (k^c, k^c) blocks of the first and second matrices in the r.h.s. vanish. Therefore, from (C.69) we get:

$$\hat{\Psi}^{*(I)}_{cc} = -\frac{1}{T}\sum_{t=1}^{T}(\psi^{(c)}_{1,t} - \psi^{(c)}_{2,t})(\psi^{(c)}_{1,t} - \psi^{(c)}_{2,t})', \tag{C.70}$$

where $\psi^{(c)}_{j,t}$ denotes the upper $(k^c, 1)$ block of vector $\psi_{j,t}$. To compute the matrix in the r.h.s., we plug the expressions $\psi_{j,t} = \frac{1}{\sqrt{N_j}}u_{j,t} + \frac{1}{T}b_{j,t} + \frac{1}{\sqrt{N_j T}}d_{j,t} + \vartheta_{j,t}$ for $j = 1, 2$ from (B.1), and use Proposition 3 and Assumptions A.1-A.4, A.5 *b)-c)* and A.6 *a)* to bound negligible terms up to $o_p(\epsilon_{N,T})$, where $\epsilon_{N,T} = (N\sqrt{T})^{-1}$.

LEMMA C.7. *Under Assumptions A.1-A.4, A.5 b)-c), A.6 a) and A.7 we have:*

$$\begin{aligned}
\hat{\Psi}^{*(I)}_{cc} = & -\frac{1}{N}\left(\frac{1}{T}\sum_{t=1}^{T}E[(\mu_N u^{(c)}_{1t} - u^{(c)}_{2t})(\mu_N u^{(c)}_{1t} - u^{(c)}_{2t})'|\mathcal{F}_t]\right) \\
& -\frac{1}{N\sqrt{T}}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[(\mu_N u^{(c)}_{1t} - u^{(c)}_{2t})(\mu_N u^{(c)}_{1t} - u^{(c)}_{2t})' - E[(\mu_N u^{(c)}_{1t} - u^{(c)}_{2t})(\mu_N u^{(c)}_{1t} - u^{(c)}_{2t})'|\mathcal{F}_t]\right]\right) \\
& -\frac{1}{T\sqrt{NT}}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[(\bar{b}^{(c)}_{1,t} - \bar{b}^{(c)}_{2,t})(\mu_N u^{(c)}_{1,t} - u^{(c)}_{2,t})' + (\mu_N u^{(c)}_{1,t} - u^{(c)}_{2,t})(\bar{b}^{(c)}_{1,t} - \bar{b}^{(c)}_{2,t})'\right]\right) \\
& -\frac{1}{T^2}\left(\frac{1}{T}\sum_{t=1}^{T}(b^{(c)}_{1,t} - b^{(c)}_{2,t})(b^{(c)}_{1,t} - b^{(c)}_{2,t})'\right) \\
& -\frac{1}{T\sqrt{NT}}\left(\frac{1}{T}\sum_{t=1}^{T}\left[(\bar{b}^{(c)}_{1,t} - \bar{b}^{(c)}_{2,t})(\mu_N d^{(c)}_{1,t} - d^{(c)}_{2,t})' + (\mu_N d^{(c)}_{1,t} - d^{(c)}_{2,t})(\bar{b}^{(c)}_{1,t} - \bar{b}^{(c)}_{2,t})'\right]\right) + o_p(\epsilon_{N,T}),
\end{aligned}$$

where $\bar{b}_{j,t}$ is defined as in Theorem 1. The terms in the parentheses are $O_p(1)$.

Lemma C.7 shows that the leading stochastic terms in $\hat{\Psi}^{*(I)}_{cc}$ are of order $O_p\left(\frac{1}{N}\right)$, $O_p\left(\frac{1}{N\sqrt{T}}\right)$, $O_p\left(\frac{1}{T\sqrt{NT}}\right)$ and $O_p\left(\frac{1}{T^2}\right)$.

ii) Asymptotic expansion of the second-order terms in the r.h.s. of (B.13)

The asymptotic expansion of the second-order term $\hat{\Psi}^{*(II)}_{cc} - \frac{1}{4}\hat{\Psi}^{*(I)}_{cc}\tilde{\Sigma}^{-1}_{cc}\hat{\Psi}^{*(I)}_{cc} + \hat{\Psi}^{*(I)}_{cs}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\hat{\Psi}^{(I)}_{sc} - \tilde{\Sigma}_{c,1}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\hat{\Psi}^{(I)}_{sc}\tilde{\Sigma}^{-1}_{cc}\hat{\Psi}^{*(I)}_{cc}$ is provided in the next lemma.

LEMMA C.8. *Under Assumptions A.1-A.4, A.5 b)-c), A.6 a) and A.7 we have:*

$$
\begin{aligned}
&\hat{\Psi}^{*(II)}_{cc} - \frac{1}{4}\hat{\Psi}^{*(I)}_{cc}\tilde{\Sigma}^{-1}_{cc}\hat{\Psi}^{*(I)}_{cc} + \hat{\Psi}^{*(I)}_{cs}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\hat{\Psi}^{(I)}_{sc} - \tilde{\Sigma}_{c,1}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\hat{\Psi}^{(I)}_{sc}\tilde{\Sigma}^{-1}_{cc}\hat{\Psi}^{*(I)}_{cc} \\
= \;& \frac{1}{T^2}\left\{\left[\frac{1}{T}\sum_{t=1}^{T}\left(b^{(c)}_{1,t} - b^{(c)}_{2,t}\right)F_t'\right]\tilde{\Sigma}^{-1}_F\left[\frac{1}{T}\sum_{t=1}^{T}F_t\left(b^{(c)}_{1,t} - b^{(c)}_{2,t}\right)'\right]\right\} \\
&+\frac{1}{T\sqrt{NT}}\left\{\left(E\left[\left(\bar{b}^{(c)}_{1,t} - \bar{b}^{(c)}_{2,t}\right)F_t'\right]\Sigma_F^{-1}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}F_t\left(\mu_N u^{(c)}_{1,t} - u^{(c)}_{2,t}\right)'\right.\right. \\
&\qquad\left.\left.+E\left[\left(\bar{b}^{(c)}_{1,t} - \bar{b}^{(c)}_{2,t}\right)F_t'\right]\Sigma_F^{-1}\frac{1}{T}\sum_{t=1}^{T}F_t\left(\mu_N d^{(c)}_{1,t} - d^{(c)}_{2,t}\right)'\right)^{+}\right\} + o_p(\epsilon_{N,T}),
\end{aligned}
$$

where $\tilde{\Sigma}_F = \frac{1}{T}\sum_{t=1}^{T}F_tF_t'$ and $A^+ := A + A'$. The terms in the curly brackets are $O_p(1)$.

From Lemmas C.7 and C.8, the asymptotic expansion of the term within the square brackets in the r.h.s of (B.13) is:

$$
\begin{aligned}
&\hat{\Psi}^{*}_{cc} - \frac{1}{4}\hat{\Psi}^{*(I)}_{cc}\tilde{\Sigma}^{-1}_{cc}\hat{\Psi}^{*(I)}_{cc} + \hat{\Psi}^{*(I)}_{cs}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\hat{\Psi}^{(I)}_{sc} - \tilde{\Sigma}_{c,1}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\hat{\Psi}^{(I)}_{sc}\tilde{\Sigma}^{-1}_{cc}\hat{\Psi}^{*(I)}_{cc} \\
= \;& -\frac{1}{N}\left(\frac{1}{T}\sum_{t=1}^{T}E[(\mu_N u^{(c)}_{1t} - u^{(c)}_{2t})(\mu_N u^{(c)}_{1t} - u^{(c)}_{2t})'|\mathcal{F}_t]\right) - \frac{1}{T^2}\left\{\frac{1}{T}\sum_{t=1}^{T}\widetilde{\Delta b}^{(c)}_t\widetilde{\Delta b}^{(c)\prime}_t\right\} \\
&-\frac{1}{N\sqrt{T}}\left\{\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[(\mu_N u^{(c)}_{1t} - u^{(c)}_{2t})(\mu_N u^{(c)}_{1t} - u^{(c)}_{2t})' - E[(\mu_N u^{(c)}_{1t} - u^{(c)}_{2t})(\mu_N u^{(c)}_{1t} - u^{(c)}_{2t})'|\mathcal{F}_t]\right]\right\} \\
&-\frac{1}{T\sqrt{NT}}\left\{\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[\Delta b^{(c)}_t(\mu_N u^{(c)}_{1,t} - u^{(c)}_{2,t})' + (\mu_N u^{(c)}_{1,t} - u^{(c)}_{2,t})\Delta b^{(c)\prime}_t\right]\right\} \\
&-\frac{1}{T\sqrt{NT}}\left\{\frac{1}{T}\sum_{t=1}^{T}\left[\Delta b^{(c)}_t(\mu_N d^{(c)}_{1,t} - d^{(c)}_{2,t})' + (\mu_N d^{(c)}_{1,t} - d^{(c)}_{2,t})\Delta b^{(c)\prime}_t\right]\right\} + o_p\left(\epsilon_{N,T}\right),
\end{aligned}
\tag{C.71}
$$

where Δb_t and $\widetilde{\Delta b}_t$ are the population and sample residuals defined in Theorem 1. For the fifth summation term in the r.h.s., let us now check that:

$$
\frac{1}{T}\sum_{t=1}^{T}\left[\Delta b^{(c)}_t(\mu_N d^{(c)}_{1,t} - d^{(c)}_{2,t})' + (\mu_N d^{(c)}_{1,t} - d^{(c)}_{2,t})\Delta b^{(c)\prime}_t\right] = O_p\left(\frac{1}{\sqrt{T}}\right). \tag{C.72}
$$

Indeed, we have:

$$\frac{1}{T}\sum_{t=1}^{T}\Delta b_t^{(c)} h_{j,t}' = E[\Delta b_t^{(c)} h_{j,t}'] + O_p\left(\frac{1}{\sqrt{T}}\right), \tag{C.73}$$

from Assumption A.4 *b)* and A.6 *a)*, and Corollary 14.3 in Davidson (1994). Then, (C.72) follows from the definition of $d_{j,t}$, the convergence in (C.73), and equality $E[\Delta b_t^{(c)} h_{j,t}'] = 0$, for $j = 1, 2$. The latter equality holds because Δb_t is the residual of a projection on F_t, and $h_{j,t}$ is spanned by F_t.
Moreover, from Assumptions A.2, A.4 *b)* and A.6 *a)*, and Corollary 14.3 in Davidson (1994), we have:

$$\tilde{V}_{jj} = I_{k_j} + O_p(T^{-1/2}), \quad j = 1, 2, \quad \tilde{V}_{12} = \left[\begin{array}{cc} I_{k^c} & 0 \\ 0 & \Phi \end{array}\right] + O_p(T^{-1/2}). \tag{C.74}$$

By plugging (C.71)-(C.72) into (B.13), and using $\dfrac{1}{T\sqrt{T}\sqrt{NT}} = o(\epsilon_{N,T})$ when $N \ll T^3$, and $\tilde{\Sigma}_{cc} = I_{k^c} + O_p(T^{-1/2})$ from (C.74), the conclusion follows.

C.8.1 Proof of Lemma C.7

We substitute the expressions $\psi_{j,t} = \frac{1}{\sqrt{N_j}} u_{j,t} + \frac{1}{T} b_{j,t} + \frac{1}{\sqrt{N_j T}} d_{j,t} + \vartheta_{j,t}$ for $j = 1, 2$ into the r.h.s. of (C.70). We use $N_2 = N$ and $N_1 = N/\mu_N^2$, and partition vectors $u_{j,t}$ and $b_{j,t}$ in block-form as:

$$u_{j,t} = \left[\begin{array}{c} u_{jt}^{(c)} \\ u_{jt}^{(s)} \end{array}\right], \quad b_{j,t} = \left[\begin{array}{c} b_{jt}^{(c)} \\ b_{jt}^{(s)} \end{array}\right], \ \ j = 1, 2.$$

Moreover, we use that from Proposition 3 the contribution of the remainder terms $\vartheta_{j,t}$ in the r.h.s. of (C.70) is of order $o_p(\epsilon_{N,T})$, and that under Assumptions A.2-A.4, A.5 *b)-c)* and A.6 *a)* we have $\frac{1}{\sqrt{T}}\sum_{t=1}^{T} u_{j,t} d_{k,t}' = O_p(1)$ and $\frac{1}{T}\sum_{t=1}^{T} d_{j,t} d_{k,t}' = O_p(1)$ (see (C.44) and (C.46)). Therefore, we get:

$$\begin{aligned}
\hat{\Psi}_{cc}^{*(I)} \;=\; & -\frac{1}{TN}\sum_{t=1}^{T}(\mu_N u_{1,t}^{(c)} - u_{2,t}^{(c)})(\mu_N u_{1,t}^{(c)} - u_{2,t}^{(c)})' \\
& -\frac{1}{T^2\sqrt{N}}\sum_{t=1}^{T}\left[(b_{1,t}^{(c)} - b_{2,t}^{(c)})(\mu_N u_{1,t}^{(c)} - u_{2,t}^{(c)})' + (\mu_N u_{1,t}^{(c)} - u_{2,t}^{(c)})(b_{1,t}^{(c)} - b_{2,t}^{(c)})'\right] \\
& -\frac{1}{T^3}\sum_{t=1}^{T}(b_{1,t}^{(c)} - b_{2,t}^{(c)})(b_{1,t}^{(c)} - b_{2,t}^{(c)})' \\
& -\frac{1}{T^2\sqrt{NT}}\sum_{t=1}^{T}\left[(b_{1,t}^{(c)} - b_{2,t}^{(c)})(\mu_N d_{1,t}^{(c)} - d_{2,t}^{(c)})' + (\mu_N d_{1,t}^{(c)} - d_{2,t}^{(c)})(b_{1,t}^{(c)} - b_{2,t}^{(c)})'\right] + o_p(\epsilon_{N,T}).
\end{aligned}$$

By recentering the first term in the r.h.s., and highlighting the convergence rates, we have:

$$\begin{aligned}
\hat{\Psi}_{cc}^{*(I)} &= -\frac{1}{N}\left(\frac{1}{T}\sum_{t=1}^{T}E[(\mu_N u_{1t}^{(c)}-u_{2t}^{(c)})(\mu_N u_{1t}^{(c)}-u_{2t}^{(c)})'|\mathcal{F}_t]\right) \qquad \text{(C.75)}\\
&\quad -\frac{1}{N\sqrt{T}}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[(\mu_N u_{1t}^{(c)}-u_{2t}^{(c)})(\mu_N u_{1t}^{(c)}-u_{2t}^{(c)})'-E[(\mu_N u_{1t}^{(c)}-u_{2t}^{(c)})(\mu_N u_{1t}^{(c)}-u_{2t}^{(c)})'|\mathcal{F}_t]\right]\right)\\
&\quad -\frac{1}{T\sqrt{NT}}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[(b_{1,t}^{(c)}-b_{2,t}^{(c)})(\mu_N u_{1,t}^{(c)}-u_{2,t}^{(c)})'+(\mu_N u_{1,t}^{(c)}-u_{2,t}^{(c)})(b_{1,t}^{(c)}-b_{2,t}^{(c)})'\right]\right)\\
&\quad -\frac{1}{T^2}\left(\frac{1}{T}\sum_{t=1}^{T}(b_{1,t}^{(c)}-b_{2,t}^{(c)})(b_{1,t}^{(c)}-b_{2,t}^{(c)})'\right)\\
&\quad -\frac{1}{T\sqrt{NT}}\left(\frac{1}{T}\sum_{t=1}^{T}\left[(b_{1,t}^{(c)}-b_{2,t}^{(c)})(\mu_N d_{1,t}^{(c)}-d_{2,t}^{(c)})'+(\mu_N d_{1,t}^{(c)}-d_{2,t}^{(c)})(b_{1,t}^{(c)}-b_{2,t}^{(c)})'\right]\right)+o_p\left(\epsilon_{N,T}\right).
\end{aligned}$$

Finally, by using $b_{j,t}=\bar{b}_{j,t}(1+O_p(T^{-1/2}+N^{-1/2}))$, where

$$\bar{b}_{j,t}=\Sigma_{\Lambda,j}^{-1}\eta_{j,t}^2 h_{j,t},$$

and the $O_p(T^{-1/2}+N^{-1/2})$ term is independent of t, and the bound $\frac{1}{T\sqrt{NT}}(\frac{1}{\sqrt{T}}+\frac{1}{\sqrt{N}})=o(\epsilon_{N,T})$ when $N\ll T^3$, we can further simplify this asymptotic expansion to get the expansion in Lemma C.7. ■

C.8.2 Proof of Lemma C.8

i) Asymptotic expansion of $\hat{\Psi}_{cc}^{*(II)}$

Let us start with $\hat{\Psi}_{cc}^{*(II)}$. From the definitions of the matrices $\hat{X}_{j,k}$ in equation (B.3), bounding the higher-order terms as in the proof of Lemma B.1, and using that $\frac{1}{T\sqrt{NT}}\leq\frac{1}{\sqrt{NT}}\leq\frac{1}{2}\left(\frac{1}{N}+\frac{1}{T^2}\right)$, we have:

$$\hat{X}_{j,k}=\frac{1}{T}\tilde{\Xi}_{j,k}+\frac{1}{\sqrt{NT}}\hat{S}_{j,k}+O_p\left(\frac{1}{N}+\frac{1}{T^2}\right), \qquad \text{(C.76)}$$

where:

$$\begin{aligned}
\tilde{\Xi}_{j,k} &= \frac{1}{T}\sum_{t=1}^{T}(h_{j,t}b_{k,t}'+b_{j,t}h_{k,t}'), && \text{(C.77)}\\
\hat{S}_{j,k} &= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}(\mu_{N,k}h_{j,t}u_{k,t}'+\mu_{N,j}u_{j,t}h_{k,t}')+\frac{1}{T}\sum_{t=1}^{T}(\mu_{N,k}h_{j,t}d_{k,t}'+\mu_{N,j}d_{j,t}h_{k,t}'), && \text{(C.78)}
\end{aligned}$$

with $\mu_{N,1}=\mu_N$ and $\mu_{N,2}=1$. Terms $\tilde{\Xi}_{j,k}$ and $\hat{S}_{j,k}$ are $O_p(1)$ under Assumptions A.2-A.4, A.5 *b)-c)* and A.6 *a)*. Then, from the definition of $\hat{\Psi}^{*(II)}$ in (B.7), the bounds $\left(\frac{1}{T}+\frac{1}{\sqrt{NT}}\right)\left(\frac{1}{N}+\frac{1}{T^2}\right)=o(\epsilon_{N,T})$ and $\left(\frac{1}{N}+\frac{1}{T^2}\right)^2=$

$o(\epsilon_{N,T})$ which hold if $T^{1/2} \ll N \ll T^{5/2}$, we get:

$$\begin{aligned}\hat{\Psi}^{*\,(II)} &= \frac{1}{T^2}\left\{-\tilde{\Xi}_{11}\tilde{V}_{11}^{-1}\left[-\tilde{\Xi}_{11}\tilde{R}+\tilde{\Xi}_{12}\tilde{B}-\tilde{B}'\tilde{\Xi}_{22}\tilde{B}+\tilde{B}'\tilde{\Xi}_{21}\right]+\left(\tilde{\Xi}_{22}\tilde{B}-\tilde{\Xi}_{21}\right)'\tilde{V}_{22}^{-1}\left(\tilde{\Xi}_{22}\tilde{B}-\tilde{\Xi}_{21}\right)\right\}\\ &\quad +\frac{1}{T\sqrt{NT}}\left\{-\tilde{\Xi}_{11}\tilde{V}_{11}^{-1}\left[-\hat{S}_{11}\tilde{R}+\hat{S}_{12}\tilde{B}-\tilde{B}'\hat{S}_{22}\tilde{B}+\tilde{B}'\hat{S}_{21}\right]\right.\\ &\qquad -\hat{S}_{11}\tilde{V}_{11}^{-1}\left[-\tilde{\Xi}_{11}\tilde{R}+\tilde{\Xi}_{12}\tilde{B}-\tilde{B}'\tilde{\Xi}_{22}\tilde{B}+\tilde{B}'\tilde{\Xi}_{21}\right]\\ &\qquad \left.+\left(\hat{S}_{22}\tilde{B}-\hat{S}_{21}\right)'\tilde{V}_{22}^{-1}\left(\tilde{\Xi}_{22}\tilde{B}-\tilde{\Xi}_{21}\right)+\left(\tilde{\Xi}_{22}\tilde{B}-\tilde{\Xi}_{21}\right)'\tilde{V}_{22}^{-1}\left(\hat{S}_{22}\tilde{B}-\hat{S}_{21}\right)\right\}\\ &\qquad +o_p(\epsilon_{N,T}).\end{aligned}$$

The second term in the r.h.s. can be further simplified by using (C.74) and $\tilde{\Xi}_{j,k} = \Xi_{j,k} + O_p(T^{-1/2})$, where $\Xi_{j,k} = E[h_{j,t}\bar{b}'_{k,t} + \bar{b}_{j,t}h'_{k,t}]$, and neglecting terms at order $o_p(\epsilon_{N,T})$ to get:

$$\begin{aligned}\hat{\Psi}^{*\,(II)} &= \frac{1}{T^2}\left\{-\tilde{\Xi}_{11}\tilde{V}_{11}^{-1}\left[-\tilde{\Xi}_{11}\tilde{R}+\tilde{\Xi}_{12}\tilde{B}-\tilde{B}'\tilde{\Xi}_{22}\tilde{B}+\tilde{B}'\tilde{\Xi}_{21}\right]+\left(\tilde{\Xi}_{22}\tilde{B}-\tilde{\Xi}_{21}\right)'\tilde{V}_{22}^{-1}\left(\tilde{\Xi}_{22}\tilde{B}-\tilde{\Xi}_{21}\right)\right\}\\ &\quad +\frac{1}{T\sqrt{NT}}\left\{-\Xi_{11}\left[-\hat{S}_{11}R+\hat{S}_{12}B-B'\hat{S}_{22}B+B'\hat{S}_{21}\right]-\hat{S}_{11}\left[-\Xi_{11}R+\Xi_{12}B-B'\Xi_{22}B+B'\Xi_{21}\right]\right.\\ &\qquad \left.+\left(\hat{S}_{22}B-\hat{S}_{21}\right)'(\Xi_{22}B-\Xi_{21})+(\Xi_{22}B-\Xi_{21})'\left(\hat{S}_{22}B-\hat{S}_{21}\right)\right\}+o_p(\epsilon_{N,T}),\end{aligned}$$

where:

$$B=\begin{bmatrix} I_{k^c} & 0\\ 0 & \Phi'\end{bmatrix},\qquad R=\begin{bmatrix} I_{k^c} & 0\\ 0 & \Phi\Phi'\end{bmatrix}.$$

Let us now compute the (cc) block of this expansion. Since $[-\Xi_{11}R+\Xi_{12}B-B'\Xi_{22}B+B'\Xi_{21}]_{(cc)} = 0$ and $\left[-\hat{S}_{11}R+\hat{S}_{12}B-B'\hat{S}_{22}B+B'\hat{S}_{21}\right]_{(cc)} = 0$, we get:

$$\begin{aligned}\hat{\Psi}^{*\,(II)}_{cc} &= \frac{1}{T^2}\left\{-\tilde{\Xi}_{11}\tilde{V}_{11}^{-1}\left[-\tilde{\Xi}_{11}\tilde{R}+\tilde{\Xi}_{12}\tilde{B}-\tilde{B}'\tilde{\Xi}_{22}\tilde{B}+\tilde{B}'\tilde{\Xi}_{21}\right]+\left(\tilde{\Xi}_{22}\tilde{B}-\tilde{\Xi}_{21}\right)'\tilde{V}_{22}^{-1}\left(\tilde{\Xi}_{22}\tilde{B}-\tilde{\Xi}_{21}\right)\right\}_{(cc)}\\ &\quad +\frac{1}{T\sqrt{NT}}\left\{-\Xi_{11,cs}\left[-\hat{S}_{11}R+\hat{S}_{12}B-B'\hat{S}_{22}B+B'\hat{S}_{21}\right]_{sc}\right.\\ &\qquad -\hat{S}_{11,cs}\left[-\Xi_{11}R+\Xi_{12}B-B'\Xi_{22}B+B'\Xi_{21}\right]_{sc}\\ &\qquad \left.+\left[\left(\hat{S}_{22}B-\hat{S}_{21}\right)'(\Xi_{22}B-\Xi_{21})+(\Xi_{22}B-\Xi_{21})'\left(\hat{S}_{22}B-\hat{S}_{21}\right)\right]_{cc}\right\}+o_p(\epsilon_{N,T}).\end{aligned}$$

By straightforward matrix algebra we have:

$$\left[-\Xi_{11}R+\Xi_{12}B-B'\Xi_{22}B+B'\Xi_{21}\right]_{sc} = -\Xi_{11,sc}+\Xi_{12,sc}-\Phi\Xi_{22,sc}+\Phi\Xi_{21,sc}, \tag{C.79}$$

$$\left[-\hat{S}_{11}R+\hat{S}_{12}B-B'\hat{S}_{22}B+B'\hat{S}_{21}\right]_{sc} = -\hat{S}_{11,sc}+\hat{S}_{12,sc}-\Phi\hat{S}_{22,sc}+\Phi\hat{S}_{21,sc}, \tag{C.80}$$

and $\left[(\Xi_{22}B-\Xi_{21})'\left(\hat{S}_{22}B-\hat{S}_{21}\right)\right]_{(cc)} = (\Xi_{22,cc}-\Xi_{21,cc})'\left(\hat{S}_{22,cc}-\hat{S}_{21,cc}\right) + (\Xi_{22,sc}-\Xi_{21,sc})' \times$

$\left(\hat{S}_{22,sc} - \hat{S}_{21,sc}\right)$. Then we get the asymptotic expansion:

$$\begin{aligned}
\hat{\Psi}^{*\,(II)}_{cc} &= \frac{1}{T^2}\left\{-\tilde{\Xi}_{11}\tilde{V}_{11}^{-1}\left[-\tilde{\Xi}_{11}\tilde{R} + \tilde{\Xi}_{12}\tilde{B} - \tilde{B}'\tilde{\Xi}_{22}\tilde{B} + \tilde{B}'\tilde{\Xi}_{21}\right] + \left(\tilde{\Xi}_{22}\tilde{B} - \tilde{\Xi}_{21}\right)'\tilde{V}_{22}^{-1}\left(\tilde{\Xi}_{22}\tilde{B} - \tilde{\Xi}_{21}\right)\right\}_{(cc)} \\
&\quad + \frac{1}{T\sqrt{NT}}\left\{-\Xi_{11,cs}\left[-\hat{S}_{11,sc} + \hat{S}_{12,sc} - \Phi\hat{S}_{22,sc} + \Phi\hat{S}_{21,sc}\right]\right. \\
&\qquad -\hat{S}_{11,cs}\left[-\Xi_{11,sc} + \Xi_{12,sc} - \Phi\Xi_{22,sc} + \Phi\Xi_{21,sc}\right] \\
&\qquad \left. + \left((\Xi_{22,cc} - \Xi_{21,cc})'\left(\hat{S}_{22,cc} - \hat{S}_{21,cc}\right) + (\Xi_{22,sc} - \Xi_{21,sc})'\left(\hat{S}_{22,sc} - \hat{S}_{21,sc}\right)\right)^{+}\right\} \\
&\qquad + o_p(\epsilon_{N,T}),
\end{aligned} \tag{C.81}$$

where $(A)^{+} = A + A'$.

ii) Asymptotic expansion of $\hat{\Psi}^{*(I)}_{cs}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\hat{\Psi}^{(I)}_{sc}$

Let us now consider the term $\hat{\Psi}^{*(I)}_{cs}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\hat{\Psi}^{(I)}_{sc}$. By the formula of the partitioned inverse for $\tilde{V}_{11}^{-1}$, and Lemmas B.1 and C.7, we have:

$$\begin{aligned}
&\hat{\Psi}^{*(I)}_{cs}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\hat{\Psi}^{(I)}_{sc} \\
&= \hat{\Psi}^{*(I)}_{cs}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\left[(\tilde{V}_{11}^{-1})_{ss}\hat{\Psi}^{*(I)}_{sc} + O_p\left(T^{-1/2}\hat{\Psi}^{*(I)}_{cc}\right)\right] \\
&= \hat{\Psi}^{*(I)}_{cs}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}(\tilde{V}_{11}^{-1})_{ss}\hat{\Psi}^{*(I)}_{sc} + O_p\left(\delta_{N,T}\frac{1}{\sqrt{T}}\left(\frac{1}{N} + \frac{1}{T^2} + \frac{1}{T\sqrt{NT}} + \epsilon_{N,T}\right)\right) \\
&= \hat{\Psi}^{*(I)}_{cs}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}(\tilde{V}_{11}^{-1})_{ss}\hat{\Psi}^{*(I)}_{sc} + o_p(\epsilon_{N,T}),
\end{aligned} \tag{C.82}$$

if $N \ll T^{5/2}$. Let us consider $\hat{\Psi}^{*(I)}_{cs}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}(\tilde{V}_{11}^{-1})_{ss}\hat{\Psi}^{*(I)}_{sc}$. By using $\hat{\Psi}^{*(I)} = -\hat{X}_{11}\tilde{R} + \hat{X}_{12}\tilde{B} - \tilde{B}'\hat{X}_{22}\tilde{B} + \tilde{B}'\hat{X}_{21}$, the expansion for $\hat{X}_{j,k}$ in (C.76), $\tilde{R}_{ss} = \Phi\Phi' + o_p(1)$, and the condition $T^{1/2} \ll N \ll T^{5/2}$ to control negligible terms, we get:

$$\begin{aligned}
&\hat{\Psi}^{*(I)}_{cs}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}(\tilde{V}_{11}^{-1})_{ss}\hat{\Psi}^{*(I)}_{sc} \\
&= \frac{1}{T^2}\left\{\left[-\tilde{\Xi}_{11}\tilde{R} + \tilde{\Xi}_{12}\tilde{B} - \tilde{B}'\tilde{\Xi}_{22}\tilde{B} + \tilde{B}'\tilde{\Xi}_{21}\right]_{cs}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}(\tilde{V}_{11}^{-1})_{ss}\left[-\tilde{\Xi}_{11}\tilde{R} + \tilde{\Xi}_{12}\tilde{B} - \tilde{B}'\tilde{\Xi}_{22}\tilde{B} + \tilde{B}'\tilde{\Xi}_{21}\right]_{sc}\right\} \\
&\quad + \frac{1}{T\sqrt{NT}}\left\{\left[-\Xi_{11}R + \Xi_{12}B - B'\Xi_{22}B + B'\Xi_{21}\right]_{cs}(I_{k_1-k^c} - \Phi\Phi')^{-1}\left[-\hat{S}_{11}R + \hat{S}_{12}B - B'\hat{S}_{22}B + B'\hat{S}_{21}\right]_{sc}\right. \\
&\qquad \left. + \left[-\hat{S}_{11}R + \hat{S}_{12}B - B'\hat{S}_{22}B + B'\hat{S}_{21}\right]_{cs}(I_{k_1-k^c} - \Phi\Phi')^{-1}\left[-\Xi_{11}R + \Xi_{12}B - B'\Xi_{22}B + B'\Xi_{21}\right]_{sc}\right\} + o_p(\epsilon_{N,T}).
\end{aligned}$$

By using equations (C.79)-(C.80) and:

$$\begin{aligned}
\left[-\Xi_{11}R + \Xi_{12}B - B'\Xi_{22}B + B'\Xi_{21}\right]_{cs} &= -\Xi_{11,cs}\Phi\Phi' + \Xi_{12,cs}\Phi' - \Xi_{22,cs}\Phi' + \Xi_{21,cs}, \\
\left[-\hat{S}_{11}R + \hat{S}_{12}B - B'\hat{S}_{22}B + B'\hat{S}_{21}\right]_{cs} &= -\hat{S}_{11,cs}\Phi\Phi' + \hat{S}_{12,cs}\Phi' - \hat{S}_{22,cs}\Phi' + \hat{S}_{21,cs},
\end{aligned}$$

we get:

$$\begin{aligned}
&\hat{\Psi}^{*(I)}_{cs}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}(\tilde{V}_{11}^{-1})_{ss}\hat{\Psi}^{*(I)}_{sc} \\
&= \frac{1}{T^2}\left\{\left[-\tilde{\Xi}_{11}\tilde{R} + \tilde{\Xi}_{12}\tilde{B} - \tilde{B}'\tilde{\Xi}_{22}\tilde{B} + \tilde{B}'\tilde{\Xi}_{21}\right]_{cs}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}(\tilde{V}_{11}^{-1})_{ss}\left[-\tilde{\Xi}_{11}\tilde{R} + \tilde{\Xi}_{12}\tilde{B} - \tilde{B}'\tilde{\Xi}_{22}\tilde{B} + \tilde{B}'\tilde{\Xi}_{21}\right]_{sc}\right\} \\
&\quad + \frac{1}{T\sqrt{NT}}\left\{\left[-\Xi_{11,cs}\Phi\Phi' + \Xi_{12,cs}\Phi' - \Xi_{22,cs}\Phi' + \Xi_{21,cs}\right](I_{k_1-k^c} - \Phi\Phi')^{-1}\left[-\hat{S}_{11,sc} + \hat{S}_{12,sc} - \Phi\hat{S}_{22,sc} + \Phi\hat{S}_{21,sc}\right]\right. \\
&\qquad \left. + \left[-\hat{S}_{11,cs}\Phi\Phi' + \hat{S}_{12,cs}\Phi' - \hat{S}_{22,cs}\Phi' + \hat{S}_{21,cs}\right](I_{k_1-k^c} - \Phi\Phi')^{-1}\left[-\Xi_{11,sc} + \Xi_{12,sc} - \Phi\Xi_{22,sc} + \Phi\Xi_{21,sc}\right]\right\} + o_p(\epsilon_{N,T}).
\end{aligned}$$

iii) Asymptotic expansion of $\hat{\Psi}_{cc}^{*\,(II)} + \hat{\Psi}_{cs}^{*(I)}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\hat{\Psi}_{sc}^{(I)}$

By putting the expansions (C.81), (C.82) and (C.83) together, we get the asymptotic expansion:

$$
\begin{aligned}
& \hat{\Psi}_{cc}^{*\,(II)} + \hat{\Psi}_{cs}^{*(I)}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\hat{\Psi}_{sc}^{(I)} \\
= \;& \frac{1}{T^2}\Bigg\{\left(-\tilde{\Xi}_{11}\tilde{V}_{11}^{-1}\left[-\tilde{\Xi}_{11}\tilde{R} + \tilde{\Xi}_{12}\tilde{B} - \tilde{B}'\tilde{\Xi}_{22}\tilde{B} + \tilde{B}'\tilde{\Xi}_{21}\right] + \left(\tilde{\Xi}_{22}\tilde{B} - \tilde{\Xi}_{21}\right)'\tilde{V}_{22}^{-1}\left(\tilde{\Xi}_{22}\tilde{B} - \tilde{\Xi}_{21}\right)\right)_{cc} \\
& + \left[-\tilde{\Xi}_{11}\tilde{R} + \tilde{\Xi}_{12}\tilde{B} - \tilde{B}'\tilde{\Xi}_{22}\tilde{B} + \tilde{B}'\tilde{\Xi}_{21}\right]_{cs}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}(\tilde{V}_{11}^{-1})_{ss}\left[-\tilde{\Xi}_{11}\tilde{R} + \tilde{\Xi}_{12}\tilde{B} - \tilde{B}'\tilde{\Xi}_{22}\tilde{B} + \tilde{B}'\tilde{\Xi}_{21}\right]_{sc}\Bigg\} \\
& + \frac{1}{T\sqrt{NT}}\Bigg\{\left[-\Xi_{11,cs} + \Xi_{12,cs}\Phi' - \Xi_{22,cs}\Phi' + \Xi_{21,cs}\right]\left(I_{k_1-k^c} - \Phi\Phi'\right)^{-1}\left[-\hat{S}_{11,sc} + \hat{S}_{12,sc} - \Phi\hat{S}_{22,sc} + \Phi\hat{S}_{21,sc}\right] \\
& + \left[-\hat{S}_{11,cs} + \hat{S}_{12,cs}\Phi' - \hat{S}_{22,cs}\Phi' + \hat{S}_{21,cs}\right]\left(I_{k_1-k^c} - \Phi\Phi'\right)^{-1}\left[-\Xi_{11,sc} + \Xi_{12,sc} - \Phi\Xi_{22,sc} + \Phi\Xi_{21,sc}\right] \\
& + \left((\Xi_{22,cc} - \Xi_{21,cc})'(\hat{S}_{22,cc} - \hat{S}_{21,cc}) + (\Xi_{22,sc} - \Xi_{21,sc})'(\hat{S}_{22,sc} - \hat{S}_{21,sc})\right)^{+}\Bigg\} + o_p(\epsilon_{N,T}).
\end{aligned}
$$

By using:

$$
\begin{aligned}
& \left[-\Xi_{11,cs} + \Xi_{12,cs}\Phi' - \Xi_{22,cs}\Phi' + \Xi_{21,cs}\right]\left(I_{k_1-k^c} - \Phi\Phi'\right)^{-1}\left[-\hat{S}_{11,sc} + \hat{S}_{12,sc} - \Phi\hat{S}_{22,sc} + \Phi\hat{S}_{21,sc}\right] \\
& + \left[-\hat{S}_{11,cs} + \hat{S}_{12,cs}\Phi' - \hat{S}_{22,cs}\Phi' + \hat{S}_{21,cs}\right]\left(I_{k_1-k^c} - \Phi\Phi'\right)^{-1}\left[-\Xi_{11,sc} + \Xi_{12,sc} - \Phi\Xi_{22,sc} + \Phi\Xi_{21,sc}\right] \\
= \;& \left[-\Xi_{11,sc} + \Phi\Xi_{21,sc} - \Phi\Xi_{22,sc} + \Xi_{12,sc}\right]'\left(I_{k_1-k^c} - \Phi\Phi'\right)^{-1}\left[-\hat{S}_{11,sc} + \hat{S}_{12,sc} - \Phi\hat{S}_{22,sc} + \Phi\hat{S}_{21,sc}\right] \\
& + \left[-\hat{S}_{11,sc} + \Phi\hat{S}_{21,sc} - \Phi\hat{S}_{22,sc} + \hat{S}_{12,sc}\right]'\left(I_{k_1-k^c} - \Phi\Phi'\right)^{-1}\left[-\Xi_{11,sc} + \Xi_{12,sc} - \Phi\Xi_{22,sc} + \Phi\Xi_{21,sc}\right] \\
= \;& \left(\left[(\Xi_{11,sc} - \Xi_{12,sc}) - \Phi(\Xi_{21,sc} - \Xi_{22,sc})\right]'\left(I_{k_1-k^c} - \Phi\Phi'\right)^{-1}\left[(\hat{S}_{11,sc} - \hat{S}_{12,sc}) - \Phi(\hat{S}_{21,sc} - \hat{S}_{22,sc})\right]\right)^{+},
\end{aligned}
$$

since $\Xi_{j,j} = \Xi_{j,j}'$ for $j = 1, 2$, $\Xi_{1,2} = \Xi_{2,1}'$, and similarly for the $\hat{S}_{j,k}$, we get:

$$
\begin{aligned}
& \hat{\Psi}_{cc}^{*\,(II)} + \hat{\Psi}_{cs}^{*(I)}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\hat{\Psi}_{sc}^{(I)} \\
= \;& \frac{1}{T^2}\Bigg\{\left(-\tilde{\Xi}_{11}\tilde{V}_{11}^{-1}\left[-\tilde{\Xi}_{11}\tilde{R} + \tilde{\Xi}_{12}\tilde{B} - \tilde{B}'\tilde{\Xi}_{22}\tilde{B} + \tilde{B}'\tilde{\Xi}_{21}\right] + \left(\tilde{\Xi}_{22}\tilde{B} - \tilde{\Xi}_{21}\right)'\tilde{V}_{22}^{-1}\left(\tilde{\Xi}_{22}\tilde{B} - \tilde{\Xi}_{21}\right)\right)_{cc} \\
& + \left[-\tilde{\Xi}_{11}\tilde{R} + \tilde{\Xi}_{12}\tilde{B} - \tilde{B}'\tilde{\Xi}_{22}\tilde{B} + \tilde{B}'\tilde{\Xi}_{21}\right]_{cs}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}(\tilde{V}_{11}^{-1})_{ss}\left[-\tilde{\Xi}_{11}\tilde{R} + \tilde{\Xi}_{12}\tilde{B} - \tilde{B}'\tilde{\Xi}_{22}\tilde{B} + \tilde{B}'\tilde{\Xi}_{21}\right]_{sc}\Bigg\} \\
& + \frac{1}{T\sqrt{NT}}\Bigg\{\left(\left[(\Xi_{11,sc} - \Xi_{12,sc}) - \Phi(\Xi_{21,sc} - \Xi_{22,sc})\right]'\left(I_{k_1-k^c} - \Phi\Phi'\right)^{-1}\left[(\hat{S}_{11,sc} - \hat{S}_{12,sc}) - \Phi(\hat{S}_{21,sc} - \hat{S}_{22,sc})\right]\right)^{+} \\
& + \left((\Xi_{22,cc} - \Xi_{21,cc})'(\hat{S}_{22,cc} - \hat{S}_{21,cc}) + (\Xi_{22,sc} - \Xi_{21,sc})'(\hat{S}_{22,sc} - \hat{S}_{21,sc})\right)^{+}\Bigg\} + o_p(\epsilon_{N,T}).
\end{aligned}
$$

Let us now rewrite the term at order $\frac{1}{T\sqrt{NT}}$. From the definitions of matrices $\Xi_{j,k}$ and $\hat{S}_{j,k}$, we have:

$$
\begin{aligned}
& (\Xi_{22,cc} - \Xi_{21,cc})'\left(\hat{S}_{22,cc} - \hat{S}_{21,cc}\right) + (\Xi_{22,sc} - \Xi_{21,sc})'\left(\hat{S}_{22,sc} - \hat{S}_{21,sc}\right) \\
= \;& E\left[(\bar{b}_{1,t}^{(c)} - \bar{b}_{2,t}^{(c)})f_t^{c\prime}\right]\frac{1}{\sqrt{T}}\sum_{t=1}^{T} f_t^c(\mu_N u_{1,t}^{(c)} - u_{2,t}^{(c)})' + E\left[(\bar{b}_{1,t}^{(c)} - \bar{b}_{2,t}^{(c)})f_t^{c\prime}\right]\frac{1}{T}\sum_{t=1}^{T} f_t^c(\mu_N d_{1,t}^{(c)} - d_{2,t}^{(c)})' \\
& + E\left[(\bar{b}_{1,t}^{(c)} - \bar{b}_{2,t}^{(c)})f_{2,t}^{s\prime}\right]\frac{1}{\sqrt{T}}\sum_{t=1}^{T} f_{2,t}^s(\mu_N u_{1,t}^{(c)} - u_{2,t}^{(c)})' + E\left[(\bar{b}_{1,t}^{(c)} - \bar{b}_{2,t}^{(c)})f_{2,t}^{s\prime}\right]\frac{1}{T}\sum_{t=1}^{T} f_{2,t}^s(\mu_N d_{1,t}^{(c)} - d_{2,t}^{(c)})',
\end{aligned}
$$

and:

$$
\begin{aligned}
& [(\Xi_{11,sc} - \Xi_{12,sc}) - \Phi(\Xi_{21,sc} - \Xi_{22,sc})]' \left(I_{k_1-k^c} - \Phi\Phi'\right)^{-1} \left[(\hat{S}_{11,sc} - \hat{S}_{12,sc}) - \Phi(\hat{S}_{21,sc} - \hat{S}_{22,sc})\right] \\
= \; & E\left[\left(\bar{b}_{1,t}^{(c)} - \bar{b}_{2,t}^{(c)}\right)(f_{1,t}^s - \Phi f_{2,t}^s)'\right] \left(I_{k_1-k^c} - \Phi\Phi'\right)^{-1} \frac{1}{\sqrt{T}} \sum_{t=1}^T (f_{1,t}^s - \Phi f_{2,t}^s) \left(\mu_N u_{1,t}^{(c)} - u_{2,t}^{(c)}\right)' \\
& + E\left[\left(\bar{b}_{1,t}^{(c)} - \bar{b}_{2,t}^{(c)}\right)(f_{1,t}^s - \Phi f_{2,t}^s)'\right] \left(I_{k_1-k^c} - \Phi\Phi'\right)^{-1} \frac{1}{T} \sum_{t=1}^T (f_{1,t}^s - \Phi f_{2,t}^s) \left(\mu_N d_{1,t}^{(c)} - d_{2,t}^{(c)}\right)' \\
= \; & E\left[\left(\bar{b}_{1,t}^{(c)} - \bar{b}_{2,t}^{(c)}\right) f_{1\perp 2,t}^{s\prime}\right] \frac{1}{\sqrt{T}} \sum_{t=1}^T f_{1\perp 2,t}^s \left(\mu_N u_{1,t}^{(c)} - u_{2,t}^{(c)}\right)' + E\left[\left(\bar{b}_{1,t}^{(c)} - \bar{b}_{2,t}^{(c)}\right) f_{1\perp 2,t}^{s\prime}\right] \frac{1}{T} \sum_{t=1}^T f_{1\perp 2,t}^s \left(\mu_N d_{1,t}^{(c)} - d_{2,t}^{(c)}\right)',
\end{aligned}
$$

where $f_{1\perp 2,t}^s := \left(I_{k_1-k^c} - \Phi\Phi'\right)^{-1/2} (f_{1,t}^s - \Phi f_{2,t}^s)$ is the residual of the L^2 orthogonal projection of $f_{1,t}^s$ onto $f_{2,t}^s$ normalized to have unit length. Moreover, since $[f_t^{c\,\prime}, f_{1\perp 2,t}^{s\,\prime}, f_{2,t}^{s\,\prime}]'$ is a linear one-to-one transformation of $F_t = [f_t^{c\,\prime}, f_{1,t}^{s\,\prime}, f_{2,t}^{s\,\prime}]'$ with unit identity variance-covariance matrix, we have that:

$$
\begin{aligned}
& E\left[(\bar{b}_{1,t}^{(c)} - \bar{b}_{2,t}^{(c)}) f_t^{c\prime}\right] \frac{1}{\sqrt{T}} \sum_{t=1}^T f_t^c (\mu_N u_{1,t}^{(c)} - u_{2,t}^{(c)})' + E\left[(\bar{b}_{1,t}^{(c)} - \bar{b}_{2,t}^{(c)}) f_{2,t}^{s\prime}\right] \frac{1}{\sqrt{T}} \sum_{t=1}^T f_{2,t}^s (\mu_N u_{1,t}^{(c)} - u_{2,t}^{(c)})' \\
& + E\left[\left(\bar{b}_{1,t}^{(c)} - \bar{b}_{2,t}^{(c)}\right) f_{1\perp 2,t}^{s\prime}\right] \frac{1}{\sqrt{T}} \sum_{t=1}^T f_{1\perp 2,t}^s \left(\mu_N u_{1,t}^{(c)} - u_{2,t}^{(c)}\right)' \\
= \; & E\left[\left(\bar{b}_{1,t}^{(c)} - \bar{b}_{2,t}^{(c)}\right) F_t'\right] \Sigma_F^{-1} \frac{1}{\sqrt{T}} \sum_{t=1}^T F_t \left(\mu_N u_{1,t}^{(c)} - u_{2,t}^{(c)}\right)',
\end{aligned}
$$

where matrix Σ_F is defined in Assumption A.2. The vector $E\left[\left(\bar{b}_{1,t}^{(c)} - \bar{b}_{2,t}^{(c)}\right) F_t'\right] \Sigma_F^{-1} F_t$ is the L^2 orthogonal projection of $\left(\bar{b}_{1,t}^{(c)} - \bar{b}_{2,t}^{(c)}\right)$ onto F_t. Similarly:

$$
\begin{aligned}
& E\left[(\bar{b}_{1,t}^{(c)} - \bar{b}_{2,t}^{(c)}) f_t^{c\prime}\right] \frac{1}{T} \sum_{t=1}^T f_t^c (\mu_N d_{1,t}^{(c)} - d_{2,t}^{(c)})' + E\left[(\bar{b}_{1,t}^{(c)} - \bar{b}_{2,t}^{(c)}) f_{2,t}^{s\prime}\right] \frac{1}{T} \sum_{t=1}^T f_{2,t}^s (\mu_N d_{1,t}^{(c)} - d_{2,t}^{(c)})' \\
& + E\left[\left(\bar{b}_{1,t}^{(c)} - \bar{b}_{2,t}^{(c)}\right) f_{1\perp 2,t}^{s\prime}\right] \frac{1}{T} \sum_{t=1}^T f_{1\perp 2,t}^s \left(\mu_N d_{1,t}^{(c)} - d_{2,t}^{(c)}\right)' \\
= \; & E\left[\left(\bar{b}_{1,t}^{(c)} - \bar{b}_{2,t}^{(c)}\right) F_t'\right] \Sigma_F^{-1} \frac{1}{T} \sum_{t=1}^T F_t \left(\mu_N d_{1,t}^{(c)} - d_{2,t}^{(c)}\right)'.
\end{aligned}
$$

Then, we get:

$$
\begin{aligned}
& \hat{\Psi}^{*\,(II)}_{cc} + \hat{\Psi}^{*(I)}_{cs}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\hat{\Psi}^{(I)}_{sc} \\
= \;& \frac{1}{T^2}\left\{ \left(-\tilde{\Xi}_{11}\tilde{V}_{11}^{-1}\left[-\tilde{\Xi}_{11}\tilde{R} + \tilde{\Xi}_{12}\tilde{B} - \tilde{B}'\tilde{\Xi}_{22}\tilde{B} + \tilde{B}'\tilde{\Xi}_{21}\right] + \left(\tilde{\Xi}_{22}\tilde{B} - \tilde{\Xi}_{21}\right)'\tilde{V}_{22}^{-1}\left(\tilde{\Xi}_{22}\tilde{B} - \tilde{\Xi}_{21}\right)\right)_{cc} \right. \\
& \qquad + \left[-\tilde{\Xi}_{11}\tilde{R} + \tilde{\Xi}_{12}\tilde{B} - \tilde{B}'\tilde{\Xi}_{22}\tilde{B} + \tilde{B}'\tilde{\Xi}_{21}\right]_{cs}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}(\tilde{V}_{11}^{-1})_{ss} \times \\
& \qquad \left. \left[-\tilde{\Xi}_{11}\tilde{R} + \tilde{\Xi}_{12}\tilde{B} - \tilde{B}'\tilde{\Xi}_{22}\tilde{B} + \tilde{B}'\tilde{\Xi}_{21}\right]_{sc} \right\} \\
& + \frac{1}{T\sqrt{NT}}\left\{ \left(E\left[\left(\bar{b}^{(c)}_{1,t} - \bar{b}^{(c)}_{2,t}\right)F_t'\right]\Sigma_F^{-1}\frac{1}{\sqrt{T}}\sum_{t=1}^{T} F_t\left(\mu_N u^{(c)}_{1,t} - u^{(c)}_{2,t}\right)' \right. \right. \\
& \qquad \left. \left. + \; E\left[\left(\bar{b}^{(c)}_{1,t} - \bar{b}^{(c)}_{2,t}\right)F_t'\right]\Sigma_F^{-1}\frac{1}{T}\sum_{t=1}^{T} F_t\left(\mu_N d^{(c)}_{1,t} - d^{(c)}_{2,t}\right)' \right)^{+} \right\} + o_p(\epsilon_{N,T}).
\end{aligned}
\tag{C.83}
$$

Let us now rework the term at order T^{-2}. For this purpose we use the equations:

$$
\begin{aligned}
\left[-\tilde{\Xi}_{11}\tilde{R} + \tilde{\Xi}_{12}\tilde{B} - \tilde{B}'\tilde{\Xi}_{22}\tilde{B} + \tilde{B}'\tilde{\Xi}_{21}\right]_{cc} &= -\tilde{\Xi}_{11,cc} + \tilde{\Xi}_{12,cc} - \tilde{\Xi}_{22,cc} + \tilde{\Xi}_{21,cc} = 0, \\
\left[-\tilde{\Xi}_{11}\tilde{R} + \tilde{\Xi}_{12}\tilde{B} - \tilde{B}'\tilde{\Xi}_{22}\tilde{B} + \tilde{B}'\tilde{\Xi}_{21}\right]_{sc} &= -\tilde{\Xi}_{11,sc} + \tilde{\Xi}_{12,sc} - \tilde{B}'_{cs}\tilde{\Xi}_{22,cc} \\
& \quad -\tilde{B}'_{ss}\tilde{\Xi}_{22,sc} + \tilde{B}'_{cs}\tilde{\Xi}_{21,cc} + \tilde{B}'_{ss}\tilde{\Xi}_{21,sc}, \\
\left[-\tilde{\Xi}_{11}\tilde{R} + \tilde{\Xi}_{12}\tilde{B} - \tilde{B}'\tilde{\Xi}_{22}\tilde{B} + \tilde{B}'\tilde{\Xi}_{21}\right]_{cs} &= -\tilde{\Xi}_{11,cc}\tilde{R}_{cs} - \tilde{\Xi}_{11,cs}\tilde{R}_{ss} + \tilde{\Xi}_{12,cc}\tilde{B}_{cs} + \tilde{\Xi}_{12,cs}\tilde{B}_{ss} \\
& \quad -\tilde{\Xi}_{22,cc}\tilde{B}_{cs} - \tilde{\Xi}_{22,cs}\tilde{B}_{ss} + \tilde{\Xi}_{21,cs}.
\end{aligned}
$$

Then, a block product computation yields:

$$
\begin{aligned}
& \left(-\tilde{\Xi}_{11}\tilde{V}_{11}^{-1}\left[-\tilde{\Xi}_{11}\tilde{R} + \tilde{\Xi}_{12}\tilde{B} - \tilde{B}'\tilde{\Xi}_{22}\tilde{B} + \tilde{B}'\tilde{\Xi}_{21}\right]\right)_{cc} \\
& \qquad + \left[-\tilde{\Xi}_{11}\tilde{R} + \tilde{\Xi}_{12}\tilde{B} - \tilde{B}'\tilde{\Xi}_{22}\tilde{B} + \tilde{B}'\tilde{\Xi}_{21}\right]_{cs}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}(\tilde{V}_{11}^{-1})_{ss} \times \\
& \qquad \left[-\tilde{\Xi}_{11}\tilde{R} + \tilde{\Xi}_{12}\tilde{B} - \tilde{B}'\tilde{\Xi}_{22}\tilde{B} + \tilde{B}'\tilde{\Xi}_{21}\right]_{sc} \\
= \;& -\left[\tilde{\Xi}_{11,cc}(\tilde{V}_{11}^{-1})_{cs} + \tilde{\Xi}_{11,cs}(\tilde{V}_{11}^{-1})_{ss}\right] \times \\
& \qquad \left[-\tilde{\Xi}_{11,sc} + \tilde{\Xi}_{12,sc} - \tilde{B}'_{cs}\tilde{\Xi}_{22,cc} - \tilde{B}'_{ss}\tilde{\Xi}_{22,sc} + \tilde{B}'_{cs}\tilde{\Xi}_{21,cc} + \tilde{B}'_{ss}\tilde{\Xi}_{21,sc}\right] \\
& + \left[-\tilde{\Xi}_{11,cc}\tilde{R}_{cs} - \tilde{\Xi}_{11,cs}\tilde{R}_{ss} + \tilde{\Xi}_{12,cc}\tilde{B}_{cs} + \tilde{\Xi}_{12,cs}\tilde{B}_{ss} - \tilde{\Xi}_{22,cc}\tilde{B}_{cs} - \tilde{\Xi}_{22,cs}\tilde{B}_{ss} + \tilde{\Xi}_{21,cs}\right] \\
& \times (I_{k_1-k^c} - \tilde{R}_{ss})^{-1}(\tilde{V}_{11}^{-1})_{ss}\left[-\tilde{\Xi}_{11,sc} + \tilde{\Xi}_{12,sc} - \tilde{B}'_{cs}\tilde{\Xi}_{22,cc} - \tilde{B}'_{ss}\tilde{\Xi}_{22,sc} + \tilde{B}'_{cs}\tilde{\Xi}_{21,cc} + \tilde{B}'_{ss}\tilde{\Xi}_{21,sc}\right] \\
= \;& \left[-\tilde{\Xi}_{11,cc}\left((\tilde{V}_{11}^{-1})_{cs}(\tilde{V}_{11}^{-1})_{ss}^{-1}(I_{k_1-k^c} - \tilde{R}_{ss}) + \tilde{R}_{cs}\right) \right. \\
& \qquad \left. -\tilde{\Xi}_{11,cs} + \tilde{\Xi}_{12,cc}\tilde{B}_{cs} + \tilde{\Xi}_{12,cs}\tilde{B}_{ss} - \tilde{\Xi}_{22,cc}\tilde{B}_{cs} - \tilde{\Xi}_{22,cs}\tilde{B}_{ss} + \tilde{\Xi}_{21,cs}\right] \\
& \times (I_{k_1-k^c} - \tilde{R}_{ss})^{-1}(\tilde{V}_{11}^{-1})_{ss}\left[-\tilde{\Xi}_{11,sc} + \tilde{\Xi}_{12,sc} - \tilde{B}'_{cs}\tilde{\Xi}_{22,cc} - \tilde{B}'_{ss}\tilde{\Xi}_{22,sc} + \tilde{B}'_{cs}\tilde{\Xi}_{21,cc} + \tilde{B}'_{ss}\tilde{\Xi}_{21,sc}\right].
\end{aligned}
$$

Let us show that the term $(\tilde{V}_{11}^{-1})_{cs}(\tilde{V}_{11}^{-1})_{ss}^{-1}(I_{k_1-k^c} - \tilde{R}_{ss}) + \tilde{R}_{cs}$ vanishes. Indeed, from equation (C.67) we have:

$$\begin{aligned}
(\tilde{V}_{11}^{-1})_{cs}(\tilde{V}_{11}^{-1})_{ss}^{-1}(I_{k_1-k^c}-\tilde{R}_{ss})+\tilde{R}_{cs} &= \left[(\tilde{V}_{11}^{-1})_{cs}(\tilde{V}_{11}^{-1})_{ss}^{-1}+\tilde{R}_{cs}(I_{k_1-k^c}-\tilde{R}_{ss})^{-1}\right](I_{k_1-k^c}-\tilde{R}_{ss}) \\
&= \left[(\tilde{V}_{11}^{-1})_{cs}(\tilde{V}_{11}^{-1})_{ss}^{-1}+\tilde{\Sigma}_{cc}^{-1}\tilde{\Sigma}_{c,1}\right](I_{k_1-k^c}-\tilde{R}_{ss}) \\
&= \tilde{\Sigma}_{cc}^{-1}\left[\tilde{\Sigma}_{cc}(\tilde{V}_{11}^{-1})_{cs}+\tilde{\Sigma}_{c,1}(\tilde{V}_{11}^{-1})_{ss}\right](\tilde{V}_{11}^{-1})_{ss}^{-1}(I_{k_1-k^c}-\tilde{R}_{ss}) \\
&= \tilde{\Sigma}_{cc}^{-1}\left[(\tilde{V}_{11})_{cc}(\tilde{V}_{11}^{-1})_{cs}+(\tilde{V}_{11})_{cs}(\tilde{V}_{11}^{-1})_{ss}\right](\tilde{V}_{11}^{-1})_{ss}^{-1}(I_{k_1-k^c}-\tilde{R}_{ss}) \\
&= 0.
\end{aligned}$$

Therefore, we get:

$$\begin{aligned}
&\left(-\tilde{\Xi}_{11}\tilde{V}_{11}^{-1}\left[-\tilde{\Xi}_{11}\tilde{R}+\tilde{\Xi}_{12}\tilde{B}-\tilde{B}'\tilde{\Xi}_{22}\tilde{B}+\tilde{B}'\tilde{\Xi}_{21}\right]\right)_{cc} \\
&\qquad +\left[-\tilde{\Xi}_{11}\tilde{R}+\tilde{\Xi}_{12}\tilde{B}-\tilde{B}'\tilde{\Xi}_{22}\tilde{B}+\tilde{B}'\tilde{\Xi}_{21}\right]_{cs}(I_{k_1-k^c}-\tilde{R}_{ss})^{-1}(\tilde{V}_{11}^{-1})_{ss}\times \\
&\qquad \left[-\tilde{\Xi}_{11}\tilde{R}+\tilde{\Xi}_{12}\tilde{B}-\tilde{B}'\tilde{\Xi}_{22}\tilde{B}+\tilde{B}'\tilde{\Xi}_{21}\right]_{sc} \\
=\;& \left[-\tilde{\Xi}_{11,cs}+\tilde{\Xi}_{12,cc}\tilde{B}_{cs}+\tilde{\Xi}_{12,cs}\tilde{B}_{ss}-\tilde{\Xi}_{22,cc}\tilde{B}_{cs}-\tilde{\Xi}_{22,cs}\tilde{B}_{ss}+\tilde{\Xi}_{21,cs}\right] \\
&\times(I_{k_1-k^c}-\tilde{R}_{ss})^{-1}(\tilde{V}_{11}^{-1})_{ss}\left[-\tilde{\Xi}_{11,sc}+\tilde{\Xi}_{12,sc}-\tilde{B}_{cs}'\tilde{\Xi}_{22,cc}-\tilde{B}_{ss}'\tilde{\Xi}_{22,sc}+\tilde{B}_{cs}'\tilde{\Xi}_{21,cc}+\tilde{B}_{ss}'\tilde{\Xi}_{21,sc}\right] \\
=\;& \left[-\tilde{\Xi}_{11,sc}+\tilde{\Xi}_{12,sc}-\tilde{B}_{cs}'\tilde{\Xi}_{22,cc}-\tilde{B}_{ss}'\tilde{\Xi}_{22,sc}+\tilde{B}_{cs}'\tilde{\Xi}_{21,cc}+\tilde{B}_{ss}'\tilde{\Xi}_{21,sc}\right]' \\
&\times(I_{k_1-k^c}-\tilde{R}_{ss})^{-1}(\tilde{V}_{11}^{-1})_{ss}\left[-\tilde{\Xi}_{11,sc}+\tilde{\Xi}_{12,sc}-\tilde{B}_{cs}'\tilde{\Xi}_{22,cc}-\tilde{B}_{ss}'\tilde{\Xi}_{22,sc}+\tilde{B}_{cs}'\tilde{\Xi}_{21,cc}+\tilde{B}_{ss}'\tilde{\Xi}_{21,sc}\right].
\end{aligned}$$

Let us consider the term $-\tilde{\Xi}_{11,sc}$ + $\tilde{\Xi}_{12,sc}$ - $\tilde{B}_{cs}'\tilde{\Xi}_{22,cc}$ - $\tilde{B}_{ss}'\tilde{\Xi}_{22,sc}$ + $\tilde{B}_{cs}'\tilde{\Xi}_{21,cc}$ + $\tilde{B}_{ss}'\tilde{\Xi}_{21,sc}$ = $-\left[(\tilde{\Xi}_{11,sc}-\tilde{\Xi}_{12,sc})-\tilde{B}_{cs}'(\tilde{\Xi}_{21,cc}-\tilde{\Xi}_{22,cc})-\tilde{B}_{ss}'(\tilde{\Xi}_{21,sc}-\tilde{\Xi}_{22,sc})\right]$. Using $\tilde{\Xi}_{11,sc}-\tilde{\Xi}_{12,sc}=\frac{1}{T}\sum_t f_{1,t}^s\left(b_{1,t}^{(c)}-b_{2,t}^{(c)}\right)'$, $\tilde{\Xi}_{21,cc}-\tilde{\Xi}_{22,cc}=\frac{1}{T}\sum_t f_t^c\left(b_{1,t}^{(c)}-b_{2,t}^{(c)}\right)'$ and $\tilde{\Xi}_{21,sc}-\tilde{\Xi}_{22,sc}=\frac{1}{T}\sum_t f_{2,t}^s\left(b_{1,t}^{(c)}-b_{2,t}^{(c)}\right)'$, we can write it as:

$$\begin{aligned}
&-\tilde{\Xi}_{11,sc}+\tilde{\Xi}_{12,sc}-\tilde{B}_{cs}'\tilde{\Xi}_{22,cc}-\tilde{B}_{ss}'\tilde{\Xi}_{22,sc}+\tilde{B}_{cs}'\tilde{\Xi}_{21,cc}+\tilde{B}_{ss}'\tilde{\Xi}_{21,sc} \\
&\qquad =-\frac{1}{T}\sum_{t=1}^{T}\left[f_{1,t}^s-\tilde{B}_{cs}'f_t^c-\tilde{B}_{ss}'f_{2,t}^s\right]\left(b_{1,t}^{(c)}-b_{2,t}^{(c)}\right)'.
\end{aligned}$$

Noting that

$$\tilde{B}'=\tilde{V}_{12}\tilde{V}_{22}^{-1}=\left[\begin{array}{cc} I & 0 \\ \tilde{B}_{cs}' & \tilde{B}_{ss}' \end{array}\right],$$

we deduce that:

$$\tilde{f}_{1\perp 2c,t}=f_{1,t}^s-\tilde{B}_{cs}'f_t^c-\tilde{B}_{ss}'f_{2,t}^s, \qquad t=1,...,T,$$

are the residuals in the sample orthogonal projection of $f_{1,t}^s$ on $f_{2,t}^s$ and f_t^c. Let us now show that $(I_{k_1-k^c}-$

$\tilde{R}_{ss})^{-1}(\tilde{V}_{11}^{-1})_{ss}$ is the inverse of the sample variance of that residuals. Indeed, the sample variance is:

$$
\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T}\tilde{f}_{1\perp 2c,t}\tilde{f}'_{1\perp 2c,t} &= \frac{1}{T}\sum_{t=1}^{T}\left[f_{1,t}^{s}-\tilde{B}'_{cs}f_t^{c}-\tilde{B}'_{ss}f_{2,t}\right]f_{1,t}^{s\prime} \\
&= \tilde{\Sigma}_{11}-\tilde{B}'_{cs}\tilde{\Sigma}_{c,1}-\tilde{B}'_{ss}\tilde{\Sigma}_{2,1}=\left(\tilde{V}_{11}-\tilde{B}'\tilde{V}_{21}\right)_{ss} \\
&= \left(\tilde{V}_{11}(I_{k_1}-\tilde{R})\right)_{ss} \\
&= -\tilde{\Sigma}_{1c}\tilde{R}_{cs}+\tilde{\Sigma}_{11}(I_{k_1-k^c}-R_{ss}) \\
&= \left[-\tilde{\Sigma}_{1c}\tilde{R}_{cs}(I_{k_1-k^c}-R_{ss})^{-1}+\tilde{\Sigma}_{11}\right](I_{k_1-k^c}-R_{ss}) \\
&= \left(-\tilde{\Sigma}_{1c}\tilde{\Sigma}_{cc}^{-1}\tilde{\Sigma}_{c1}+\tilde{\Sigma}_{11}\right)(I_{k_1-k^c}-R_{ss})=[(\tilde{V}_{11}^{-1})_{ss}]^{-1}(I_{k_1-k^c}-R_{ss}),
\end{aligned}
$$

from Equation (C.67). By gathering these results, we get:

$$
\begin{aligned}
&\left(-\tilde{\Xi}_{11}\tilde{V}_{11}^{-1}\left[-\tilde{\Xi}_{11}\tilde{R}+\tilde{\Xi}_{12}\tilde{B}-\tilde{B}'\tilde{\Xi}_{22}\tilde{B}+\tilde{B}'\tilde{\Xi}_{21}\right]\right)_{cc} \\
&\qquad+\left[-\tilde{\Xi}_{11}\tilde{R}+\tilde{\Xi}_{12}\tilde{B}-\tilde{B}'\tilde{\Xi}_{22}\tilde{B}+\tilde{B}'\tilde{\Xi}_{21}\right]_{cs}(I_{k_1-k^c}-\tilde{R}_{ss})^{-1}(\tilde{V}_{11}^{-1})_{ss}\times \\
&\qquad\quad\left[-\tilde{\Xi}_{11}\tilde{R}+\tilde{\Xi}_{12}\tilde{B}-\tilde{B}'\tilde{\Xi}_{22}\tilde{B}+\tilde{B}'\tilde{\Xi}_{21}\right]_{sc} \\
=\;&\left[\frac{1}{T}\sum_{t=1}^{T}\left(b_{1,t}^{(c)}-b_{2,t}^{(c)}\right)\tilde{f}'_{1\perp 2c,t}\right]\left(\frac{1}{T}\sum_{t=1}^{T}\tilde{f}_{1\perp 2c,t}\tilde{f}'_{1\perp 2c,t}\right)^{-1}\left[\frac{1}{T}\sum_{t=1}^{T}\tilde{f}_{1\perp 2c,t}\left(b_{1,t}^{(c)}-b_{2,t}^{(c)}\right)'\right].
\end{aligned}
$$

Let us now consider the term $\left[\left(\tilde{\Xi}_{22}\tilde{B}-\tilde{\Xi}_{21}\right)'\tilde{V}_{22}^{-1}\left(\tilde{\Xi}_{22}\tilde{B}-\tilde{\Xi}_{21}\right)\right]_{cc}$ also showing at order T^{-2} in the r.h.s. of the asymptotic expansion (C.83). Direct computation yields:

$$
\begin{aligned}
&\left[\left(\tilde{\Xi}_{22}\tilde{B}-\tilde{\Xi}_{21}\right)'\tilde{V}_{22}^{-1}\left(\tilde{\Xi}_{22}\tilde{B}-\tilde{\Xi}_{21}\right)\right]_{cc} \\
&\qquad\qquad=\left[(\tilde{\Xi}_{22,cc}-\tilde{\Xi}_{21,cc})'\vdots(\tilde{\Xi}_{22,sc}-\tilde{\Xi}_{21,sc})'\right]\tilde{V}_{22}^{-1}\begin{bmatrix}\tilde{\Xi}_{22,cc}-\tilde{\Xi}_{21,cc}\\ \tilde{\Xi}_{22,sc}-\tilde{\Xi}_{21,sc}\end{bmatrix} \\
&\quad=\left[\frac{1}{T}\sum_{t=1}^{T}\left(b_{1,t}^{(c)}-b_{2,t}^{(c)}\right)h'_{2,t}\right]\left(\frac{1}{T}\sum_{t=1}^{T}h_{2,t}h'_{2,t}\right)^{-1}\left[\frac{1}{T}\sum_{t=1}^{T}h_{2,t}\left(b_{1,t}^{(c)}-b_{2,t}^{(c)}\right)'\right].
\end{aligned}
$$

Hence, the term at order T^{-2} in the r.h.s. of (C.83) becomes:

$$\begin{aligned}
& \left(-\tilde{\Xi}_{11}\tilde{V}_{11}^{-1}\left[-\tilde{\Xi}_{11}\tilde{R}+\tilde{\Xi}_{12}\tilde{B}-\tilde{B}'\tilde{\Xi}_{22}\tilde{B}+\tilde{B}'\tilde{\Xi}_{21}\right]+\left(\tilde{\Xi}_{22}\tilde{B}-\tilde{\Xi}_{21}\right)'\tilde{V}_{22}^{-1}\left(\tilde{\Xi}_{22}\tilde{B}-\tilde{\Xi}_{21}\right)\right)_{cc} \\
& \quad +\left[-\tilde{\Xi}_{11}\tilde{R}+\tilde{\Xi}_{12}\tilde{B}-\tilde{B}'\tilde{\Xi}_{22}\tilde{B}+\tilde{B}'\tilde{\Xi}_{21}\right]_{cs}(I_{k_1-k^c}-\tilde{R}_{ss})^{-1}(\tilde{V}_{11}^{-1})_{ss}\times \\
& \quad \left[-\tilde{\Xi}_{11}\tilde{R}+\tilde{\Xi}_{12}\tilde{B}-\tilde{B}'\tilde{\Xi}_{22}\tilde{B}+\tilde{B}'\tilde{\Xi}_{21}\right]_{sc} \\
= & \left[\frac{1}{T}\sum_{t=1}^{T}\left(b_{1,t}^{(c)}-b_{2,t}^{(c)}\right)h_{2,t}'\right]\left(\frac{1}{T}\sum_{t=1}^{T}h_{2,t}h_{2,t}'\right)^{-1}\left[\frac{1}{T}\sum_{t=1}^{T}h_{2,t}\left(b_{1,t}^{(c)}-b_{2,t}^{(c)}\right)'\right] \\
& +\left[\frac{1}{T}\sum_{t=1}^{T}\left(b_{1,t}^{(c)}-b_{2,t}^{(c)}\right)\tilde{f}_{1\perp 2c,t}'\right]\left(\frac{1}{T}\sum_{t=1}^{T}\tilde{f}_{1\perp 2c,t}\tilde{f}_{1\perp 2c,t}'\right)^{-1}\left[\frac{1}{T}\sum_{t=1}^{T}\tilde{f}_{1\perp 2,c}\left(b_{1,t}^{(c)}-b_{2,t}^{(c)}\right)'\right] \\
= & \left[\frac{1}{T}\sum_{t=1}^{T}\left(b_{1,t}^{(c)}-b_{2,t}^{(c)}\right)F_t'\right]\tilde{\Sigma}_F^{-1}\left[\frac{1}{T}\sum_{t=1}^{T}F_t\left(b_{1,t}^{(c)}-b_{2,t}^{(c)}\right)'\right],
\end{aligned} \tag{C.84}$$

where:

$$\tilde{\Sigma}_F=\frac{1}{T}\sum_{t=1}^{T}F_tF_t',$$

because $\tilde{f}_{1\perp 2c,t}$ is orthogonal in-sample to $h_{2,t}$, and $(\tilde{f}_{1\perp 2,c}', h_{2,t}')'$ is a linear transformation of $(f_t^{c\prime}, f_{1,t}^{s\prime}, f_{2,t}^{s\prime})'$. By substituting (C.84) into (C.83), we get:

$$\begin{aligned}
& \hat{\Psi}_{cc}^{*\,(II)}+\hat{\Psi}_{cs}^{*(I)}(I_{k_1-k^c}-\tilde{R}_{ss})^{-1}\hat{\Psi}_{sc}^{(I)} \\
= & \frac{1}{T^2}\left\{\left[\frac{1}{T}\sum_{t=1}^{T}\left(b_{1,t}^{(c)}-b_{2,t}^{(c)}\right)F_t'\right]\tilde{\Sigma}_F^{-1}\left[\frac{1}{T}\sum_{t=1}^{T}F_t\left(b_{1,t}^{(c)}-b_{2,t}^{(c)}\right)'\right]\right\} \\
& +\frac{1}{T\sqrt{NT}}\left\{\left(E\left[\left(\bar{b}_{1,t}^{(c)}-\bar{b}_{2,t}^{(c)}\right)F_t'\right]\Sigma_F^{-1}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}F_t\left(\mu_N u_{1,t}^{(c)}-u_{2,t}^{(c)}\right)'\right.\right. \\
& \quad \left.\left.+\ E\left[\left(\bar{b}_{1,t}^{(c)}-\bar{b}_{2,t}^{(c)}\right)F_t'\right]\Sigma_F^{-1}\frac{1}{T}\sum_{t=1}^{T}F_t\left(\mu_N d_{1,t}^{(c)}-d_{2,t}^{(c)}\right)'\right)^{+}\right\}+o_p(\epsilon_{N,T}).
\end{aligned} \tag{C.85}$$

iv) Conclusion

We finally consider the other second-order terms in the r.h.s. of (B.13).
By $\hat{\Psi}_{cc}^{*(I)} = O_p\left(\frac{1}{N}+\frac{1}{T^2}+\frac{1}{T\sqrt{NT}}+\epsilon_{N,T}\right)$ from Lemma C.7, we have:

$$\hat{\Psi}_{cc}^{*(I)}\tilde{\Sigma}_{cc}^{-1}\hat{\Psi}_{cc}^{*(I)}=o_p(\epsilon_{N,T}), \tag{C.86}$$

if $T^{1/2}\ll N\ll T^3$. Moreover, by using $\tilde{\Sigma}_{c,1}(I_{k_1-k^c}-\tilde{R}_{ss})^{-1}=O_p(T^{-1/2})$ from (C.74), $\hat{\Psi}_{sc}^{(I)}=O_p(\delta_{N,T})$, and $\hat{\Psi}_{cc}^{*(I)}=O_p\left(\frac{1}{N}+\frac{1}{T^2}+\frac{1}{T\sqrt{NT}}+\epsilon_{N,T}\right)$, we have:

$$\begin{aligned}
\tilde{\Sigma}_{c,1}(I_{k_1-k^c}-\tilde{R}_{ss})^{-1}\hat{\Psi}_{sc}^{(I)}\tilde{\Sigma}_{cc}^{-1}\hat{\Psi}_{cc}^{*(I)} &= O_p\left[\frac{1}{\sqrt{T}}\delta_{N,T}\left(\frac{1}{N}+\frac{1}{T^2}+\frac{1}{T\sqrt{NT}}+\epsilon_{N,T}\right)\right] \\
&= o_p\left(\epsilon_{N,T}\right),
\end{aligned} \tag{C.87}$$

if $T^{1/2} \ll N \ll T^3$. From (C.85), (C.86) and (C.87), the conclusion follows. ∎

C.9 Proof of Lemma B.6

We show the conditions in parts (i)-(iv) of Lemma B.6. Part (i) follows by the Law of Iterated Expectation and $E(U_t|\mathcal{F}_t) = 0$, which is implied by Assumption A.4 *a)*. Part (ii) is implied by Assumptions A.3, A.4 *b)* and A.5 *b)*. The NED property in part (iii) holds true because conditional expectations given $\mathcal{F}_t$ can be well approximated by elements in the sigma-field $\mathcal{V}_{t-m}^{t+m}$ generated by the mixing process (V_t), for large m, by Assumptions A.3, A.4 *b)*, A.5 *b)* and A.6 *a)-c)*, as we show in the next lemma.

LEMMA C.9. *Assumptions A.3, A.4 b), A.5 b) and A.6 a)-c) imply part (iii) in Lemma B.6.*

To check part (iv) in Lemma B.6 we use:

$$\begin{aligned}\lim_{T,N\to\infty} V\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\mathcal{Z}_{N,t}\right) &= \lim_{T,N\to\infty}\frac{1}{T}\sum_{h=-T+1}^{T-1}(T-|h|)Cov\left(\mathcal{Z}_{N,t},\mathcal{Z}_{N,t-h}\right)\\ &= \lim_{N\to\infty}\sum_{h=-\infty}^{\infty}Cov\left(\mathcal{Z}_{N,t},\mathcal{Z}_{N,t-h}\right),\end{aligned}$$

where the first equality follows from stationarity of the data. The series converges because the zero-mean process $\mathcal{Z}_{N,t}$ is a L^2-mixingale with size -1, [1] by Theorem 17.5 in Davidson (1994) and Conditions (ii)-(iii), which implies $\|Cov\left(\mathcal{Z}_{N,t},\mathcal{Z}_{N,t-h}\right)\| = \left\|E\left[E(\mathcal{Z}_{N,t}|\mathcal{V}_{t-h})\mathcal{Z}_{N,t-h}'\right]\right\| \leq \|E(\mathcal{Z}_{N,t}|\mathcal{V}_{t-h})\|_2\|\mathcal{Z}_{N,t-h}\|_2 = O\left(h^{-\psi}\right)$, uniformly in $N_1, N_2 \geq 1$, for some $\psi > 1$. The latter uniform bound also allows for an application of the Lebesgue Lemma to get:

$$\Omega_U = \lim_{T,N\to\infty} V\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\mathcal{Z}_{N,t}\right) = \sum_{h=-\infty}^{\infty}\Gamma(h),$$

where $\Gamma(h) = \lim_{N\to\infty} Cov\left(\mathcal{Z}_{N,t},\mathcal{Z}_{N,t-h}\right)$, which yields equation (B.16). The computations in Subsection B.1.6, and in particular Lemma B.7, show that the limit in $\Gamma(h)$ is well-defined.

C.9.1 Proof of Lemma C.9

Assumption A.6 *a)* gives the strong mixing condition for process V_t. Since $U_t = \mu_N\left(\tilde{\Sigma}_{\Lambda,1}^{-1}\xi_{1,t}\right)^{(c)} - \left(\tilde{\Sigma}_{\Lambda,2}^{-1}\xi_{2,t}\right)^{(c)}$, where $\tilde{\Sigma}_{\Lambda,j} = \Lambda_j'\Lambda_j/N_j$ for $j = 1,2$, process U_t is function of the components of process V_t. Therefore, to prove the NED property for process $\mathcal{Z}_{N,t}$, we simply have to show that processes $X_{N,t} = E(U_t'U_t|\mathcal{F}_t)$ and $Y_{N,t} = \Delta b_t^{(c)\prime}U_t$ are L^2-NED on (V_t). For the first process we have:

$$\begin{aligned}\|X_{N,t} - E(X_{N,t}|\mathcal{V}_{t-m}^{t+m})\|_2 &\leq \|X_{N,t} - E(X_{N,t}|F_t,...,F_{t-m})\|_2\\ &= \|E(U_t'U_t|\mathcal{F}_t) - E(U_t'U_t|F_t,...,F_{t-m})\|_2 = O(m^{-\psi}),\end{aligned}$$

for $\psi > 1$, by the Law of Iterated Expectation and Assumption A.6 *b)*. For the second process we have:

$$\begin{aligned}\|Y_{N,t} - E(Y_{N,t}|\mathcal{V}_{t-m}^{t+m})\|_2 &\leq \|U_t'(\Delta b_t^{(c)} - E[\Delta b_t^{(c)}|\mathcal{V}_{t-m}^{t+m}])\|_2 \leq E\left[\|U_t\|^2\|\Delta b_t^{(c)} - E(\Delta b_t^{(c)}|\mathcal{V}_{t-m}^{t+m})\|^2\right]^{1/2}\\ &\leq \|U_t\|_{2r}\|\Delta b_t^{(c)} - E(\Delta b_t^{(c)}|\mathcal{V}_{t-m}^{t+m})\|_{2p},\end{aligned}$$

[1] That is, $\|E[\mathcal{Z}_{N,t}|\mathcal{V}_{t-m}]\|_2 \leq \zeta(m)$, uniformly in $t \geq 1$ and $N_1, N_2 \geq 1$, where $\zeta(m) = O(m^{-\psi})$ for some $\psi > 1$.

where $p = r/(r-1)$, by the Holder inequality. Term $\|U_t\|_{2r}$ is bounded uniformly in $N_1, N_2 \geq 1$ and $t \geq 1$ by Assumptions A.3 and A.5 *b)*. Moreover, by the definition of process Δb_t we get:

$$\begin{aligned}
\|\Delta b_t^{(c)} - E(\Delta b_t^{(c)}|\mathcal{V}_{t-m}^{t+m})\|_{2p} &\leq \sum_{j=1}^{2} \|\bar{b}_{j,t}^{(c)} - E(\bar{b}_{j,t}^{(c)}|\mathcal{V}_{t-m}^{t+m})\|_{2p} \leq \sum_{j=1}^{2} \|\Sigma_{\Lambda,j}^{-1}\| \|h_{j,t}(\eta_{j,t}^2 - E(\eta_{j,t}^2|\mathcal{V}_{t-m}^{t+m}))\|_{2p} \\
&\leq \sum_{j=1}^{2} \|\Sigma_{\Lambda,j}^{-1}\| \|h_{j,t}\|_{4p} \|\eta_{j,t}^2 - E(\eta_{j,t}^2|\mathcal{V}_{t-m}^{t+m})\|_{4p}.
\end{aligned}$$

The latter term is $O(m^{-\psi})$ with $\psi > 1$ by Assumptions A.3, A.4 *b)* A.6 *c)* and $4p \leq 8$. The conclusion follows. ■

C.10 Proof of Lemma B.7

First, let us show that we can interchange the limit $N \to \infty$ and the outer expectation in the r.h.s. of equation (B.18), i.e.:

$$\lim_{N\to\infty} E\left[Cov\left(\mathcal{Z}_{N,t}, \mathcal{Z}_{N,t-h}|\mathcal{F}_t\right)\right] = E\left[\lim_{N\to\infty} Cov\left(\mathcal{Z}_{N,t}, \mathcal{Z}_{N,t-h}|\mathcal{F}_t\right)\right]. \tag{C.88}$$

Indeed, by the Cauchy-Schwarz inequality, we have the bound $Cov\left(\mathcal{Z}_{N,t}, \mathcal{Z}_{N,t-h}|\mathcal{F}_t\right) \leq \chi_t \chi_{t-h}$, P-a.s., uniformly in $N_1, N_2 \geq 1$, where $\chi_t := \sup_{N_1,N_2\geq 1} E\left[\|\mathcal{Z}_{N,t}\|^2|\mathcal{F}_t\right]^{1/2}$. The uniform upper bound $\chi_t\chi_{t-h}$ is integrable, because $E[\chi_t\chi_{t-h}] \leq E[\chi_t^2]^{1/2} E[\chi_{t-h}^2]^{1/2}$ by Cauchy-Schwarz, and $E[\chi_t^2] = E\left[\sup_{N_1,N_2\geq 1} E\left(\|\mathcal{Z}_{N,t}\|^2|\mathcal{F}_t\right)\right]$ $\leq CE\left[\sup_{N_1,N_2\geq 1} E\left(\|U_t\|^4|\mathcal{F}_t\right)\right] < \infty$, for a constant C, by Assumption A.5 *b)*. Then, (C.88) follows by an application of the Lebesgue Lemma.

Second, let us write $Cov\left(\mathcal{Z}_{N,t}, \mathcal{Z}_{N,t-h}|\mathcal{F}_t\right)$ in block form using $\mathcal{Z}_{N,t} = [U_t'U_t - E(U_t'U_t|\mathcal{F}_t), \Delta b_t^{(c)\prime}U_t]'$ and show that:

$$\begin{aligned}
&\lim_{N\to\infty} \begin{bmatrix} Cov(U_t'U_t, U_{t-h}'U_{t-h}|\mathcal{F}_t) & Cov(U_t'U_t, \Delta b_{t-h}^{(c)\prime}U_{t-h}|\mathcal{F}_t) \\ Cov(\Delta b_t^{(c)\prime}U_t, U_{t-h}'U_{t-h}|\mathcal{F}_t) & Cov(\Delta b_t^{(c)\prime}U_t, \Delta b_{t-h}^{(c)\prime}U_{t-h}|\mathcal{F}_t) \end{bmatrix} \\
&= \begin{bmatrix} Cov(U_t^{\infty\prime}U_t^{\infty}, U_{t-h}^{\infty\prime}U_{t-h}^{\infty}|\mathcal{F}_t) & Cov(U_t^{\infty\prime}U_t^{\infty}, \Delta b_{t-h}^{(c)\prime}U_{t-h}^{\infty}|\mathcal{F}_t) \\ Cov(\Delta b_t^{(c)\prime}U_t^{\infty}, U_{t-h}^{\infty\prime}U_{t-h}^{\infty}|\mathcal{F}_t) & Cov(\Delta b_t^{(c)\prime}U_t^{\infty}, \Delta b_{t-h}^{(c)\prime}U_{t-h}^{\infty}|\mathcal{F}_t) \end{bmatrix}, \quad P-a.s.
\end{aligned} \tag{C.89}$$

We focus on the convergence of the upper-left block; the arguments for the other blocks are similar. We have $Cov(U_t'U_t, U_{t-h}'U_{t-h}|\mathcal{F}_t) = E[(U_t'U_t)(U_{t-h}'U_{t-h})|\mathcal{F}_t] - E[U_t'U_t|\mathcal{F}_t]E[U_{t-h}'U_{t-h}|\mathcal{F}_t]$. Let us prove that:

$$\lim_{N\to\infty} E[(U_t'U_t)(U_{t-h}'U_{t-h})|\mathcal{F}_t] = E[(U_t^{\infty\prime}U_t^{\infty})(U_{t-h}^{\infty\prime}U_{t-h}^{\infty})|\mathcal{F}_t], \qquad P-a.s. \tag{C.90}$$

By definition of conditional expectation, this is equivalent to:

$$E\left[\lim_{N\to\infty} E[(U_t'U_t)(U_{t-h}'U_{t-h})|\mathcal{F}_t]1_A\right] = E\left[(U_t^{\infty\prime}U_t^{\infty})(U_{t-h}^{\infty\prime}U_{t-h}^{\infty})1_A\right],$$

for any measurable set $A \in \mathcal{F}_t$. By Assumption A.5 *b)* and the Lebesgue Lemma, we can interchange the limes and the expectation in the l.h.s., and by the Law of Iterated Expectation we get:

$$\lim_{N\to\infty} E\left[(U_t'U_t)(U_{t-h}'U_{t-h})1_A\right] = E\left[(U_t^{\infty\prime}U_t^{\infty})(U_{t-h}^{\infty\prime}U_{t-h}^{\infty})1_A\right]. \tag{C.91}$$

Now, by (B.19) and stable convergence, we have $(U_t'U_t)(U_{t-h}'U_{t-h})1_A \xrightarrow{d} (U_t^{\infty\prime}U_t^{\infty})(U_{t-h}^{\infty\prime}U_{t-h}^{\infty})1_A$. Moreover, by Assumption A.5 *b)*, we have uniform integrability: $\sup_{N\geq 1} E[|(U_t'U_t)(U_{t-h}'U_{t-h})1_A|^\rho] < \infty$, for some $\rho > 1$. Therefore, by the Corollary of Theorem 25.12 on page 338 in Billingsley (1995), we get (C.91). By similar arguments applied to $E[U_t'U_t|\mathcal{F}_t]$ and $E[U_{t-h}'U_{t-h}|\mathcal{F}_t]$, and to the other blocks of the matrix in the l.h.s. of (C.89), equation (C.89) follows. Combining (C.88) and (C.89), the statement of Lemma B.7 follows. ■

C.11 Proof of Lemma B.8

The proof of Lemma B.8 deploys the following uniform asymptotic expansions of factors and loadings estimates:

$$\hat{f}_t^c = \hat{\mathcal{H}}_c^{-1}\left[f_t^c + \frac{1}{\sqrt{N_1}}u_{1,t}^{(c)}\right] + o_p\left(T^{-1/2}\right), \tag{C.92}$$

$$\hat{f}_{j,t}^s = \hat{\mathcal{H}}_{s,j}^{-1}\left[\tilde{f}_{j,t}^s + \frac{1}{\sqrt{N_j}}u_{j,t}^{(s)}\right] + o_p(T^{-1/2}), \qquad j = 1,2, \tag{C.93}$$

$$\hat{\lambda}_{j,i}^c = \hat{\mathcal{H}}_c'\left[\lambda_{j,i}^c + \tilde{\Sigma}_{cc}^{-1}\tilde{\Sigma}_{c,j}\lambda_{j,i}^s + \frac{1}{\sqrt{T}}w_{j,i}^c\right] + o_p\left(T^{-1/2}\right), \qquad j = 1,2, \tag{C.94}$$

$$\hat{\lambda}_{j,i}^s = \hat{\mathcal{H}}_{s,j}'\left[\lambda_{j,i}^s + \frac{1}{\sqrt{T}}w_{j,i}^s\right] + o_p\left(T^{-1/2}\right), \qquad j = 1,2, \tag{C.95}$$

where the $o_p(T^{-1/2})$ terms are uniform w.r.t. $1 \leq t \leq T$ and $1 \leq i \leq N_j$, vector $u_{j,t}$ is defined in Proposition 3, $\tilde{f}_{j,t}^s = f_{j,t}^s - \tilde{\Sigma}_{j,c}\tilde{\Sigma}_{cc}^{-1}f_t^c$, $w_{j,i}^c = \tilde{\Sigma}_{cc}^{-1}\frac{1}{\sqrt{T}}\sum_{t=1}^T f_t^c\varepsilon_{j,i,t}$ and $w_{j,i}^s = (\frac{1}{T}\tilde{F}_j^{s\prime}\tilde{F}_j^s)^{-1}\frac{1}{\sqrt{T}}\sum_{t=1}^T \tilde{f}_{j,t}^s\varepsilon_{j,i,t}$, and matrices $\hat{\mathcal{H}}_c$ and $\hat{\mathcal{H}}_{s,j}$ are such that $\hat{\mathcal{H}}_c'\hat{\mathcal{H}}_c = I_{k^c} + o_p(1)$ and $\hat{\mathcal{H}}_{s,j}'\hat{\mathcal{H}}_{s,j} = I_{k_j^s} + o_p(1)$.
These asymptotic expansions hold under Assumptions A.1-A.4, A.5 *b)-c)*, A.6 *a)*, A.7, A.8, and are derived in Proposition D.4 in Appendix D.4.

C.11.1 Proof of Lemma B.8 Part (i)

To derive the asymptotic expansion of matrix $\hat{\Lambda}_j'\hat{\Lambda}_j/N_j$, we work with the matrix versions of the asymptotic expansions in equations (C.94) and (C.95). Stacking the loadings $\hat{\lambda}_{j,i}^c$ in matrix $\hat{\Lambda}_j^c = [\hat{\lambda}_{j,1}^c, ..., \hat{\lambda}_{j,N_j}^c]'$ we get:

$$\hat{\Lambda}_j^c = \left[\Lambda_j^c + \frac{1}{\sqrt{T}}(G_j^c + \Lambda_j^s\sqrt{T}\tilde{\Sigma}_{j,c}\tilde{\Sigma}_{cc}^{-1})\right]\hat{\mathcal{H}}_c + o_p\left(T^{-1/2}\right),$$

where

$$G_j^c = \frac{1}{\sqrt{T}}\varepsilon_j'F^c, \tag{C.96}$$

and $o_p(T^{-1/2})$ denotes a matrix whose rows are $(k^c, 1)$ vectors uniformly of order $o_p(T^{-1/2})$. Similarly, stacking the loadings $\hat{\lambda}_{j,i}^s$ in matrix $\hat{\Lambda}_j^s = [\hat{\lambda}_{j,1}^s, ..., \hat{\lambda}_{j,N_j}^s]'$ we get:

$$\hat{\Lambda}_j^s = \left[\Lambda_j^s + \frac{1}{\sqrt{T}}G_j^s\right]\hat{\mathcal{H}}_{s,j} + o_p\left(T^{-1/2}\right),$$

where

$$G_j^s = \frac{1}{\sqrt{T}}\varepsilon_j'F_j^s. \tag{C.97}$$

By gathering these expansions into matrix $\hat{\Lambda}_j = [\hat{\Lambda}_j^c \ \vdots \ \hat{\Lambda}_j^s]$, we get:

$$\hat{\Lambda}_j = \left(\Lambda_j + \frac{1}{\sqrt{T}} G_j + \frac{1}{\sqrt{T}} \Lambda_j Q_j\right) \hat{\mathcal{U}}_j + o_p\left(T^{-1/2}\right), \qquad j = 1, 2, \tag{C.98}$$

where

$$G_j = \begin{bmatrix} G_j^c & \vdots & G_j^s \end{bmatrix} = \frac{1}{\sqrt{T}} \varepsilon_j' H_j, \qquad H_j = [F^c \vdots F_j^s], \tag{C.99}$$

$$\hat{\mathcal{U}}_j = \begin{bmatrix} \hat{\mathcal{H}}_c & 0 \\ 0 & \hat{\mathcal{H}}_{s,j} \end{bmatrix}, \tag{C.100}$$

$$Q_j = \begin{bmatrix} 0 & 0 \\ \sqrt{T}\tilde{\Sigma}_{j,c}\tilde{\Sigma}_{cc}^{-1} & 0 \end{bmatrix}. \tag{C.101}$$

To compute $\frac{\hat{\Lambda}_j'\hat{\Lambda}_j}{N_j}$, we consider the matrix product:

$$\begin{aligned} & \frac{1}{N_j}\left[\Lambda_j + \frac{1}{\sqrt{T}} G_j + \frac{1}{\sqrt{T}} \Lambda_j Q_j\right]' \left[\Lambda_j + \frac{1}{\sqrt{T}} G_j + \frac{1}{\sqrt{T}} \Lambda_j Q_j\right] \\ = \ & \frac{1}{N_j}\Lambda_j'\Lambda_j + \frac{1}{N_j\sqrt{T}}\left(\Lambda_j' G_j + G_j'\Lambda_j\right) + \frac{1}{N_j T} G_j' G_j + \frac{1}{\sqrt{T}}\left[\left(\frac{1}{N_j}\Lambda_j'\Lambda_j\right) Q_j + Q_j'\left(\frac{1}{N_j}\Lambda_j'\Lambda_j\right)\right] \\ & + \frac{1}{N_j T}\left(Q_j'\Lambda_j' G_j + G_j'\Lambda_j Q_j\right) + \frac{1}{T} Q_j'\left(\frac{1}{N_j}\Lambda_j'\Lambda_j\right) Q_j. \end{aligned} \tag{C.102}$$

Let us now bound the different terms. We have:

$$\frac{1}{\sqrt{N_j}}\Lambda_j' G_j = \frac{1}{\sqrt{N_j T}}\Lambda_j'\varepsilon_j' H_j = \frac{1}{\sqrt{N_j T}}\sum_{i=1}^{N_j}\sum_{t=1}^{T}\lambda_{j,i} h_{j,t}'\varepsilon_{j,it} = O_p(1),$$

and:

$$\frac{1}{N_j} G_j' G_j = \frac{1}{N_j}\sum_{i=1}^{N_j}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T} h_{j,t}\varepsilon_{j,i,t}\right)\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T} h_{j,t}\varepsilon_{j,i,t}\right)' = O_p(1),$$

by arguments similar to the proof of Lemma B.1. Thus, by using these bounds and $\Lambda_j'\Lambda_j/N_j = O(1)$ and $Q_j = O_p(1)$, from equation (C.102) we get:

$$\begin{aligned} \frac{1}{N_j}\left[\Lambda_j + \frac{1}{\sqrt{T}} G_j + \frac{1}{\sqrt{T}} \Lambda_j Q_j\right]' \left[\Lambda_j + \frac{1}{\sqrt{T}} G_j + \frac{1}{\sqrt{T}} \Lambda_j Q_j\right] = \ & \frac{1}{N_j}\Lambda_j'\Lambda_j + \frac{1}{\sqrt{T}}\left(L_{\Lambda,j} + L_{\Lambda,j}'\right) \\ & + O_p\left(\frac{1}{\sqrt{NT}} + \frac{1}{T}\right), \end{aligned}$$

where

$$L_{\Lambda,j} = \left(\frac{\Lambda_j'\Lambda_j}{N_j}\right) Q_j. \tag{C.103}$$

Therefore we have:

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

C.11.2 Proof of Lemma B.8 Part (ii)

a) Asymptotic expansion of $\hat{\Gamma}_j$

We start by deriving the uniform asymptotic expansion for the residuals. The asymptotic expansions in (C.92)-(C.95) allow to compute the asymptotic expansion of $\hat{\varepsilon}_{j,i,t}$:

$$\begin{aligned}
\hat{\varepsilon}_{j,i,t} &= y_{j,i,t} - \hat{\lambda}_{j,i}^{c\,\prime}\hat{f}_t^c - \hat{\lambda}_{j,i}^{s\,\prime}\hat{f}_{j,t}^s \;=\; \varepsilon_{j,i,t} - \left[\hat{\lambda}_{j,i}^{c\,\prime}\hat{f}_t^c - \lambda_{j,i}^{c\,\prime}f_t^c\right] - \left[\hat{\lambda}_{j,i}^{s\,\prime}\hat{f}_{j,t}^s - \lambda_{j,i}^{s\,\prime}f_{j,t}^s\right] \\
&= \varepsilon_{j,i,t} - \left[\left(\lambda_{j,i}^c + \tilde{\Sigma}_{cc}^{-1}\tilde{\Sigma}_{c,j}\lambda_{j,i}^s + \frac{1}{\sqrt{T}}w_{j,i}^c + o_p(T^{-1/2})\right)'\left(f_t^c + \frac{1}{\sqrt{N_1}}u_{1,t}^{(c)} + o_p(T^{-1/2})\right) - \lambda_{j,i}^{c\,\prime}f_t^c\right] \\
&\quad - \left[\left(\lambda_{j,i}^s + \frac{1}{\sqrt{T}}w_{j,i}^s + o_p(T^{-1/2})\right)'\left(f_{j,t}^s - \tilde{\Sigma}_{j,c}\tilde{\Sigma}_{cc}^{-1}f_t^c + \frac{1}{\sqrt{N_j}}u_{j,t}^{(s)} + o_p(T^{-1/2})\right) - \lambda_{j,i}^{s\,\prime}f_{j,t}^s\right] \\
&= \varepsilon_{j,i,t} - \left(\frac{1}{\sqrt{N_1}}\lambda_{j,i}^{c\,\prime}u_{1,t}^{(c)} + \frac{1}{\sqrt{T}}w_{j,i}^{c\prime}f_t^c\right) - \left(\frac{1}{\sqrt{N_j}}\lambda_{j,i}^{s\,\prime}u_{j,t}^{(s)} + \frac{1}{\sqrt{T}}w_{j,i}^{s\prime}f_{j,t}^s\right) + o_p\left(T^{-1/2}\right). \qquad \text{(C.104)}
\end{aligned}$$

Here the $o_p(T^{-1/2})$ term is uniform w.r.t. $1 \le i \le N_j$, $1 \le t \le T$ by the bounds in the next Lemma C.10 and Assumption A.8 *d)*.

LEMMA C.10. *Let $X = O_{p,\ell}(a_{N,T})$ mean $X = O_p[a_{N,T}(\log T)^{\bar{b}}]$ for some $\bar{b} > 0$. Under Assumption A.8 we have the following uniform bounds:*

$$\begin{aligned}
\sup_{1\le t\le T}\|h_{j,t}\| &= O_{p,\ell}(1), &\text{(C.105)}\\
\sup_{1\le t\le T}\|u_{j,t}\| &= O_{p,\ell}(1), &\text{(C.106)}\\
\sup_{1\le i\le N_j}\|\frac{1}{T}\sum_{t=1}^T h_{j,t}\varepsilon_{j,i,t}\| &= O_{p,\ell}(T^{-\eta/2}), &\text{(C.107)}
\end{aligned}$$

where $\eta \ge 1/2$.

If we adopt $\hat{f}_t^c$ to compute residuals in panel $j = 1$, and $\hat{f}_t^{c*}$ for $j = 2$, we have:

$$\hat{\varepsilon}_{j,i,t} = \varepsilon_{j,i,t} - \frac{1}{\sqrt{T}}\left(w_{j,i}^{c\prime}f_t^c + w_{j,i}^{s\prime}f_{j,t}^s\right) - \frac{1}{\sqrt{N_j}}\left(\lambda_{j,i}^{c\,\prime}u_{j,t}^{(c)} + \lambda_{j,i}^{s\,\prime}u_{j,t}^{(s)}\right) + o_p\left(T^{-1/2}\right). \qquad \text{(C.108)}$$

Equation (C.108) allows us to compute:

$$\begin{aligned}\hat{\gamma}_{j,ii} &= \frac{1}{T}\sum_{t=1}^{T}\hat{\varepsilon}_{j,i,t}^2 \;=\; \frac{1}{T}\sum_{t=1}^{T}\left[\varepsilon_{j,i,t} - \frac{1}{\sqrt{T}}\left(w_{j,i}^{c\prime}f_t^c + w_{j,i}^{s\prime}f_{j,t}^s\right) - \frac{1}{\sqrt{N_j}}\left(\lambda_{j,i}^{c\,\prime}u_{j,t}^{(c)} + \lambda_{j,i}^{s\,\prime}u_{j,t}^{(s)}\right)\right]^2 + o_p\left(T^{-1/2}\right)\\
&= \frac{1}{T}\sum_{t=1}^{T}\varepsilon_{j,i,t}^2 - \frac{2}{T\sqrt{T}}\sum_{t=1}^{T}\varepsilon_{j,i,t}\left(w_{j,i}^{c\,\prime}f_t^c + w_{j,i}^{s\,\prime}f_{j,t}^s\right) - \frac{2}{T\sqrt{N_j}}\sum_{t=1}^{T}\varepsilon_{j,i,t}\left(\lambda_{j,i}^{c\,\prime}u_{j,t}^{(c)} + \lambda_{j,i}^{s\,\prime}u_{j,t}^{(s)}\right)\\
&\quad + \frac{1}{T^2}\sum_{t=1}^{T}\left(w_{j,i}^{c\,\prime}f_t^c + w_{j,i}^{s\,\prime}f_{j,t}^s\right)^2 + \frac{1}{TN_j}\sum_{t=1}^{T}\left(\lambda_{j,i}^{c\,\prime}u_{j,t}^{(c)} + \lambda_{j,i}^{s\,\prime}u_{j,t}^{(s)}\right)^2\\
&\quad + \frac{2}{T\sqrt{TN_j}}\sum_{t=1}^{T}\left(w_{j,i}^{c\,\prime}f_t^c + w_{j,i}^{s\,\prime}f_{j,t}^s\right)\left(\lambda_{j,i}^{c\,\prime}u_{j,t}^{(c)} + \lambda_{j,i}^{s\,\prime}u_{j,t}^{(s)}\right) + o_p\left(T^{-1/2}\right).\end{aligned}$$

By solving out the parentheses, using $w_{j,i}^c = \frac{1}{\sqrt{T}}\sum_{t=1}^T \varepsilon_{j,i,t}f_t^c = O_p(1)$, $w_{j,i}^s = \frac{1}{\sqrt{T}}\sum_{t=1}^T \varepsilon_{j,i,t}f_{j,t}^s = O_p(1)$, $\frac{1}{\sqrt{T}}\sum_{t=1}^T \varepsilon_{j,i,t}u_{j,t}^{(c)} = O_p(1)$ and $\frac{1}{\sqrt{T}}\sum_{t=1}^T \varepsilon_{j,i,t}u_{j,t}^{(s)} = O_p(1)$, uniformly in $1 \leq i \leq N_j$, we get:

$$\hat{\gamma}_{j,ii} = \frac{1}{T}\sum_{t=1}^{T}\varepsilon_{j,i,t}^2 + O_p\left(\frac{1}{N}\right) + o_p\left(T^{-1/2}\right),$$

uniformly in $1 \leq i \leq N_j$. Using that $1/N = o(1/\sqrt{T})$ when $\sqrt{T} \ll N$, we get:

$$\hat{\gamma}_{j,ii} = \frac{1}{T}\sum_{t=1}^{T}\varepsilon_{j,i,t}^2 + o_p\left(T^{-1/2}\right) = \gamma_{j,ii} + \frac{1}{\sqrt{T}}w_{j,i}^{\varepsilon} + o_p\left(T^{-1/2}\right),$$

uniformly in $1 \leq i \leq N_j$, where

$$w_{j,i}^{\varepsilon} := \frac{1}{\sqrt{T}}\sum_{t=1}^{T}(\varepsilon_{j,i,t}^2 - \gamma_{j,ii}).$$

Therefore, we have:

$$\hat{\Gamma}_j = \Gamma_j + \frac{1}{\sqrt{T}}W_j^{\varepsilon} + o_p\left(T^{-1/2}\right), \tag{C.109}$$

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

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

From (C.98) and (C.109) we have:

$$\frac{1}{N_j}\hat{\Lambda}_j'\hat{\Gamma}_j\hat{\Lambda}_j = \hat{\mathcal{U}}_j'\hat{\Omega}_{jj}^*\hat{\mathcal{U}}_j + o_p\left(T^{-1/2}\right), \tag{C.110}$$

where we define:

$$\begin{aligned}
\hat{\Omega}^*_{jj} &:= \frac{1}{N_j}\left(\Lambda_j + \frac{1}{\sqrt{T}}G_j + \frac{1}{\sqrt{T}}\Lambda_j Q_j\right)'\left(\Gamma_j + \frac{1}{\sqrt{T}}W^{\varepsilon}_j\right)\left(\Lambda_j + \frac{1}{\sqrt{T}}G_j + \frac{1}{\sqrt{T}}\Lambda_j Q_j\right) \\
&= \tilde{\Omega}_{jj} + \hat{\Omega}^*_{jj,I} + \hat{\Omega}^*_{jj,II} + \hat{\Omega}^{*\prime}_{jj,II} + \hat{\Omega}^*_{jj,III} + \hat{\Omega}^{*\prime}_{jj,III} + \hat{\Omega}^*_{jj,IV} + \hat{\Omega}^*_{jj,V} \\
&\quad + \frac{1}{\sqrt{T}}(\tilde{\Omega}_{jj}Q_j + Q_j'\tilde{\Omega}_{jj}) + \frac{1}{\sqrt{T}}(\hat{\Omega}^*_{jj,I}Q_j + Q_j'\hat{\Omega}^*_{jj,I}) + \frac{1}{\sqrt{T}}(\hat{\Omega}^*_{jj,II}Q_j + Q_j'\hat{\Omega}^{*\prime}_{jj,II}) \\
&\quad + \frac{1}{\sqrt{T}}(\hat{\Omega}^*_{jj,III}Q_j + Q_j'\hat{\Omega}^{*\prime}_{jj,III}) + \frac{1}{T}Q_j'\tilde{\Omega}_{jj}Q_j + \frac{1}{T}Q_j'\hat{\Omega}^*_{jj,I}Q_j,
\end{aligned}$$

and:

$$\begin{aligned}
\tilde{\Omega}_{jj} &:= \frac{1}{N_j}\Lambda_j'\Gamma_j\Lambda_j, \\
\hat{\Omega}^*_{jj,I} &:= \frac{1}{N_j\sqrt{T}}\Lambda_j'W^{\varepsilon}_j\Lambda_j = O_p\left(\frac{1}{\sqrt{NT}}\right), \\
\hat{\Omega}^*_{jj,II} &:= \frac{1}{N_j\sqrt{T}}G_j'\Gamma_j\Lambda_j = O_p\left(\frac{1}{\sqrt{NT}}\right), \\
\hat{\Omega}^*_{jj,III} &:= \frac{1}{N_jT}G_j'W^{\varepsilon}_j\Lambda_j = O_p\left(\frac{1}{T}\right), \\
\hat{\Omega}^*_{jj,IV} &:= \frac{1}{N_jT}G_j'\Gamma_jG_j = O_p\left(\frac{1}{T}\right), \\
\hat{\Omega}^*_{jj,V} &:= \frac{1}{N_jT\sqrt{T}}G_j'W^{\varepsilon}_jG_j = O_p\left(\frac{1}{T\sqrt{T}}\right).
\end{aligned}$$

Collecting the previous results, we get:

$$\hat{\Omega}^*_{jj} = \tilde{\Omega}_{jj} + \frac{1}{\sqrt{T}}\left(L_{\Omega,j} + L'_{\Omega,j}\right) + O_p\left(\frac{1}{\sqrt{NT}} + \frac{1}{T}\right), \tag{C.111}$$

where:

$$L_{\Omega,j} = \tilde{\Omega}_{jj}Q_j. \tag{C.112}$$

By substituting into equation (C.110) we get:

$$\frac{1}{N_j}\hat{\Lambda}_j'\hat{\Gamma}_j\hat{\Lambda}_j = \hat{\mathcal{U}}_j'\left[\tilde{\Omega}_{jj} + \frac{1}{\sqrt{T}}\left(L_{\Omega,j} + L'_{\Omega,j}\right)\right]\hat{\mathcal{U}}_j + o_p\left(T^{-1/2}\right), \qquad j = 1, 2.$$

■

C.11.3 Proof of Lemma C.10

We prove the uniform bounds in (C.105) and (C.107). The proof of bound (C.106) follows by similar arguments. *Proof of (C.105).* Let $\delta = c(\log T)^{\bar{b}}$, for constants $c > 0$ and $\bar{b} = 1/b$, where $b > 0$ is defined in Assumption

A.8 a). Then:

$$\begin{aligned} P[\sup_{1\le t\le T}\|h_{j,t}\|\ge\delta] &\le \sum_{t=1}^{T}P[\|h_{j,t}\|\ge\delta]\le c_1T\exp(-c_2\delta^b)=c_1T\exp[-c_2c^b(\log T)] \\ &= c_1T^{1-c_2c^b}=o(1), \end{aligned}$$

if $c>(1/c_2)^{1/b}$. Thus, $\sup\limits_{1\le t\le T}\|h_{j,t}\|=O_p[(\log T)^{\bar{b}}]$.

Proof of (C.107). Let $\delta=c(\log T)^{1/2}T^{-\eta/2}$, for constants $c>0$ and η, where $\eta\ge 1/2$ is defined in Assumption A.8 c). Then:

$$\begin{aligned} P[\sup_{1\le i\le N_j}\|\frac{1}{T}\sum_{t=1}^{T}h_{j,t}\varepsilon_{j,i,t}\|\ge\delta] &\le \sum_{i=1}^{N_j}P[\|\frac{1}{T}\sum_{t=1}^{T}h_{j,t}\varepsilon_{j,i,t}\|\ge\delta]\le N_j\sup_{1\le i\le N_j}P[\|\frac{1}{T}\sum_{t=1}^{T}h_{j,t}\varepsilon_{j,i,t}\|\ge\delta] \\ &\le c_1N_jT\exp(-c_2\delta^2T^{\eta})+c_3TN_j\delta^{-1}\exp(-c_4T^{\bar{\eta}}) \\ &= c_1N_jT\exp(-c_2c^2(\log T))+c_3TN_j\delta^{-1}\exp(-c_4T^{\bar{\eta}}) \\ &= O(T^{7/2-c_2c^2})+o(1)=o(1), \end{aligned}$$

if $c>(\frac{7}{2c_2})^{1/2}$. Thus, $\sup\limits_{1\le i\le N_j}\|\frac{1}{T}\sum_{t=1}^{T}h_{j,t}\varepsilon_{j,i,t}\|=O_p[(\log T)^{1/2}T^{-\eta/2}]=O_{p,\ell}(T^{-\eta/2})$.

C.12 Proof of Lemma B.9

We assume that estimator $\hat{f}_t^c$ is used to get factor loadings on panel $j=1$, and estimator $\hat{f}_t^{c\,*}$ is used to get factor loadings on panel $j=2$. Recall $\hat{\Sigma}_U=(N_2/N_1)\hat{\Sigma}_{u,11}^{(cc)}+\hat{\Sigma}_{u,22}^{(cc)}$. Let r be the true number of common factors, and let k^c denote the number of common factors used in the estimation procedure. We consider the case with $r<k^c\le\underline{k}\equiv\min\{k_1,k_2\}$.

Let us first consider panel $j=1$. The common factor estimator is $\hat{f}_t^c=\hat{W}_1'\hat{h}_{1,t}$ where $\hat{W}_1$ is the $k_1\times k^c$ matrix whose columns are eigenvectors of $\hat{R}$ associated with the k^c largest eigenvalues, normalized to have $\hat{W}_1'\hat{W}_1=I_{k^c}$. Without loss of generality, let $\hat{\mathcal{H}}_j=I_{k_j}$ in Proposition 3. Then, we have $\hat{R}=R+o_p(1)$, where $R=\begin{pmatrix} I_r & 0\\ 0 & \Phi\Phi'\end{pmatrix}$. The large-sample limit of $\hat{W}_1$ is the matrix of normalized eigenvectors associated to the k^c largest eigenvalues of matrix R. These eigenvalues are 1, with multiplicity r, and $\rho_{r+1}^2,...,\rho_{k^c}^2$, that are the k^c-r largest eigenvalues of matrix $\Phi\Phi'$ (assumed distinct, to simplify the proof). Let α denote the $(k_1-r)\times(k^c-r)$ matrix whose columns are the corresponding normalized eigenvectors of $\Phi\Phi'$. Then, we have $\hat{W}_1=W_1+o_p(1)$ where

$$W_1=\begin{bmatrix}\mathcal{U} & 0\\ 0 & \alpha\end{bmatrix},$$

$r\times r$ matrix $\mathcal{U}$ is possibly stochastic and such that $\mathcal{U}'\mathcal{U}=I_r$, and $\alpha'\alpha=I_{k^c-r}$. For later use, we denote by β the $(k_1-r)\times(k_1-k^c)$ matrix whose columns are an orthonormal basis of the orthogonal complement to the columns space of α. Then, $[\alpha\vdots\beta]$ is an orthogonal matrix, $\beta'\beta=I_{k_1-k^c}$, $\alpha'\beta=0$, and:

$$\alpha\alpha'+\beta\beta'=I_{k_1-r}. \tag{C.113}$$

From Proposition 3 with $\hat{\mathcal{H}}_j=I_{k_j}$ we have $\hat{h}_{j,t}\simeq h_{j,t}$, where symbol $\simeq$ means equality up to terms that are asymptotically negligible for determining large-sample limits. Then:

$$\hat{f}_t^c \simeq W_1' h_{1,t} = \begin{bmatrix} \mathcal{U}' f_t^c \\ \alpha' f_{1,t}^s \end{bmatrix}.$$

Let us consider the estimation of the factor loadings on the panel with $j = 1$. From (C.113) the model for this panel can be written as:

$$\begin{aligned} y_{1,i,t} &= f_t^{c\,\prime} \lambda_{1,i}^c + f_{1,t}^{s\,\prime} \lambda_{1,i}^s + \varepsilon_{1,i,t} = [\mathcal{U} f_t^{c\,\prime}][\mathcal{U} \lambda_{1,i}^c] + [\alpha' f_{1,t}^s]'[\alpha' \lambda_{1,i}^s] + [\beta' f_{1,t}^s]'[\beta' \lambda_{1,i}^s] + \varepsilon_{1,i,t} \\ &= \underline{f}_t^{c\,\prime} \underline{\lambda}_{1,i}^c + \underline{f}_{1,t}^{s\,\prime} \underline{\lambda}_{1,i}^s + \varepsilon_{1,i,t}, \end{aligned}$$

where $\underline{f}_t^c = \begin{bmatrix} \mathcal{U}' f_t^c \\ \alpha' f_{1,t}^s \end{bmatrix}$, $\underline{\lambda}_{1,i}^c = \begin{bmatrix} \mathcal{U}' \lambda_{1,i}^c \\ \alpha' \lambda_{1,i}^s \end{bmatrix}$, $\underline{f}_{1,t}^s = \beta' f_{1,t}^s$ and $\underline{\lambda}_{1,i}^s = \beta' \lambda_{1,i}^s$. Note that the transformed factors $\underline{f}_t^c$ and $\underline{f}_{1,t}^s$ are orthogonal, and have dimensions k^c and $k_1 - k^c$ respectively. Since $\hat{f}_t^c$ converges to $\underline{f}_t^c$, by regressing $y_{1,i,t}$ onto $\hat{f}_t^c$ we estimate $\underline{\lambda}_{1,i}^c$. Then, the residuals satisfy the model:

$$\xi_{1,i,t} \simeq \underline{f}_{1,t}^{s\,\prime} \underline{\lambda}_{1,i}^s + \varepsilon_{1,i,t}.$$

The frequency-specific factor is estimated by extracting the first $k_1 - k^c$ principal components from the residuals, which yields asymptotically $\hat{f}_{1,t}^s \simeq \mathcal{V} \underline{f}_t^s$, where $\mathcal{V}$ is an orthogonal matrix. So for the estimated factor loadings we have:

$$\hat{\lambda}_{1,i}^c \simeq \underline{\lambda}_{1,i}^c = \begin{bmatrix} \mathcal{U}' \lambda_{1,i}^c \\ \alpha' \lambda_{1,i}^s \end{bmatrix}, \qquad \hat{\lambda}_{1,i}^s \simeq \mathcal{V} \underline{\lambda}_{1,i}^s = \mathcal{V} \beta' \lambda_{1,i}^s.$$

Thus, $\hat{\lambda}_{1,i}$ is asymptotically an orthogonal transformation of $\lambda_{1,i}$, i.e. $\hat{\lambda}_{1,i} \simeq \mathcal{R}_1 \lambda_{1,i}$, say. Using $\hat{\varepsilon}_{1,i,t} \simeq \varepsilon_{1,i,t}$, we get $\hat{\Sigma}_{u,11} \simeq \mathcal{R}_1 \Sigma_{u,11} \mathcal{R}_1'$, which implies $\hat{\Sigma}_{u,11} = O_p(1)$.
Let us now consider the estimation of factor loadings in panel $j = 2$. By paralleling the above arguments, we have $\hat{\Sigma}_{u,22} = O_p(1)$. Thus, $\|\hat{\Sigma}_U\| = O_p(1)$. The conclusion follows. ■

D Additional theoretical and empirical results

Section D.1 discusses the separation of common and group specific factors for identification purposes in generic group-factor models. Section D.2 provides details about an alternative identification strategy, different from the canonical correlation analysis proposed in Proposition 1 of Andreou, Gagliardini, Ghysels, and Rubin (2019), for the common and group-specific factor spaces in a group-factor model. Section D.3 discusses the identification of the mixed frequency factor model in the cases of stock-sampling, and of general linear aggregation schemes for the LF observables. Sections D.4 and D.5 provide uniform asymptotic expansions and asymptotic distributions of factors and loadings estimators in a group factor model. Section D.6 provides the asymptotic distribution of factors and loadings estimators in a mixed frequency model. Section D.7 contains a digression on some technical assumptions. Section D.8 contains a discussion of properties of an iterative PCA estimator for group factor models. A description of the practical implementation of our estimation and testing procedures appears in Section D.9. Section D.10 describes exhaustively the dataset used in the empirical application of Section 7. Section D.11 presents additional empirical results.

D.1 Separation of common and group-specific factors

The following proposition gives a sufficient condition for the identification of the group factor model (2.1) - (2.2) when the factor dimensions k^c, k_1^s, k_2^s are known.

PROPOSITION D.1. *Assume that the 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_1, N_2 *large enough. Then, the factor model is identifiable: the data* $[y_{1,t}', y_{2,t}']'$ *satisfy a group factor model as (2.1) - (2.2) with stacked factor* $(f_t^{c\,\prime}, f_{1,t}^{s\,\prime}, f_{2,t}^{s\,\prime})'$ *replaced by* $(\tilde{f}_t^{c\,\prime}, \tilde{f}_{1,t}^{s\,\prime}, \tilde{f}_{2,t}^{s\,\prime})'$ *defined by the linear transformation*

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

if, and only if, the matrix $A = (A_{i,j})$ *is a block-diagonal orthogonal matrix.*

The full-rank condition in Proposition D.1 is a standard condition for separate identification of the pervasive factor spaces in the two subgroups. The identification condition in Proposition D.1 is implied by Assumption A.3, and implies that the matrix of loadings in the right hand side of equation (2.1) is full-rank. Proposition D.1 shows that this condition - together with the normalization restrictions in (2.2) - is also sufficient for identifiability of the common factor f_t^c, the group-specific factors $f_{j,t}^s$, and the factor loadings Λ_j^c, Λ_j^s, up to separate rotations. Hence, the rotational invariance of model (2.1) - (2.2) maintains the interpretation of common factor and group-specific factors.

D.1.1 Proof of Proposition D.1

By replacing equation (D.1) into model (2.1), we get

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} \Lambda_1^c A_{11} + \Lambda_1^s A_{21} & \Lambda_1^c A_{12} + \Lambda_1^s A_{22} & \Lambda_1^c A_{13} + \Lambda_1^s A_{23} \\ \Lambda_2^c A_{11} + \Lambda_2^s A_{31} & \Lambda_2^c A_{12} + \Lambda_2^s A_{32} & \Lambda_2^c A_{13} + \Lambda_2^s A_{33} \end{bmatrix} \begin{bmatrix} \tilde{f}_t^c \\ \tilde{f}_{1,t}^s \\ \tilde{f}_{2,t}^s \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}.$$

This factor model satisfies the restrictions in the loading matrix appearing in equation (2.1) if, and only if, $\Lambda_1^c A_{13} + \Lambda_1^s A_{23} = 0$, and $\Lambda_2^c A_{12} + \Lambda_2^s A_{32} = 0$, which can be written as linear homogeneous systems of equations for the elements of matrices $[A_{13}' \; A_{23}']'$ and $[A_{12}' \; A_{32}']'$:

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

Since $\begin{bmatrix} \Lambda_1^c \vdots \Lambda_1^s \end{bmatrix}$ and $\begin{bmatrix} \Lambda_2^c \vdots \Lambda_2^s \end{bmatrix}$ are full column rank, it follows that $A_{13} = 0$, $A_{23} = 0$, $A_{12} = 0$, and $A_{32} = 0$. Therefore, the transformation of the factors that is compatible with the restrictions on the loading matrix in equation (2.1) 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:

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

The transformed factors satisfy the normalization restrictions in (2.2) if, and only if,

$$Cov(\tilde{f}_{1,t}^s,\ \tilde{f}_t^c) = -A_{22}^{-1} A_{21} A_{11}^{-1} (A_{11}^{-1})' = 0, \tag{D.2}$$

$$Cov(\tilde{f}_{2,t}^s,\ \tilde{f}_t^c) = -A_{33}^{-1} A_{31} A_{11}^{-1} (A_{11}^{-1})' = 0, \tag{D.3}$$

$$V(\tilde{f}_t^c) = A_{11}^{-1} (A_{11}^{-1})' = I_{k^c}, \tag{D.4}$$

$$V(\tilde{f}_{1,t}^s) = A_{22}^{-1} (A_{22}^{-1})' + A_{22}^{-1} A_{21} A_{11}^{-1} (A_{11}^{-1})' A_{21}' (A_{22}^{-1})' = I_{k_1^s}, \tag{D.5}$$

$$V(\tilde{f}_{2,t}^s) = A_{33}^{-1} (A_{33}^{-1})' + A_{33}^{-1} A_{31} A_{11}^{-1} (A_{11}^{-1})' A_{31}' (A_{33}^{-1})' = I_{k_2^s}. \tag{D.6}$$

Since the matrices A_{11}, A_{22} and A_{33} are nonsingular, equations (D.2) and (D.3) imply $A_{21} = 0$, and $A_{31} = 0$. Then, from equations (D.4) - (D.6), we get that matrices A_{11}, A_{22} and A_{33} are orthogonal.

■

D.2 Identification of the common and group-specific factor spaces from the variance-covariance matrix of stacked factors

In this section we provide an identification strategy for the common and group-specific factor spaces in a group-factor model, which is alternative to canonical correlation analysis proposed in Proposition 1. The identification of the factor spaces is achieved through an eigenvalue-eigenvector decomposition of the variance-covariance matrix of the stacked principal components extracted separately from the two different groups of data.
We define the matrices $w_j = [w_{j,1}, ..., w_{j,\underline{k}}]$, $j = 1, 2$, with the canonical directions. These matrices are such that $w_j' V_{jj} w_j = I_{\underline{k}}$, $j = 1, 2$. Moreover, when $\rho_\ell \neq 0$, then

$$w_{1,\ell} = \frac{1}{\rho_\ell} V_{11}^{-1} V_{12} w_{2,\ell}\,, \qquad w_{2,\ell} = \frac{1}{\rho_\ell} V_{22}^{-1} V_{21} w_{1,\ell}\,. \tag{D.7}$$

The principal components are normalized such that $V(h_{j,t}) = I_{k_j}$, for $j = 1, 2$.

LEMMA D.2. *Let $h_t = [h_{1,t}'\ \ h_{2,t}']'$, be a random vector, such that $V_{11} = V(h_{1,t}) = I_{k_1}$, $V_{22} = V(h_{2,t}) = I_{k_2}$, $V_{12} = Cov(h_{1,t}, h_{2,t})$ and let $V(h_t)$ be the variance-covariance matrix of vector h_t:*

$$V(h_t) = \begin{bmatrix} I_{k_1} & V_{12} \\ V_{21} & I_{k_2} \end{bmatrix}.$$

Let $r = rank(V_{21})$, with $r \leq \underline{k} = \min(k_1, k_2)$. Then, matrix $V(h_t)$ has $2r$ eigenvalues $1 \pm \rho_\ell$, $\ell = 1, ..., r$, with multiplicity 1, corresponding to the non-zero canonical correlations,$\rho_\ell \neq 0$ between $h_{1,t}$ and $h_{2,t}$, and the eigenvalue 1 with multiplicity $k_1 + k_2 - 2r$. The eigenvectors of $V(h_t)$ associated with the eigenvalues $1 \pm \rho_\ell$, $\ell = 1, ..., r$ are

$$v_\ell^{\pm} = \begin{bmatrix} w_{1,\ell} \\ \pm w_{2,\ell} \end{bmatrix},\ \ell = 1, ..., r,$$

where $w_{1,\ell}$ (resp. $w_{2,\ell}$), are the normalized eigenvectors of $R = V_{12} V_{21}$ (resp. $R^* = V_{21} V_{12}$), associated with eigenvalues ρ_ℓ^2.

From Proposition 1 and Lemma D.2 we get the next Proposition, which yields an identification result for group factor models.

PROPOSITION D.3. *i) The number k^c of common factors is equal to the multiplicity of the eigenvalue 2 of matrix $V(h_t)$. ii) Let W be the (k_1+k_2, k^c) matrix whose columns are the orthonormal eigenvectors associated with the k^c eigenvalues of $V(h_t)$ equal to 2. Then, $f_t^c = \dfrac{1}{\sqrt{2}}W'h_t$ (up to a one-to-one transformation).*

Proposition D.3 is analogous to Proposition 3.1 in Chen (2012). Our derivation of Proposition D.3 as a consequence of Lemma D.2 and Proposition 1 allows us to clarify the link between eigenvalues equal to 2 of the stacked variance-covariance matrix and unit canonical correlations. Moreover, Lemma D.2 and Proposition D.3 admit a rather straightforward generalization to the case of a generic number of groups. Those results would be the basis for extending the results of our paper to more than two groups, i.e., more than two sampling frequencies.

The sample counterparts of the results in Lemma D.2 and Proposition D.3 suggest an alternative estimator of the common factor, which has been sometimes applied in the literature (see e.g. Goyal, Pérignon, and Villa (2008) and the reference therein). In the notation of Section 3 of the paper, a consistent estimator of matrix $V(h_t)$ is $\hat{V} = \frac{1}{T}\hat{H}'\hat{H}$ where $\hat{H} := [\hat{H}_1 \vdots \hat{H}_2]$, and it holds:

$$\hat{V}\hat{W} = \hat{W}(I_{k^c} + \hat{\Lambda}), \tag{D.8}$$

where $\hat{W} := \frac{1}{\sqrt{2}}\begin{bmatrix} \hat{W}_1 \\ \hat{W}_2 \end{bmatrix}$ is a matrix with orthonormal columns in which we stack the canonical directions in the two groups, and $\hat{\Lambda}$ is the diagonal matrix that collects the k^c largest estimated canonical correlations. Now, by using that matrices $\hat{H}'\hat{H}$ and $\hat{H}\hat{H}'$ have the same non zero eigenvalues, and by pre-multiplying equation (D.8) times $\frac{1}{\sqrt{2}}\hat{H}$ we get $(\frac{1}{T}\hat{H}\hat{H}')\hat{F}^{c\star} = \hat{F}^{c\star}(I_{k^c} + \hat{\Lambda})$, where $\hat{F}^{c\star} := \frac{1}{\sqrt{2}}\hat{H}\hat{W}$. Thus, we get a $T \times k^c$ matrix of common factor estimates $\hat{F}^{c\star}$ as the matrix of eigenvectors to the k^c largest eigenvalues (equal to $1 + \hat{\rho}_\ell$, $\ell = 1, ..., k^c$) of matrix $\frac{1}{T}\hat{H}\hat{H}' = \frac{1}{T}(\hat{H}_1\hat{H}_1' + \hat{H}_2\hat{H}_2')$. We have $\hat{F}^{c\star} = \frac{1}{2}(\hat{H}_1\hat{W}_1 + \hat{H}_2\hat{W}_2) = \frac{1}{2}(\hat{F}^c + \hat{F}^{c*})$, i.e., the average of the two estimators in Definition 1.

D.2.1 Proof of Lemma D.2

Let ρ_i, $i = 1, ..., \underline{k}$, be the canonical correlations between $h_{1,t}$ and $h_{2,t}$. From Anderson (2003) and Magnus and Neudecker (2007), ρ_i^2 corresponds to the i-th ordered eigenvalue of matrix $R = V_{12}V_{21}$. Let $1 + \mu$, say, be an eigenvalue of matrix $V(h_t)$, and $Z = [Z_1' \ Z_2']' \in \mathbb{R}^{k_1+k_2}$ be the associated (normalized) eigenvector. We have:

$$V(h_t)Z \quad = \quad (1 + \mu)Z \ .$$

Rewriting matrix $V(h_t)$ as:

$$V(h_t) \quad = \quad I_{k_1+k_2} + \begin{bmatrix} 0 & V_{12} \\ V_{21} & 0 \end{bmatrix}, \tag{D.9}$$

we get:

$$\begin{bmatrix} 0 & V_{12} \\ V_{21} & 0 \end{bmatrix}\begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} \quad = \quad \mu \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix}. \tag{D.10}$$

The last equation implies:

$$
\begin{aligned}
V_{12}Z_2 &= \mu Z_1, && \text{(D.11)}\\
V_{21}Z_1 &= \mu Z_2, && \text{(D.12)}
\end{aligned}
$$

and:

$$
\begin{aligned}
V_{12}V_{21}Z_1 &= \mu^2 Z_1, && \text{(D.13)}\\
V_{21}V_{12}Z_2 &= \mu^2 Z_2. && \text{(D.14)}
\end{aligned}
$$

If $Z_1 \neq 0$, then μ^2 is an eigenvalue of $V_{12}V_{21}$, i.e. a squared canonical correlation, and if $Z_2 \neq 0$, then μ^2 is an eigenvalue of $V_{21}V_{12}$. From the condition $rank(V_{21}) = r$, with $r \leq \underline{k}$, there are r canonical correlations different from zero: $\rho_1 \geq ... \geq \rho_r > 0$. Let $w_{1,\ell}$, $\ell = 1, ..., r$, be the associated eigenvectors of $R = V_{12}V_{21}$, and $w_{2,\ell}$, $\ell = 1, ..., r$ the corresponding eigenvectors of $R^* = V_{21}V_{12}$. Then, the scalars

$$
\mu_{\ell,\pm} = \pm\rho_\ell, \qquad \ell = 1, ..., r,
$$

and the vectors

$$
v_\ell^{\pm} = \left[\begin{array}{c} w_{1,\ell} \\ \pm\frac{1}{\rho_\ell} V_{21} w_{1,\ell} \end{array} \right] = \left[\begin{array}{c} w_{1,\ell} \\ \pm w_{2,\ell} \end{array} \right] \tag{D.15}
$$

solve equation (D.10). Here, we use $\frac{1}{\rho_\ell}V_{21}w_{1,\ell} = w_{2,\ell}$, from property (D.7). Therefore, $1 \pm \rho_\ell$ are eigenvalues of $V(h_t)$ associated with eigenvectors $v_\ell^{\pm}$, with $\ell = 1, ..., r$.
Let us now consider the solutions of equation (D.10) with $\mu = 0$. We have:

$$
\begin{aligned}
V_{12}Z_2 &= 0, && \text{(D.16)}\\
V_{21}Z_1 &= 0. && \text{(D.17)}
\end{aligned}
$$

From $rank(V_{12}) = r$, the null space of matrix V_{12} is $(k_2 - r)$-dimensional. Let the columns of the $(k_2, k_2 - r)$ full column rank matrix $\tilde{Z}_2$ span the $(k_2 - r)$-dimensional space of solutions of equation (D.16). Similarly, let the columns of the $(k_1, k_1 - r)$ full column rank matrix $\tilde{Z}_1$ span the $(k_1 - r)$-dimensional space of solutions of equation (D.17). Define the $(k_1 + k_2, 2(k_2 - r))$ matrix:

$$
\tilde{Z}_0 = \left[\begin{array}{cc|cc} \tilde{Z}_1 & O_{k_1 \times (k_2 - k_1)} & -\tilde{Z}_1 & O_{k_1 \times (k_2 - k_1)} \\ \multicolumn{2}{c|}{\tilde{Z}_2} & \multicolumn{2}{c}{\tilde{Z}_2} \end{array} \right].
$$

Any column of this matrix is a solution of (D.10) with $\mu = 0$. Since matrices $\tilde{Z}_1$ and $\tilde{Z}_2$ are full column rank, the column rank of matrix Z_0 is $2(k_1 - r) + (k_2 - k_1) = k_1 + k_2 - 2r$. Therefore, there are $k_1 + k_2 - 2r$ linearly independent eigenvectors of $\left[\begin{array}{cc} 0 & V_{12} \\ V_{21} & 0 \end{array} \right]$ associated with the eigenvalue 0. These vectors are eigenvectors of $V(h_t)$ associated with the eigenvalue 1. ∎

D.2.2 Proof of Proposition D.3

From Lemma D.2, $V(h_t)$ has eigenvalue 2 if, and only if, there is a canonical correlation equal to 1. Part *i)* follows from Proposition 1 *i)*. Moreover, from Proposition 1 and Lemma D.2 the columns of matrix $W = \frac{1}{\sqrt{2}} \left[\begin{array}{c} W_1 \\ W_2 \end{array} \right]$ are orthonormal eigenvectors of $V(h_t)$ associated with eigenvalue 2, since $W'W = \frac{1}{2}(W_1'W_1 +$

$W_2'W_2) = I_{k^c}$. Finally, $\frac{1}{\sqrt{2}}W'h_t = \frac{1}{2}(W_1'h_{1,t} + W_2'h_{2,t}) = \frac{1}{2}(S_1' + S_2')f_t^c$, from (C.1) and (C.2), which implies part *ii)*. ■

D.3 Identification with stock-sampling and generic linear aggregation schemes

In the case of stock-sampling, the low frequency observations of $x_{m,t}^{L*}$ in the mixed-frequency factor model (5.1) are the values of $x_{M,t}^{L*}$, i.e. $x_t^L = x_{M,t}^{L*}$. Then, the model for the observable variables becomes:

$$\begin{aligned} x_{m,t}^H &= \Lambda_{HC} g_{m,t}^C + \Lambda_H g_{m,t}^H + e_{m,t}^H, \quad m = 1, ..., M, \\ x_t^L &= \Lambda_{LC} g_{M,t}^C + \Lambda_L g_{M,t}^L + e_{M,t}^L. \end{aligned}$$

We stack the observations $x_t^H = x_{M,t}^H$ and x_t^L of the last high frequency subperiod and write:

$$\begin{bmatrix} x_t^H \\ x_t^L \end{bmatrix} = \begin{bmatrix} \Lambda_{HC} & \Lambda_H & 0 \\ \Lambda_{LC} & 0 & \Lambda_L \end{bmatrix} \begin{bmatrix} g_{M,t}^C \\ g_{M,t}^H \\ g_{M,t}^L \end{bmatrix} + \begin{bmatrix} e_{M,t}^H \\ e_{M,t}^L \end{bmatrix}. \tag{D.18}$$

This last equation corresponds to a group factor model, with common factor $g_{M,t}^C$ and "group-specific" factors $g_{M,t}^H$, $g_{M,t}^L$. Therefore, the factor values $g_{M,t}^C$, $f_{M,t}^H$, $f_{M,t}^L$, and the factor loadings Λ_{HC}, Λ_{LC}, Λ_H, Λ_L, are identifiable up to a sign as proved in Section 2.

Once the factor loadings are identified from equation (D.18), 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 Λ_{HC} and Λ_H in (5.1). 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 $i = 1, 2, ..., N_H$, for any $m = 1, ..., M-1$ and any t. To summarize, with stock-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. We cannot estimate $g_{m,t}^L$, for $m < M$, as only $g_{M,t}^L$ is identified by the last paired panel data set consisting of $x_{M,t}^H$ combined with x_t^L. This is not surprising, since we have no HF observation available for the LF process.

Flow sampling and stock-sampling are examples of linear aggregation schemes. The case of a general linear aggregation scheme relating the LF observations x_t^L to the unobservable variables $x_{m,t}^{L*}$, can be described using the cumulator $x_{m,t}^{L*,c}$ defined by the process:

$$x_{m,t}^{L*,c} = a_m x_{m-1,t}^{L*,c} + b_m x_{m,t}^{L*}, \qquad x_{0,t}^{L*,c} \equiv 0,$$

and letting $x_t^L = x_{M,t}^{L*,c}$ for $m = M$, while $x_{m,t}^{L*,c}$ is not observed otherwise. A similar representation is used by Harvey (1989) and Nunes (2005), among others, and includes both stock and flow aggregation as special cases. More specifically, stock-sampling corresponds to the case $a_m = 1(m \neq M)$ and $b_m = 1(m = M)$, where $1(.)$ denotes the indicator function. Flow sampling can be represented setting $a_m = 1(m \neq 1)$ and $b_m = 1/M$ for all m. As the aggregation scheme is linear, it is straightforward to show that applying it to the HF observables and stacking them together with the LF ones, a representation analogous to the one in equation (5.3) is obtained. Then, the identification of the loadings, the aggregated factors, and the common and high frequency factors, follows as in the cases of flow and stock-sampling.

D.4 Uniform asymptotic expansions of factor values and factor loadings in the group factor model

We derive uniform asymptotic expansions for the estimators of the factor values and factor loadings in Definitions 1 and 2 and equations (3.3) and (3.4), up to terms $o_p(\bar{N}^{-1/2})$, where $\bar{N} := \max\{N_1, T\}$.

PROPOSITION D.4. *i) Under Assumption A.1 with $\mu > 0$, and Assumptions A.2-A.4, A.5 b)-c), A.6 a), A.7, A.8, the asymptotic expansions of the factors estimators are given by:*

$$\hat{f}_t^c \;=\; \hat{\mathcal{H}}_c^{-1}\left[f_t^c + \frac{1}{\sqrt{N_1}}u_{1,t}^{(c)} + \frac{1}{T}\beta_{1,t}^c\right] + o_p\left(\bar{N}^{-1/2}\right), \tag{D.19}$$

and:

$$\hat{f}_{j,t}^s = \hat{\mathcal{H}}_{s,j}^{-1}\left[\tilde{f}_{j,t}^s + \frac{1}{\sqrt{N_j}}u_{j,t}^{(s)} + \frac{1}{T}\beta_{j,t}^s\right] + o_p(\bar{N}^{-1/2}), \qquad j = 1, 2, \tag{D.20}$$

where $\tilde{f}_{j,t}^s = f_{j,t}^s - \tilde{\Sigma}_{j,c}\tilde{\Sigma}_{cc}^{-1}f_t^c$ and the o_p terms are uniform w.r.t. $1 \le t \le T$. The asymptotic expansions of the loadings estimators are:

$$\hat{\lambda}_{j,i}^c = \hat{\mathcal{H}}_c'\left[\lambda_{j,i}^c + \tilde{\Sigma}_{cc}^{-1}\tilde{\Sigma}_{c,j}\lambda_{j,i}^s + \frac{1}{\sqrt{T}}w_{j,i}^c + \frac{1}{T}\beta_{\Lambda,j,i}^c\right] + o_p\left(\bar{N}^{-1/2}\right), \qquad j \;=\; 1,\; 2, \tag{D.21}$$

and:

$$\hat{\lambda}_{j,i}^s = \hat{\mathcal{H}}_{s,j}'\left[\lambda_{j,i}^s + \frac{1}{\sqrt{T}}w_{j,i}^s + \frac{1}{T}\beta_{\Lambda,j,i}^s\right] + o_p\left(\bar{N}^{-1/2}\right), \qquad j = 1, 2, \tag{D.22}$$

where the o_p terms are uniform w.r.t. $1 \le i \le N_j$. Matrices $\hat{\mathcal{H}}_c$ and $\hat{\mathcal{H}}_{s,j}$ are such that:

$$\hat{\mathcal{H}}_c\hat{\mathcal{H}}_c' = \tilde{\Sigma}_{cc} + o_p(\bar{N}^{-1/2}), \qquad \hat{\mathcal{H}}_{s,j}\hat{\mathcal{H}}_{s,j}' = (\frac{1}{T}\tilde{F}_j^{s\,\prime}\tilde{F}_j^s)^{-1} + o_p(\bar{N}^{-1/2}), \quad j = 1, 2, \tag{D.23}$$

where $\tilde{F}_j^s = [\tilde{f}_{j,1}^s, ..., \tilde{f}_{j,T}^s]'$. Vector $u_{j,t}$ is defined in Proposition 3, and $w_{j,i}^c = \tilde{\Sigma}_{cc}^{-1}\frac{1}{\sqrt{T}}\sum_{t=1}^T f_t^c\varepsilon_{j,i,t}$ and $w_{j,i}^s = (\frac{1}{T}\tilde{F}_j^{s\,\prime}\tilde{F}_j^s)^{-1}\frac{1}{\sqrt{T}}\sum_{t=1}^T \tilde{f}_{j,t}^s\varepsilon_{j,i,t}$. The bias terms are such that:

$$\begin{aligned}
\beta_{1,t}^c &= \bar{b}_{1,t}^{(c)} - E[(\bar{b}_{1,t}^{(c)} - \bar{b}_{2,t}^{(c)})(f_{1,t}^s - \Phi f_{2,t}^s)'](I_{k_1-k^c} - \Phi\Phi')^{-1}f_{1,t}^s, \\
\beta_{j,t}^s &= [\Sigma_{\lambda,j}^{(ss)}]^{-1}\left(\eta_{j,t}^2 f_{j,t}^s - \Sigma_{\lambda,j}^{(ss)}E[f_{j,t}^s\beta_{j,t}^{c\,\prime}]f_t^c - \Sigma_{\lambda,j}^{(sc)}\tilde{\beta}_{j,t}^c\right), \\
\beta_{\Lambda,j,i}^c &= E[\beta_{1,t}^c f_{j,t}^{s\,\prime}]\lambda_{j,i}^s - E[f_t^c\beta_{1,t}^{c\,\prime}]\lambda_{j,i}^c, \\
\beta_{\Lambda,j,i}^s &= -E[f_{j,t}^s\beta_{j,t}^{c\,\prime}]\lambda_{j,i}^c - E[f_{j,t}^s\beta_{j,t}^{s\,\prime}]\lambda_{j,i}^s,
\end{aligned}$$

where vector $\bar{b}_{j,t}$ is defined in Theorem 1, and $\tilde{\beta}_{j,t}^c = \beta_{j,t}^c - E[\beta_{j,t}^c f_t^{c\,\prime}]f_t^c$.
ii) If we relax the regularity conditions to allow for $\mu \ge 0$ in Assumption A.1, the asymptotic expansions in (D.19)-(D.23) hold with remainder term of uniform order $o_p(T^{-1/2})$ (and bias terms of order T^{-1} absorbed into the remainder term).

In the asymptotic expansion of $\hat{f}_t^c$, the stochastic term at order $N_1^{-1/2}$ comes from the estimation of the principal components in the first subgroup. The bias at order T^{-1} consists both of a term arising from principal components estimation, namely $T^{-1}\bar{b}_{1,t}^{(c)}$, and another term induced by estimation of the canonical directions associated with the unit canonical correlations. Vector $\hat{f}_{j,t}^s$ estimates the residual of the sample projection of $f_{j,t}$ onto f_t^c at rate $N_j^{-1/2}$. The bias at order T^{-1} is induced by both extraction of PC from the panel of residuals and the bias

in estimator $\hat{f}_t^c$.
In the asymptotic expansion of $\hat{\lambda}_{j,i}^c$, the term $\tilde{\Sigma}_{cc}^{-1}\tilde{\Sigma}_{c,j}\lambda_{j,i}^s$ is induced by the fact that the common and frequency-specific factors are not orthogonal in-sample. The deterministic term at order T^{-1} is induced by the bias in the common factor estimates. The expansion of $\hat{\lambda}_{j,i}^c$ does not contain explicitly a bias component at order N_j^{-1}, since $N_j^{-1} = o_p(\bar{N}^{-1/2})$ under Assumption A.1.
The uniform asymptotic expansions at order $o_p(T^{-1/2})$ in Proposition D.4 ii) suffice for the proof of Theorem 2. We need the more accurate expansions at order $o_p(\bar{N}^{-1/2})$ in the proof of Proposition D.7, where the error-in-variable from estimation of factor loadings has to be controlled.

D.4.1 Proof of Proposition D.4

We start by providing some uniform bounds in Subsection D.4.1 a), that are instrumental for the rest of the proof of Proposition D.4. Then, in Subsections D.4.1 b)-e) we establish the uniform asymptotic expansions of factors and loadings up to order $o_p(\bar{N}^{-1/2})$, where $\bar{N} = \max\{N_1, T\}$ (proof of part i)). Finally, in Subsection D.4.1 f) we show how to get the uniform asymptotic expansions up to order $o_p(T^{-1/2})$ under a less restrictive asymptotic scheme (proof of part ii)).

a) Uniform bounds

Let $X = O_{p,\ell}(a_{N,T})$ mean $X = O_p[a_{N,T}(\log T)^{\bar{b}}]$ for some $\bar{b} > 0$. Under Assumption A.8 we have the following uniform bounds, which complement those in Lemma C.10:

$$\sup_{1\leq t\leq T} \|b_{j,t}\| = O_{p,\ell}(1), \tag{D.24}$$
$$\sup_{1\leq t\leq T} \|d_{j,t}\| = O_{p,\ell}(1), \tag{D.25}$$
$$\sup_{1\leq t\leq T} \|\hat{h}_{j,t}\| = O_{p,\ell}(1), \tag{D.26}$$
$$\sup_{1\leq t\leq T} \|\beta_{j,t}^c\| = O_{p,\ell}(1), \tag{D.27}$$
$$\sup_{1\leq i\leq N_j} \|\frac{1}{T}\sum_{t=1}^{T} \beta_{j,t}^c \varepsilon_{j,i,t}\| = O_{p,\ell}(T^{-\eta/2}), \tag{D.28}$$
$$\sup_{1\leq i\leq N_j} \|\frac{1}{T}\sum_{t=1}^{T} \varepsilon_{j,i,t}^2\| = O_p(1), \tag{D.29}$$
$$\sup_{1\leq i\leq N_j} \frac{1}{N_j T}\sum_{\ell=1,\ell\neq i}^{N_j}\sum_{t=1}^{T} \lambda_{j,\ell}\varepsilon_{j,\ell,t}\varepsilon_{j,i,t} = O_{p,\ell}(\frac{1}{\sqrt{NT^\eta}}) + O(\frac{1}{N}), \tag{D.30}$$

where $\eta \geq 1/2$. We prove below the uniform bound in (D.30). The proofs of the other ones follow by similar arguments.
Proof of (D.30). We have:

$$\frac{1}{N_j T}\sum_{\ell=1,\ell\neq i}^{N_j}\sum_{t=1}^{T} \lambda_{j,\ell}\varepsilon_{j,\ell,t}\varepsilon_{j,i,t} = \frac{1}{\sqrt{N_j}}\left[\frac{1}{T}\sum_{t=1}^{T}\left(\frac{1}{\sqrt{N_j}}\sum_{\ell=1,\ell\neq i}^{N_j} \varepsilon_{j,\ell,t}\varepsilon_{j,i,t} - E[\frac{1}{\sqrt{N_j}}\sum_{\ell=1,\ell\neq i}^{N_j} \lambda_{j,\ell}\varepsilon_{j,\ell,t}\varepsilon_{j,i,t}]\right)\right] + \frac{1}{N_j}\sum_{\ell=1,\ell\neq i}^{N_j} \lambda_{j,\ell}E[\varepsilon_{j,\ell,t}\varepsilon_{j,i,t}].$$

From Assumption A.8 c) we have $\frac{1}{T}\sum_{t=1}^{T}\left(\frac{1}{\sqrt{N_j}}\sum_{\ell=1,\ell\neq i}^{N_j}\varepsilon_{j,\ell,t}\varepsilon_{j,i,t} - E[\frac{1}{\sqrt{N_j}}\sum_{\ell=1,\ell\neq i}^{N_j}\lambda_{j,\ell}\varepsilon_{j,\ell,t}\varepsilon_{j,i,t}]\right) = O_{p,\ell}(T^{-\eta/2})$, uniformly in $1 \leq i \leq N_j$, similarly as in the proof of (C.107). From Assumptions A.8 b) and d) we have $\sum_{\ell=1,\ell\neq i}^{N_j}\lambda_{j,\ell}E[\varepsilon_{j,\ell,t}\varepsilon_{j,i,t}] = O(1)$, uniformly in $1 \leq i \leq N_j$. Then, (D.30) follows.

b) Asymptotic expansion of $\hat{f}_t^c$

Let us start by establishing the asymptotic expansion of $\hat{f}_t^c$ up to order $o_p(\bar{N}^{-1/2})$. Equation (B.12) and $\hat{\Psi} = O_p(\delta_{N,T})$ imply $\hat{W}_1^* = [E_c + E_s(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\hat{\Psi}_{sc}^{(I)}]\hat{\mathcal{U}} + O_p(\delta_{N,T}^2)$. 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:

$$\begin{aligned}
\hat{f}_t^c &= \hat{W}_1'\hat{h}_{1,t} = \hat{D}\hat{\mathcal{U}}'\left[E_c'\hat{h}_{1,t} + \hat{\Psi}_{sc}^{(I)\prime}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}E_s'\hat{h}_{1,t}\right] + O_{p,l}\left(\delta_{N,T}^2\right) \\
&= \hat{D}\hat{\mathcal{U}}'\left[f_t^c + \frac{1}{\sqrt{N_1}}u_{1,t}^{(c)} + \frac{1}{T}b_{1,t}^{(c)} + \frac{1}{\sqrt{N_1 T}}d_{1,t}^{(c)} + \vartheta_{1,t}^{(c)}\right. \\
&\quad \left. + \hat{\Psi}_{sc}^{(I)\prime}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}\left(f_{1,t}^s + \frac{1}{\sqrt{N_1}}u_{1,t}^{(s)} + \frac{1}{T}b_{1,t}^{(s)} + \frac{1}{\sqrt{N_1 T}}d_{1,t}^{(s)} + \vartheta_{1,t}^{(s)}\right)\right] + O_{p,l}\left(\delta_{N,T}^2\right), \qquad \text{(D.31)}
\end{aligned}$$

uniformly in $1 \leq t \leq T$, where we use the expansion of the factor estimates in Proposition 3, and (D.26). Under Assumption A.1 with $\mu > 0$, $N = N_2$ and N_1 grow at the same rate such that $T^{1/2} \ll N \ll T^{5/2}$. Therefore, $(\log T)^{\bar{b}}\delta_{N,T}^2 = o(\bar{N}^{-1/2})$, for any $\bar{b} > 0$, $\frac{1}{\sqrt{N_1}}\delta_{N,T} = o(\bar{N}^{-1/2})$ and $\frac{1}{T}\delta_{N,T} = o(\bar{N}^{-1/2})$ under Assumption A.1 with $\mu > 0$. By using uniform bounds in Lemma C.10 and (D.24)-(D.25), and keeping only terms up to $o_p(\bar{N}^{-1/2})$, we get:

$$\hat{f}_t^c = \hat{\mathcal{H}}_c^{-1}\left[f_t^{(c)} + \frac{1}{\sqrt{N_1}}u_{1,t}^{(c)} + \frac{1}{T}\bar{b}_{1,t}^{(c)} + \hat{\Psi}_{sc}^{(I)\prime}(I_{k_1-k^c} - \tilde{R}_{ss})^{-1}f_{1,t}^s\right] + o_p\left(\bar{N}^{-1/2}\right), \qquad \text{(D.32)}$$

uniformly in $1 \leq t \leq T$, where $\hat{\mathcal{H}}_c^{-1} = \hat{D}\hat{\mathcal{U}}'$.

To further develop this asymptotic expansion, we need the asymptotic behavior of $\hat{\Psi}_{sc}^{(I)}$. From equation $\hat{\Psi} = \tilde{V}_{11}^{-1}\hat{\Psi}^*$ (see Lemma B.2) we have $\hat{\Psi}_{sc}^{(I)} = (\tilde{V}_{11}^{-1})_{sc}\hat{\Psi}_{cc}^{*(I)} + (\tilde{V}_{11}^{-1})_{ss}\hat{\Psi}_{sc}^{*(I)}$. From Lemma C.7, we have $\hat{\Psi}_{cc}^{*(I)} = O_p\left(\frac{1}{N} + \frac{1}{T^2} + \frac{1}{T\sqrt{NT}}\right) = o_p(\bar{N}^{-1/2})$ under Assumption A.1 with $\mu > 0$. Moreover, from (B.6) and Lemma B.3 we get:

$$\hat{\Psi}_{sc}^{*(I)} = -(\hat{X}_{11,sc} - \hat{X}_{12,sc}) + \tilde{B}_{cs}'(\hat{X}_{21,cc} - \hat{X}_{22,cc}) + \tilde{B}_{ss}'(\hat{X}_{21,sc} - \hat{X}_{22,sc}).$$

From Lemmas B.1 and B.3, and equation (C.74), the second term in the r.h.s. is $O_p(T^{-1/2}\delta_{N,T}) = o_p(\bar{N}^{-1/2})$ under Assumption A.1 with $\mu > 0$. Now, we substitute in the definitions of terms $\hat{X}_{j,k}$ from (B.3), and use that $\frac{1}{T}\sum_{t=1}^{T}\psi_{j,t}\psi_{k,t}' = o_p(\bar{N}^{-1/2})$. We get:

$$\hat{\Psi}_{sc}^{*(I)} = -\frac{1}{T}\sum_{t=1}^{T}(f_{1,t}^s - \tilde{B}_{ss}'f_{2,t}^s)[\psi_{1,t}^{(c)} - \psi_{2,t}^{(c)}]' + o_p(\bar{N}^{-1/2}).$$

By using the definition of $\psi_{j,t}$, $\tilde{B}_{ss} = \Phi' + O_p(T^{-1/2})$, and keeping terms up to $o_p(\bar{N}^{-1/2})$, we get:

$$\begin{aligned}
\hat{\Psi}_{sc}^{*(I)} &= -\frac{1}{T}\left(\frac{1}{T}\sum_{t=1}^{T}(f_{1,t}^{s} - \Phi f_{2,t}^{s})[\bar{b}_{1,t}^{(c)} - \bar{b}_{2,t}^{(c)}]'\right) + o_p(\bar{N}^{-1/2}) \\
&= -\frac{1}{T}E[(f_{1,t}^{s} - \Phi f_{2,t}^{s})(\bar{b}_{1,t}^{(c)} - \bar{b}_{2,t}^{(c)})'] + o_p(\bar{N}^{-1/2}).
\end{aligned}$$

Thus, by using $(\tilde{V}_{11}^{-1})_{ss} = I_{k_1-k^c} + O_p(T^{-1/2})$ and $N \ll T^3$, we get:

$$\hat{\Psi}_{sc}^{(I)} = -\frac{1}{T}E[(f_{1,t}^{s} - \Phi f_{2,t}^{s})(\bar{b}_{1,t}^{(c)} - \bar{b}_{2,t}^{(c)})'] + o_p(\bar{N}^{-1/2}). \tag{D.33}$$

Thus, from (D.32) and (D.33), and by using $(I_{k_1-k^c} - \tilde{R}_{ss})^{-1} = (I_{k_1-k^c} - \Phi\Phi')^{-1} + O_p(T^{-1/2})$ and $N \ll T^3$, we get:

$$\hat{f}_t^c = \hat{\mathcal{H}}_c^{-1}\left[f_t^{(c)} + \frac{1}{\sqrt{N_1}}u_{1,t}^{(c)} + \frac{1}{T}\beta_{1,t}^{c}\right] + o_p\left(\bar{N}^{-1/2}\right), \tag{D.34}$$

uniformly in $1 \leq t \leq T$, where:

$$\beta_{1,t}^{c} = \bar{b}_{1,t}^{(c)} - E[(\bar{b}_{1,t}^{(c)} - \bar{b}_{2,t}^{(c)})(f_{1,t}^{s} - \Phi f_{2,t}^{s})'](I_{k_1-k^c} - \Phi\Phi')^{-1}f_{1,t}^{s},$$

which yields (D.19).
The asymptotic expansion for estimator $\hat{f}_t^{c\,*}$ is obtained by interchanging the roles of panels $j = 1$ and $j = 2$. Hence,

$$\hat{f}_t^{c\,*} = \hat{\mathcal{H}}_{*c}^{-1}\left[f_t^{(c)} + \frac{1}{\sqrt{N_2}}u_{2,t}^{(c)} + \frac{1}{T}\beta_{2,t}^{c}\right] + o_p\left(\bar{N}^{-1/2}\right),$$

uniformly in $1 \leq t \leq T$, where:

$$\beta_{2,t}^{c} = \bar{b}_{2,t}^{(c)} - E[(\bar{b}_{2,t}^{(c)} - \bar{b}_{1,t}^{(c)})(f_{2,t}^{s} - \Phi' f_{1,t}^{s})'](I_{k_2-k^c} - \Phi'\Phi)^{-1}f_{2,t}^{s}.$$

Finally, let us show the asymptotic expansion for $\hat{\mathcal{H}}_c\hat{\mathcal{H}}_c'$. Substituting the expression of $\hat{f}_t^c$ from equation (D.34) into the equality $\frac{1}{T}\sum_{t=1}^{T}\hat{f}_t^c\hat{f}_t^{c\prime} = I_{k^c}$, we get:

$$\begin{aligned}
I_{k^c} &= \hat{\mathcal{H}}_c^{-1}\frac{1}{T}\sum_{t=1}^{T}\left(f_t^c + \frac{1}{\sqrt{N_1}}u_{1,t}^{(c)} + \frac{1}{T}\beta_{1,t}^{c}\right)\left(f_t^c + \frac{1}{\sqrt{N_1}}u_{1,t}^{(c)} + \frac{1}{T}\beta_{1,t}^{c}\right)'\left(\hat{\mathcal{H}}_c^{-1}\right)' + o_p\left(\bar{N}^{-1/2}\right) \\
&= \hat{\mathcal{H}}_c^{-1}\tilde{\Sigma}_{cc}\left(\hat{\mathcal{H}}_c^{-1}\right)' + o_p\left(\bar{N}^{-1/2}\right),
\end{aligned} \tag{D.35}$$

using arguments similar to the proof of Lemma B.1 and Assumption A.1 with $\mu > 0$. Thus, we get $\hat{\mathcal{H}}_c\hat{\mathcal{H}}_c' = \tilde{\Sigma}_{cc} + o_p\left(\bar{N}^{-1/2}\right)$, which yields the first equation in (D.23). By using (C.74) it follows:

$$\hat{\mathcal{H}}_c\hat{\mathcal{H}}_c' = I_{k^c} + O_p(T^{-1/2}). \tag{D.36}$$

c) Asymptotic expansion of $\hat{\lambda}_{j,i}^c$

Let us now derive the asymptotic expansion of the loading estimator $\hat{\lambda}_{j,i}^c = \hat{F}^{c\,\prime}y_{j,i}/T$ up to order $o_p\left(\bar{N}^{-1/2}\right)$, where $y_{j,i}$ is the i-th column of matrix Y_j and $\hat{F}^c = [\hat{f}_1^c, ..., \hat{f}_T^c]'$. From equation (D.34) we have $\hat{F}^c =$

$\left(F^c + \frac{1}{\sqrt{N_1}} U_1^c + \frac{1}{T} B_1^c\right) \left(\hat{\mathcal{H}}_c^{-1}\right)' + o_p\left(\bar{N}^{-1/2}\right)$, where $U_1^c = [u_{1,1}^{(c)}, ..., u_{1,T}^{(c)}]'$ and $B_1^c = [\beta_{1,1}^c, ..., \beta_{1,T}^c]'$, which implies:

$$\hat{F}^c \hat{\mathcal{H}}_c' - F^c = \frac{1}{\sqrt{N_1}} U_1^c + \frac{1}{T} B_1^c + o_p\left(\bar{N}^{-1/2}\right). \tag{D.37}$$

Here $o_p\left(\bar{N}^{-1/2}\right)$ denotes a matrix whose rows are uniformly of stochastic order $o_p\left(\bar{N}^{-1/2}\right)$. Then:

$$\begin{aligned}
\hat{\lambda}_{j,i}^c &= \frac{1}{T}\hat{F}^{c\,\prime} y_{j,i} = \frac{1}{T}\hat{F}^{c\,\prime}\left(F^c \lambda_{j,i}^c + F_j^s \lambda_{j,i}^s + \varepsilon_{j,i}\right) \\
&= \frac{1}{T}\hat{F}^{c\,\prime}\left(\left[\hat{F}^c \hat{\mathcal{H}}_c' - \left(\hat{F}^c \hat{\mathcal{H}}_c' - F^c\right)\right]\lambda_{j,i}^c + F_j^s \lambda_{j,i}^s + \varepsilon_{j,i}\right) \\
&= \hat{\mathcal{H}}_c' \lambda_{j,i}^c - \frac{1}{T}\hat{F}^{c\,\prime}\left(\hat{F}^c \hat{\mathcal{H}}_c' - F^c\right)\lambda_{j,i}^c + \frac{1}{T}\hat{F}^{c\,\prime} F_j^s \lambda_{j,i}^s + \frac{1}{T}\hat{F}^{c\,\prime}\varepsilon_{j,i}, \qquad j = 1, 2.
\end{aligned}$$

By writing $\hat{F}^c = \left[F^c + (\hat{F}^c \hat{\mathcal{H}}_c' - F^c)\right](\hat{\mathcal{H}}_c')^{-1}$, and rearranging terms, we get:

$$\begin{aligned}
\hat{\lambda}_{j,i}^c &= \hat{\mathcal{H}}_c' \Bigg\{\lambda_{j,i}^c + (\hat{\mathcal{H}}_c')^{-1}(\hat{\mathcal{H}}_c)^{-1}\frac{1}{T}F^{c\,\prime}\varepsilon_{j,i} + (\hat{\mathcal{H}}_c')^{-1}(\hat{\mathcal{H}}_c)^{-1}\frac{1}{T}F^{c\,\prime}F_j^s \lambda_{j,i}^s \\
&\quad + (\hat{\mathcal{H}}_c')^{-1}(\hat{\mathcal{H}}_c)^{-1}\frac{1}{T}(\hat{F}^c \hat{\mathcal{H}}_c' - F^c)'\varepsilon_{j,i} + (\hat{\mathcal{H}}_c')^{-1}(\hat{\mathcal{H}}_c)^{-1}\frac{1}{T}(\hat{F}^c \hat{\mathcal{H}}_c' - F^c)' F_j^s \lambda_{j,i}^s \\
&\quad - (\hat{\mathcal{H}}_c')^{-1}(\hat{\mathcal{H}}_c)^{-1}\frac{1}{T}\left[F^c + (\hat{F}^c \hat{\mathcal{H}}_c' - F^c)\right]'\left(\hat{F}^c \hat{\mathcal{H}}_c' - F^c\right)\lambda_{j,i}^c\Bigg\}.
\end{aligned} \tag{D.38}$$

We use equation (D.37) to bound the different terms. We have:

$$\begin{aligned}
\frac{1}{T}(\hat{F}^c \hat{\mathcal{H}}_c' - F^c)'\varepsilon_{1,i} &= \frac{1}{\sqrt{N_1}T} U_1^{c\,\prime}\varepsilon_{1,i} + \frac{1}{T^2}B_1^{c\,\prime}\varepsilon_{1,i} + o_p(\bar{N}^{-1/2}) \\
&= (\Lambda_1'\Lambda_1/N_1)^{-1}\frac{1}{N_1 T}\sum_{\ell=1}^{N_1}\sum_{t=1}^{T}\lambda_{1,\ell}\varepsilon_{1,\ell,t}\varepsilon_{1,i,t} + o_p(\bar{N}^{-1/2}) \\
&= (\Lambda_1'\Lambda_1/N_1)^{-1}\frac{1}{N_1 T}\sum_{t=1}^{T}\lambda_{1,i}\varepsilon_{1,i,t}^2 + (\Lambda_1'\Lambda_1/N_1)^{-1}\frac{1}{N_1 T}\sum_{\ell=1,\ell\neq i}^{N_1}\sum_{t=1}^{T}\lambda_{1,\ell}\varepsilon_{1,\ell,t}\varepsilon_{1,i,t} \\
&\quad + o_p(\bar{N}^{-1/2}) = O_p(N_1^{-1}) + O_{p,l}[(N_1 T^{\eta})^{-1/2}] + o_p(\bar{N}^{-1/2}),
\end{aligned}$$

uniformly in $1 \leq i \leq N_1$, using bounds (D.28)-(D.29) and Assumption A.8 *d)*. A similar bound holds for $j = 2$. Since N_1 grows at the same rate as N and $T^{1/2} \ll N$, we have $N_1^{-1} = o(\bar{N}^{-1/2})$. Moreover, from $\eta \geq 1/2$ and $T^{1/2} \ll N$, we have $O_{p,l}[(N_1 T^{\eta})^{-1/2}] = o_p(\bar{N}^{-1/2})$. Hence, $\frac{1}{T}(\hat{F}^c \hat{\mathcal{H}}_c' - F^c)'\varepsilon_{j,i} = o_p(\bar{N}^{-1/2})$, uniformly in $1 \leq i \leq N_1$. Moreover:

$$\begin{aligned}
\frac{1}{T}(\hat{F}^c \hat{\mathcal{H}}_c' - F^c)' F_j^s &= \frac{1}{T\sqrt{N_1}} U_1^{c\,\prime} F_j^s + \frac{1}{T^2} B_1^{c\,\prime} F_j^s + o_p(\bar{N}^{-1/2}) \\
&= \frac{1}{T}E[\beta_{1,t}^c f_{j,t}^{s\,\prime}] + O_p((N_1 T)^{-1/2}) + o_p(\bar{N}^{-1/2}) = \frac{1}{T}E[\beta_{1,t}^c f_{j,t}^{s\,\prime}] + o_p(\bar{N}^{-1/2}),
\end{aligned}$$

and:

$$
\begin{aligned}
& \frac{1}{T}\left[F^c + (\hat{F}^c\hat{\mathcal{H}}_c' - F^c)\right]'\left(\hat{F}^c\hat{\mathcal{H}}_c' - F^c\right) \\
= \;& \frac{1}{T\sqrt{N_1}}F^{c\,\prime}U_1^c + \frac{1}{T^2}F^{c\,\prime}B_1^c + \frac{1}{N_1 T}U_1^{c\,\prime}U_1^c + \frac{1}{T^2\sqrt{N_1}}(U_1^{c\,\prime}B_1^c + B_1^{c\,\prime}U_1^c) + \frac{1}{T^3}B_1^{c\,\prime}B_1^c + o_p(\bar{N}^{-1/2}) \\
= \;& \frac{1}{T}E[f_t^c\beta_{1,t}^{c\,\prime}] + O_p((N_1T)^{-1/2} + N_1^{-1}) + o_p(\bar{N}^{-1/2}) = \frac{1}{T}E[f_t^c\beta_{1,t}^{c\,\prime}] + o_p(\bar{N}^{-1/2}).
\end{aligned}
$$

Further, from (D.36) we have $(\hat{\mathcal{H}}_c')^{-1}(\hat{\mathcal{H}}_c)^{-1} = (\hat{\mathcal{H}}_c\hat{\mathcal{H}}_c')^{-1} = \tilde{\Sigma}_{cc}^{-1} + o_p(\bar{N}^{-1/2}) = I_{k^c} + O_p(T^{-1/2})$. Then, from (D.38) and Assumption A.8 *d)* we get:

$$
\hat{\lambda}_{j,i}^c \;=\; \hat{\mathcal{H}}_c'\left[\lambda_{j,i}^c + \tilde{\Sigma}_{cc}^{-1}\frac{1}{T}F^{c\,\prime}\varepsilon_{j,i} + \tilde{\Sigma}_{cc}^{-1}\frac{1}{T}F^{c\,\prime}F_j^s\lambda_{j,i}^s + \frac{1}{T}\left(E[\beta_{1,t}^c f_{j,t}^{s\,\prime}]\lambda_{j,i}^s - E[f_t^c\beta_{1,t}^{c\,\prime}]\lambda_{j,i}^c\right)\right] + o_p(\bar{N}^{-1/2}),
$$

uniformly in $1 \le i \le N_j$. The last equation can be rewritten as

$$
\hat{\lambda}_{j,i}^c = \hat{\mathcal{H}}_c'\left[\lambda_{j,i}^c + \tilde{\Sigma}_{cc}^{-1}\tilde{\Sigma}_{c,j}\lambda_{j,i}^s + \frac{1}{\sqrt{T}}w_{j,i}^c + \frac{1}{T}\beta_{\Lambda,j,i}^c\right] + o_p\left(\bar{N}^{-1/2}\right), \qquad j \;=\; 1,\; 2, \tag{D.39}
$$

where:

$$
\begin{aligned}
w_{j,i}^c \;&:=\; \tilde{\Sigma}_{cc}^{-1}\frac{1}{\sqrt{T}}F^{c\,\prime}\varepsilon_{j,i} = \tilde{\Sigma}_{cc}^{-1}\frac{1}{\sqrt{T}}\sum_{t=1}^T f_t^c\varepsilon_{j,i,t}, \\
\tilde{\Sigma}_{cc} \;&=\; \frac{1}{T}F^{c\,\prime}F^c = \frac{1}{T}\sum_{t=1}^T f_t^c f_t^{c\,\prime}, \qquad \tilde{\Sigma}_{c,j} = \frac{1}{T}F^{c\,\prime}F_j^s = \frac{1}{T}\sum_{t=1}^T f_t^c f_{j,t}^{s\prime}, \\
\beta_{\Lambda,j,i}^c \;&=\; E[\beta_{1,t}^c f_{j,t}^{s\,\prime}]\lambda_{j,i}^s - E[f_t^c\beta_{1,t}^{c\,\prime}]\lambda_{j,i}^c.
\end{aligned}
$$

If we use $\hat{f}_t^{c\,*}$ for group $j = 2$, the bias is:

$$
\beta_{\Lambda,j,i}^c = E[\beta_{j,t}^c f_{j,t}^{s\,\prime}]\lambda_{j,i}^s - E[f_t^c\beta_{j,t}^{c\,\prime}]\lambda_{j,i}^c, \tag{D.40}
$$

where $\beta_{j,i}^c$ is the bias at order T^{-1} of $\hat{f}_t^{c\,*}$.

d) Asymptotic expansion of $\hat{f}_{j,t}^s$

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 $\xi_{j,i,t} := y_{j,i,t} - \hat{f}_t^{c\,\prime}\hat{\lambda}_{j,i}^c$, where we replace $\hat{f}_t^c$ with $\hat{f}_t^{c\,*}$ for $j = 2$. By substituting the asymptotic expansions in equations (D.34) and (D.39), have:

$$
\begin{aligned}
\xi_{j,i,t} \;&=\; f_{j,t}^{s\,\prime}\lambda_{j,i}^s + \varepsilon_{j,i,t} - \left(\hat{f}_t^{c\,\prime}\hat{\lambda}_{j,i}^c - f_t^{c\,\prime}\lambda_{j,i}^c\right) \\
&=\; f_{j,t}^{s\,\prime}\lambda_{j,i}^s + \varepsilon_{j,i,t} \\
&\quad -\left[\left(f_t^c + \frac{1}{\sqrt{N_j}}u_{j,t}^{(c)} + \frac{1}{T}\beta_{j,t}^c\right)'\left(\lambda_{j,i}^c + \tilde{\Sigma}_{cc}^{-1}\tilde{\Sigma}_{c,j}\lambda_{j,i}^s + \frac{1}{\sqrt{T}}w_{j,i}^c + \frac{1}{T}\beta_{\Lambda,j,i}^c\right) - f_t^{c\,\prime}\lambda_{j,i}^c\right] + o_p(\bar{N}^{-1}) \\
&=\; \tilde{f}_{j,t}^{s\,\prime}\lambda_{j,i}^s + e_{j,i,t} + o_p(\bar{N}^{-1}),
\end{aligned} \tag{D.41}
$$

where we define:

$$
\begin{aligned}
\tilde{f}^s_{j,t} &:= f^s_{j,t} - \tilde{\Sigma}_{j,c}\tilde{\Sigma}_{cc}^{-1} f^c_t, \\
e_{j,i,t} &:= \varepsilon_{j,i,t} - \frac{1}{\sqrt{T}} f^{c\prime}_t w^c_{j,i} - \frac{1}{\sqrt{N_j}} u^{(c)\,\prime}_{j,t} \lambda^c_{j,i} - \frac{1}{T}\left(f^{c\,\prime}_t \beta^c_{\Lambda,j,i} + \beta^{c\,\prime}_{j,t}\lambda^c_{j,i}\right).
\end{aligned}
$$

The term $o_p(\bar{N}^{-1})$ is uniform in $i = 1, ..., N_j$ and $t = 1, ..., T$ by bounds (C.105)-(C.106) and (D.26)-(D.27), and Assumption A.8 *d)*. Then, the residuals $\xi_{j,i,t}$, with $i = 1, ..., N_j$ and $t = 1, ..., T$, satisfy an approximate factor structure with factors $\tilde{f}^s_{j,t}$, loadings $\lambda^s_{j,i}$ and errors $e_{j,i,t}$, up to $o_p(\bar{N}^{-1/2})$. The error terms contain a factor structure at order T^{-1}.

From the asymptotic theory of the PC estimators in large panels we have an asymptotic expansion as that in Proposition 3:

$$
\hat{f}^s_{j,t} = \hat{\mathcal{H}}^{-1}_{s,j}\left[\tilde{f}^s_{j,t} + \frac{1}{\sqrt{N_j}} v^{*\,s}_{j,t} + \frac{1}{T} b^{*\,s}_{j,t} + \frac{1}{\sqrt{N_j T}} d^{*\,s}_{j,t} + \vartheta^{*\,s}_{j,t}\right], \qquad j = 1, 2, \tag{D.42}
$$

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

$$
\begin{aligned}
v^{*\,s}_{j,t} &= \left(\frac{1}{N_j}\Lambda^{s\prime}_j \Lambda^s_j\right)^{-1} \frac{1}{\sqrt{N_j}} \Lambda^{s\prime}_j e_{j,t} \\
b^{*\,s}_{j,t} &= \left(\frac{1}{N_j}\Lambda^{s\,\prime}_j \Lambda^s_j\right)^{-1} \left(\frac{1}{T}\tilde{F}^{s\,\prime}_j \tilde{F}^s_j\right)^{-1} (\eta^*_{j,t})^2 \tilde{f}^s_{j,t}, \\
d^{*\,s}_{j,t} &= \left(\frac{1}{N_j}\Lambda^{s\,\prime}_j \Lambda^s_j\right)^{-1} \left(\frac{1}{T}\tilde{F}^{s\,\prime}_j \tilde{F}^s_j\right)^{-1} \left(\frac{1}{N_j T}\sum_{i=1}^{N_j}\sum_{r=1}^{T} e_{j,i,r}\tilde{f}^s_{j,r}\lambda^{s\,\prime}_{j,i}\right)\tilde{f}^s_{j,t},
\end{aligned}
$$

where $(\eta^*_{j,t})^2 = \underset{N_j\to\infty}{plim} \frac{1}{N_j}\sum_{i=1}^{N_j} E[e^2_{j,i,t}|\mathcal{F}_t]$ and $\tilde{F}^s_j$ denotes the matrix with rows $\tilde{f}^{s\,\prime}_{j,t}$. We have

$$
\begin{aligned}
\frac{1}{N_j}\Lambda^{s\,\prime}_j e_{j,t} &= \frac{1}{N_j}\sum_{i=1}^{N_j}\lambda^s_{j,i}\varepsilon_{j,i,t} - \frac{1}{\sqrt{T}}\left(\frac{1}{N_j}\sum_{i=1}^{N_j}\lambda^s_{j,i} w^{c\,\prime}_{j,i}\right) f^c_t - \frac{1}{\sqrt{N_j}}\left(\frac{\Lambda^{s\prime}_j\Lambda^c_j}{N_j}\right) u^{(c)}_{j,t} \\
&\quad -\frac{1}{T}\left(\frac{1}{N_j}\sum_{i=1}^{N_j}\lambda^s_{j,i}\beta^{c\,\prime}_{\Lambda,j,i}\right) f^c_t - \frac{1}{T}\left(\frac{1}{N_j}\Lambda^{s\,\prime}_j\Lambda^c_j\right)\beta^c_{j,t}.
\end{aligned}
$$

We have $\frac{1}{N_j}\sum_{i=1}^{N_j}\lambda^s_{j,i} w^{c\,\prime}_{j,i} = O_p(N_j^{-1/2})$, $\frac{1}{N_j}\Lambda^{s\,\prime}_j\Lambda^c_j = \Sigma^{(sc)}_{\lambda,j} + O(N_j^{-1/2})$ and from (D.40):

$$
\begin{aligned}
\frac{1}{N_j}\sum_{i=1}^{N_j}\lambda^s_{j,i}\beta^{c\,\prime}_{\Lambda,j,i} &= \left(\frac{1}{N_j}\Lambda^{s\,\prime}_j\Lambda^s_j\right) E[f^s_{j,t}\beta^{c\,\prime}_{j,t}] - \left(\frac{1}{N_j}\Lambda^{s\,\prime}_j\Lambda^c_j\right) E[\beta^c_{j,t} f^{c\,\prime}_t] \\
&= \Sigma^{(ss)}_{\lambda,j} E[f^s_{j,t}\beta^{c\,\prime}_{j,t}] - \Sigma^{(sc)}_{\lambda,j} E[\beta^c_{j,t} f^{c\,\prime}_t] + O(N_j^{-1/2}).
\end{aligned}
$$

Thus:

$$
\frac{1}{N_j}\Lambda^{s\,\prime}_j e_{j,t} = \frac{1}{N_j}\sum_{i=1}^{N_j}\lambda^s_{j,i}\varepsilon_{j,i,t} - \frac{1}{\sqrt{N_j}}\left(\frac{\Lambda^{s\prime}_j\Lambda^c_j}{N_j}\right) u^{(c)}_{j,t} - \frac{1}{T}\left(\Sigma^{(ss)}_{\lambda,j} E[f^s_{j,t}\beta^{c\,\prime}_{j,t}] f^c_t + \Sigma^{(sc)}_{\lambda,j}\tilde{\beta}^c_{j,t}\right) + o_p(\bar{N}^{-1/2}),
$$

uniformly w.r.t. $t = 1, ..., T$, where $\tilde{\beta}^c_{j,t} := \beta^c_{j,t} - E[\beta^c_{j,t} f^{c\,\prime}_t] f^c_t$ is the residual of the orthogonal projection of $\beta^c_{j,t}$ onto f^c_t, and:

$$\frac{1}{\sqrt{N_j}} v^{*\,s}_{j,t} = \frac{1}{\sqrt{N_j}} v^s_{j,t} - \frac{1}{T}[\Sigma^{(ss)}_{\lambda,j}]^{-1}\left(\Sigma^{(ss)}_{\lambda,j} E[f^s_{j,t}\beta^{c\,\prime}_{j,t}] f^c_t + \Sigma^{(sc)}_{\lambda,j}\tilde{\beta}^c_{j,t}\right) + o_p(\bar{N}^{-1/2}),$$

where $v^s_{j,t} = \left(\frac{\Lambda^{s\prime}_j \Lambda^s_j}{N_j}\right)^{-1} \frac{1}{\sqrt{N_j}} \sum_{i=1}^{N_j} \lambda^s_{j,i}\varepsilon_{j,i,t} - \left(\frac{\Lambda^{s\prime}_j \Lambda^s_j}{N_j}\right)^{-1}\left(\frac{\Lambda^{s\prime}_j \Lambda^c_j}{N_j}\right) u^{(c)}_{j,t}$. Moreover:

$$b^{*\,s}_{j,t} = [\Sigma^{(ss)}_{\lambda,j}]^{-1}\eta^2_{j,t} f^s_{j,t} + O_p(T^{-1/2} + N^{-1/2}).$$

Therefore, we have:

$$\hat{f}^s_{j,t} = \hat{\mathcal{H}}^{-1}_{s,j}\left[f^s_{j,t} - \tilde{\Sigma}_{jc}\tilde{\Sigma}^{-1}_{cc} f^c_t + \frac{1}{\sqrt{N_j}} v^s_{j,t} + \frac{1}{T}\beta^s_{j,t}\right] + o_p(N_j^{-1/2}), \qquad j = 1, 2, \tag{D.43}$$

uniformly w.r.t. $t = 1, ..., T$, where:

$$\beta^s_{j,t} = [\Sigma^{(ss)}_{\lambda,j}]^{-1}\left(\eta^2_{j,t} f^s_{j,t} - \Sigma^{(ss)}_{\lambda,j} E[f^s_{j,t}\beta^{c\,\prime}_{j,t}] f^c_t - \Sigma^{(sc)}_{\lambda,j}\tilde{\beta}^c_{j,t}\right).$$

Let us now show that $v^s_{j,t} = u^{(s)}_{j,t}$, the lower k^s_j-dimensional component of $u_{j,t}$. For this purpose, let us denote by $\tilde{\Sigma}_{ab}$ and $(\tilde{\Sigma}^{-1})_{ab}$, with $a, b = c, s$ the blocks of matrix $\tilde{\Sigma} \equiv \tilde{\Sigma}_{\lambda,j}$ and of its inverse $\tilde{\Sigma}^{-1}$. Then, we have:

$$v^s_{j,t} = \tilde{\Sigma}^{-1}_{ss}\frac{1}{\sqrt{N_j}}\sum_{i=1}^{N_j}\lambda^s_{j,i}\varepsilon_{j,i,t} - \tilde{\Sigma}^{-1}_{ss}\tilde{\Sigma}_{sc} u^{(c)}_{j,t},$$

and:

$$u^{(c)}_{j,t} = (\tilde{\Sigma}^{-1})_{cc}\frac{1}{\sqrt{N_j}}\sum_{i=1}^{N_j}\lambda^c_{j,i}\varepsilon_{j,i,t} + (\tilde{\Sigma}^{-1})_{cs}\frac{1}{\sqrt{N_j}}\sum_{i=1}^{N_j}\lambda^s_{j,i}\varepsilon_{j,i,t}.$$

Therefore, we get:

$$v^s_{j,t} = \tilde{\Sigma}^{-1}_{ss}[I_{k_j} - \tilde{\Sigma}_{sc}(\tilde{\Sigma}^{-1})_{cs}]\frac{1}{\sqrt{N_j}}\sum_{i=1}^{N_j}\lambda^s_{j,i}\varepsilon_{j,i,t} - \tilde{\Sigma}^{-1}_{ss}\tilde{\Sigma}_{sc}(\tilde{\Sigma}^{-1})_{cc}\frac{1}{\sqrt{N_j}}\sum_{i=1}^{N_j}\lambda^c_{j,i}\varepsilon_{j,i,t}.$$

From the property of the matrix inverse, $I_{k_j} - \tilde{\Sigma}_{sc}(\tilde{\Sigma}^{-1})_{cs} = \tilde{\Sigma}_{ss}(\tilde{\Sigma}^{-1})_{ss}$ and $\tilde{\Sigma}_{sc}(\tilde{\Sigma}^{-1})_{cc} = -\tilde{\Sigma}_{ss}(\tilde{\Sigma}^{-1})_{sc}$. Therefore, we get:

$$v^s_{j,t} = (\tilde{\Sigma}^{-1})_{ss}\frac{1}{\sqrt{N_j}}\sum_{i=1}^{N_j}\lambda^s_{j,i}\varepsilon_{j,i,t} + (\tilde{\Sigma}^{-1})_{sc}\frac{1}{\sqrt{N_j}}\sum_{i=1}^{N_j}\lambda^c_{j,i}\varepsilon_{j,i,t} = \left(\tilde{\Sigma}^{-1}\frac{1}{\sqrt{N_j}}\sum_{i=1}^{N_j}\lambda_{j,i}\varepsilon_{j,i,t}\right)^{(s)} = u^{(s)}_{j,t}.$$

Plugging the latter equation in (D.43) yields (D.20).

e) Asymptotic expansion of $\hat{\lambda}^s_{j,i}$

Let us now derive the asymptotic expansion of factor loadings estimator $\hat{\lambda}^s_{j,i}$ up to order $o_p\left(\bar{N}^{-1/2}\right)$. The analysis parallels the one in Subsection D.4.1 c). We have $\hat{\lambda}^s_{j,i} = \hat{F}^{s\,\prime}_j \xi_{j,i}/T$, where $\xi_{j,i}$ is the i-th column of matrix Ξ_j and $\hat{F}_j^s = [\hat{f}_{j,1}^s, ..., \hat{f}_{j,T}^s]'$. From equation (D.43) we have $\hat{F}_j^s = \left(\tilde{F}_j^s + \frac{1}{\sqrt{N_j}} U_j^s + \frac{1}{T} B_j^s\right)\left(\hat{\mathcal{H}}_{s,j}^{-1}\right)' + o_p\left(\bar{N}^{-1/2}\right)$, where $U_j^s = [u_{j,1}^{(s)}, ..., u_{j,T}^{(s)}]'$ and $B_j^s = [\beta_{j,1}^s, ..., \beta_{j,T}^s]'$, which implies:

$$\hat{F}_j^s \hat{\mathcal{H}}_{j,s}' - \tilde{F}_j^s = \frac{1}{\sqrt{N_j}} U_j^s + \frac{1}{T} B_j^s + o_p\left(\bar{N}^{-1/2}\right). \tag{D.44}$$

Then:

$$\begin{aligned} \hat{\lambda}_{j,i}^s &= \frac{1}{T}\hat{F}_j^{s\prime}\xi_{j,i} \;=\; \frac{1}{T}\hat{F}_j^{s\prime}\left(\tilde{F}_j^s \lambda_{j,i}^s + e_{j,i}\right) + o_p(\bar{N}^{-1/2}) \\ &= \frac{1}{T}\hat{F}_j^{s\prime}\left(\left[\hat{F}_j^s \hat{\mathcal{H}}_{j,s}' - \left(\hat{F}_j^s \hat{\mathcal{H}}_{j,s}' - \tilde{F}_j^s\right)\right]\lambda_{j,i}^s + e_{j,i}\right) + o_p(\bar{N}^{-1/2}) \\ &= \hat{\mathcal{H}}_{j,s}'\lambda_{j,i}^s - \frac{1}{T}\hat{F}_j^{s\prime}\left(\hat{F}_j^s \hat{\mathcal{H}}_{j,s}' - \tilde{F}_j^s\right)\lambda_{j,i}^s + \frac{1}{T}\hat{F}_j^{s\prime} e_{j,i} + o_p(\bar{N}^{-1/2}), \qquad j = 1, 2, \end{aligned}$$

uniformly in $i = 1, ..., N_j$. By writing $\hat{F}_j^s = \left[\tilde{F}_j^s + (\hat{F}_j^s \hat{\mathcal{H}}_{j,s}' - \tilde{F}_j^s)\right](\hat{\mathcal{H}}_{j,s}')^{-1}$, and rearranging terms, we get:

$$\begin{aligned} \hat{\lambda}_{j,i}^s &= \hat{\mathcal{H}}_{s,j}'\Bigg\{\lambda_{j,i}^s + (\hat{\mathcal{H}}_{j,s}')^{-1}(\hat{\mathcal{H}}_{j,s})^{-1}\frac{1}{T}\tilde{F}_j^{s\prime} e_{j,i} \\ &\quad + (\hat{\mathcal{H}}_{j,s}')^{-1}(\hat{\mathcal{H}}_{j,s})^{-1}\frac{1}{T}(\hat{F}_j^s \hat{\mathcal{H}}_{j,s}' - \tilde{F}_j^s)' e_{j,i} \\ &\quad - (\hat{\mathcal{H}}_{j,s}')^{-1}(\hat{\mathcal{H}}_{j,s})^{-1}\frac{1}{T}\left[\tilde{F}_j^s + (\hat{F}_j^s \hat{\mathcal{H}}_{j,s}' - \tilde{F}_j^s)\right]'(\hat{F}_j^s \hat{\mathcal{H}}_{j,s}' - \tilde{F}_j^s)\lambda_{j,i}^s\Bigg\} + o_p(\bar{N}^{-1/2}). \end{aligned} \tag{D.45}$$

By using equations $e_{j,i} = \varepsilon_{j,i} - \frac{1}{\sqrt{T}} F^c w_{j,i}^c - \frac{1}{\sqrt{N_j}} U_j^c \lambda_{j,i}^c - \frac{1}{T} F^c \beta_{\Lambda,j,i}^c - \frac{1}{T} B_j^c \lambda_{j,i}^c$ and $\tilde{F}_j^{s\prime} F^c = 0$, equation (D.44), and paralleling the computations in Subsection D.4.1 c), we get:

$$\begin{aligned} \frac{1}{T}\tilde{F}_j^{s\prime} e_{j,i} &= \frac{1}{T}\tilde{F}_j^{s\prime}\varepsilon_{j,i} - \frac{1}{T}E[f_{j,t}^s \beta_{j,t}^{c\prime}]\lambda_{j,i}^c + o_p(\bar{N}^{-1/2}), \\ \frac{1}{T}(\hat{F}_j^s \hat{\mathcal{H}}_{j,s}' - \tilde{F}_j^s)' e_{j,i} &= o_p(\bar{N}^{-1/2}), \\ \frac{1}{T}\left[\tilde{F}_j^s + (\hat{F}_j^s \hat{\mathcal{H}}_{j,s}' - \tilde{F}_j^s)\right]'(\hat{F}_j^s \hat{\mathcal{H}}_{j,s}' - \tilde{F}_j^s) &= \frac{1}{T}E[f_{j,t}^s \beta_{j,t}^{s\prime}] + o_p(\bar{N}^{-1/2}), \\ (\hat{\mathcal{H}}_{j,s}')^{-1}(\hat{\mathcal{H}}_{j,s})^{-1} &= (\tilde{F}_j^{s\prime}\tilde{F}_j^s/T)^{-1} + o_p(\bar{N}^{-1/2}), \end{aligned}$$

uniformly in $i = 1, ..., N_j$. Thus, from (D.45) we get:

$$\hat{\lambda}_{j,i}^s = \hat{\mathcal{H}}_{s,j}'\left\{\lambda_{j,i}^s + (\tilde{F}_j^{s\prime}\tilde{F}_j^s/T)^{-1}\frac{1}{T}\tilde{F}_j^{s\prime}\varepsilon_{j,i} - \frac{1}{T}\left(E[f_{j,t}^s \beta_{j,t}^{c\prime}]\lambda_{j,i}^c + E[f_{j,t}^s \beta_{j,t}^{s\prime}]\lambda_{j,i}^s\right)\right\} + o_p(\bar{N}^{-1/2}),$$

uniformly in $i = 1, ..., N_j$. This equation can be written as:

$$\hat{\lambda}_{j,i}^s = \hat{\mathcal{H}}_{s,j}'\left[\lambda_{j,i}^s + \frac{1}{\sqrt{T}} w_{j,i}^s + \frac{1}{T}\beta_{\Lambda,j,i}^s\right] + o_p(\bar{N}^{-1/2}),$$

where:

$$\begin{aligned} w^s_{j,i} &= (\tilde{F}^{s\,\prime}_j \tilde{F}^s_j/T)^{-1}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\tilde{f}^s_{j,t}\varepsilon_{j,i,t}, \\ \beta^s_{\Lambda,j,i} &= -E[f^s_{j,t}\beta^{c\,\prime}_{j,t}]\lambda^c_{j,i} - E[f^s_{j,t}\beta^{s\,\prime}_{j,t}]\lambda^s_{j,i}. \end{aligned}$$

f) Asymptotic expansions up to order $o_p(T^{-1/2})$

Let us start by establishing the uniform asymptotic expansion of estimator $\hat{f}^c_t$ at order $o_p(T^{-1/2})$. From (D.31), using $(\log T)^{\bar{b}}\delta_{N,T} = o(T^{-1/2})$, for any $\bar{b} > 0$, and the uniform bounds (C.105)-(D.26), we get:

$$\hat{f}^c_t = \hat{\mathcal{H}}_c^{-1}\left(f^c_t + \frac{1}{\sqrt{N_1}}u^{(c)}_{1,t}\right) + o_p(T^{-1/2}),$$

uniformly in $t = 1, ..., T$, which yields the uniform bound for $\hat{f}^c_t$. The uniform bounds for the other estimators follow by paralleling the arguments in Subsection D.4.1 c)-e). ■

D.5 Asymptotic distribution of factors and loadings in generic group factor model

The next proposition provides the asymptotic distribution of the common and group-specific factors estimators introduced in Definitions 1 and 2 in the main body of the paper. To simplify the proof, we assume that N_1 and N_2, with $N_2 \leq N_1$, grow at the same rate, i.e., $N_2/N_1 \to \mu$ with $\mu > 0$. This condition could be relaxed at the expense of a more involved restriction on N_1, N_2, T.

PROPOSITION D.5. *Under Assumption A.1 with $\mu > 0$, and Assumptions A.2 - A.8 we have:*

$$\sqrt{N_1}\begin{bmatrix} \hat{\mathcal{H}}_c\hat{f}^c_t - f^c_t - \frac{1}{T}\beta^c_{1,t} \\ \hat{\mathcal{H}}_{s,1}\hat{f}^s_{1,t} - \left(f^s_{1,t} - (F^{s\,\prime}_1 F^c)(F^{c\,\prime}F^c)^{-1}f^c_t\right) - \frac{1}{T}\beta^s_{1,t} \end{bmatrix} \xrightarrow{d} N\left(0, \Sigma_{u,11,t}\right), \quad (\mathcal{F}_t\text{-stably}), \tag{D.46}$$

and:

$$\sqrt{N_2}\begin{bmatrix} \hat{\mathcal{H}}^*_c\hat{f}^{c\,*}_t - f^c_t - \frac{1}{T}\beta^c_{2,t} \\ \hat{\mathcal{H}}_{s,2}\hat{f}^s_{2,t} - \left(f^s_{2,t} - (F^{s\,\prime}_2 F^c)(F^{c\,\prime}F^c)^{-1}f^c_t\right) - \frac{1}{T}\beta^s_{2,t} \end{bmatrix} \xrightarrow{d} N\left(0, \Sigma_{u,22,t}\right), \quad (\mathcal{F}_t\text{-stably}), \tag{D.47}$$

for any t, where matrices $\hat{\mathcal{H}}_c$, $\hat{\mathcal{H}}^*_c$ and $\hat{\mathcal{H}}_{s,j}$ are such that $\hat{\mathcal{H}}_c\hat{\mathcal{H}}_c' = (\frac{1}{T}F^{c\,\prime}F^c)^{-1} + o_p(N_1^{-1/2})$, $\hat{\mathcal{H}}^*_c\hat{\mathcal{H}}^{*\,\prime}_c = (\frac{1}{T}F^{c\,\prime}F^c)^{-1} + o_p(N_2^{-1/2})$ and $\hat{\mathcal{H}}_{s,j}\hat{\mathcal{H}}'_{s,j} = (\frac{1}{T}\tilde{F}^{s\,\prime}_j\tilde{F}^s_j)^{-1} + o_p(N_j^{-1/2})$, we define $F^c = [f^c_1, ..., f^c_T]'$, $F^s_j = [f^s_{j,1}, ..., f^s_{j,T}]'$ and $\tilde{F}^s_j = F^s_j - F^c(F^{c\prime}F^c)^{-1}(F^{c\prime}F^s_j)$ for $j = 1, 2$, and the bias terms are:

$$\begin{aligned} \beta^c_{1,t} &= \bar{b}^{(c)}_{1,t} - E[(\bar{b}^{(c)}_{1,t} - \bar{b}^{(c)}_{2,t})(f^s_{1,t} - \Phi f^s_{2,t})'](I_{k_1-k^c} - \Phi\Phi')^{-1}f^s_{1,t}, \\ \beta^c_{2,t} &= \bar{b}^{(c)}_{2,t} - E[(\bar{b}^{(c)}_{2,t} - \bar{b}^{(c)}_{1,t})(f^s_{2,t} - \Phi' f^s_{1,t})'](I_{k_2-k^c} - \Phi'\Phi)^{-1}f^s_{2,t}, \\ \beta^s_{j,t} &= [\Sigma^{(ss)}_{\lambda,j}]^{-1}\left(\eta^2_{j,t}f^s_{j,t} - \Sigma^{(ss)}_{\lambda,j}E[f^s_{j,t}\beta^{c\,\prime}_{j,t}]f^c_t - \Sigma^{(sc)}_{\lambda,j}\tilde{\beta}^c_{j,t}\right), \quad j = 1, 2, \end{aligned}$$

and $\tilde{\beta}^c_{j,t} = \beta^c_{j,t} - E[\beta^c_{j,t}f^{c\,\prime}_t]f^c_t$ is the residual of the orthogonal projection of $\beta^c_{j,t}$ onto f^c_t.

From Proposition D.5 a linear transformation of vector $\hat{f}_t^c$ (resp. $\hat{f}_t^{c*}$) estimates the common factor f_t^c at rate $1/\sqrt{N_1}$ (resp. $1/\sqrt{N_2}$) with a bias of order $1/T$. The variance of the asymptotic Gaussian distribution is the upper-left (c, c) block of matrix $\Sigma_{u,11,t}$ (resp. $\Sigma_{u,22,t}$), 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 $o_p(N_1^{-1/2})$, and therefore asymptotically negligible. The estimator $\hat{f}_{j,t}^s$ approximates the residual of the sample projection of the group-j specific factor on the common factor, up to a linear transformation, at rate $1/\sqrt{N_j}$ and with an asymptotic bias of order $1/T$.
Let us now derive the asymptotic distribution of the factor loadings estimators in equations (3.3) and (3.4). For this purpose, we introduce the next assumption.

Assumption D.1. *We have for any $j = 1, 2$ and $i \geq 1$:*

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \begin{bmatrix} f_t^c \varepsilon_{j,i,t} \\ f_{j,t}^s \varepsilon_{j,i,t} \\ f_{j,t}^s \otimes f_t^c \end{bmatrix} \xrightarrow{d} N \left(0, \begin{bmatrix} \Phi_{j,i}^{cc} & \Phi_{j,i}^{cs} & 0 \\ \Phi_{j,i}^{sc} & \Phi_{j,i}^{ss} & 0 \\ 0 & 0 & \Psi_j \end{bmatrix} \right),$$

as $T \to \infty$, where:

$$\begin{aligned} \Phi_{j,i}^{cc} &= \sum_{h=-\infty}^{\infty} E[f_t^c f_{t-h}^{c\prime} \varepsilon_{j,i,t} \varepsilon_{j,i,t-h}], \qquad \Phi_{j,i}^{cs} = \sum_{h=-\infty}^{\infty} E[f_t^c f_{j,t-h}^{s\prime} \varepsilon_{j,i,t} \varepsilon_{j,i,t-h}] = (\Phi_{j,i}^{sc})', \\ \Phi_{j,i}^{ss} &= \sum_{h=-\infty}^{\infty} E[f_{j,t}^s f_{j,t-h}^{s\prime} \varepsilon_{j,i,t} \varepsilon_{j,i,t-h}], \qquad \Psi_j = \sum_{h=-\infty}^{\infty} E\left[f_{j,t}^s f_{j,t-h}^{s\prime} \otimes f_t^c f_{t-h}^{c\prime}\right]. \end{aligned}$$

Assumption D.1 states that time series averages of the error terms scaled by the factors, as well as time series averages of the cross-products of common and specific factors, are asymptotically Gaussian. It is used to show the asymptotic normality of the loadings estimators in Proposition D.6, and is implied by e.g. a mixing condition on the individual error series jointly with the factor process. The part of Assumption D.1 concerning scaled error terms corresponds to Assumption F.4 in Bai (2003).

PROPOSITION D.6. *Under Assumption A.1 with $\mu > 0$, Assumptions A.2 - A.8 and D.1 we have:*

$$\sqrt{T} \begin{bmatrix} \left(\hat{\mathcal{H}}_c'\right)^{-1} \hat{\lambda}_{j,i}^c - \lambda_{j,i}^c \\ \left(\hat{\mathcal{H}}_{s,j}'\right)^{-1} \hat{\lambda}_{j,i}^s - \lambda_{j,i}^s \end{bmatrix} \xrightarrow{d} N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \left(\Phi_{j,i}^{cc} + (\lambda_{j,i}^{s\prime} \otimes I_{k^c}) \Psi_j (\lambda_{j,i}^s \otimes I_{k^c})\right) & \Psi_{j,i}^{cs} \\ \Psi_{j,i}^{sc} & \Psi_{j,i}^{ss} \end{bmatrix} \right), \quad \text{(D.48)}$$

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 Proposition D.5.

The factor loadings are estimated at rate $\sqrt{T}$. Matrix $\Phi_{j,i}^{cc}$ is the asymptotic variance for cross-sectional OLS regression of data in group j on the true values of the common factor. The additional component in the asymptotic variance of estimator $\hat{\lambda}_{j,i}^c$ is due to the fact that the true values of common and group-specific factors are not orthogonal in-sample. This fact is not taken into account by the estimator of factor loadings. Finally, there are no bias terms at order N_1^{-1}, N_2^{-1} in the large sample distributions of factor loadings, since in our asymptotics $\sqrt{T}/N = o(1)$ and hence such bias terms are negligible.

D.5.1 Proof of Proposition D.5

We use the asymptotic expansions in Proposition D.4 i). Specifically, equations (D.19) and (D.20) for $j = 1$ imply:

$$\sqrt{N_1}\left[\begin{array}{c} \hat{\mathcal{H}}_c \hat{f}_t^c - f_t^c - \frac{1}{T}\beta_{1,t}^c \\ \hat{\mathcal{H}}_{s,1}\hat{f}_{1,t}^s - (f_{1,t}^s - (F_1^{s\prime}F^c)(F^{c\prime}F^c)^{-1}f_t^c) - \frac{1}{T}\beta_{1,t}^s \end{array}\right] = u_{1,t} + o_p(1).$$

From Assumptions A.3 and A.5 *a)*, we have $u_{1,t} \xrightarrow{d} N(0, \Sigma_{u,11,t})$, $\mathcal{F}_t$-stably. Then, the asymptotic distribution in (D.46) follows. The asymptotic distribution in (D.47) can be establish along similar lines.

D.5.2 Proof of Proposition D.6

We prove Proposition D.6 by the asymptotic expansions in Proposition D.4 i), by keeping only terms up to $o_p(T^{-1/2})$. Specifically, equation (D.21) implies:

$$\begin{aligned} \sqrt{T}\left[\left(\hat{\mathcal{H}}_c'\right)^{-1}\hat{\lambda}_{j,i}^c - \lambda_{j,i}^c\right] &= w_{j,i}^c + (F^{c\prime}F_j^s/\sqrt{T})\lambda_{j,i}^s + o_p(1) \\ &= \frac{1}{\sqrt{T}}\sum_{t=1}^T f_t^c(\varepsilon_{j,i,t} + f_{j,t}^{s\,\prime}\lambda_{j,i}^s) + o_p(1) \\ &= \frac{1}{\sqrt{T}}\sum_{t=1}^T \left[f_t^c\varepsilon_{j,i,t} + (\lambda_{j,i}^{s\prime}\otimes I_{k^c})(f_{j,t}^s\otimes f_t^c)\right] + o_p(1). \end{aligned}$$

Moreover, equation (D.22) imply:

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

Thus, we get:

$$\sqrt{T}\left[\begin{array}{c} \left(\hat{\mathcal{H}}_c'\right)^{-1}\hat{\lambda}_{j,i}^c - \lambda_{j,i}^c \\ \left(\hat{\mathcal{H}}_{s,j}'\right)^{-1}\hat{\lambda}_{j,i}^s - \lambda_{j,i}^s \end{array}\right] = \frac{1}{\sqrt{T}}\sum_{t=1}^T \left[\begin{array}{c} f_t^c\varepsilon_{j,i,t} + (\lambda_{j,i}^{s\prime}\otimes I_{k^c})(f_{j,t}^s\otimes f_t^c) \\ f_{j,t}^s\varepsilon_{j,i,t} \end{array}\right] + o_p(1).$$

Then, Assumption D.1 yields (D.48). ■

D.6 Asymptotic distribution of factor estimates in a mixed frequency model

For the mixed frequency factor model in equation (5.1) with flow sampling, the Assumptions A.1-A.9 are meant to apply with errors $\varepsilon_{1,i,t} = \bar{e}_t^{H,i}$ and $\varepsilon_{2,i,t} = \bar{e}_t^{L,i}$, vectors of common and group-specific factors $f_t^c = \bar{g}_t^C$, $f_{1,t}^s = \bar{g}_t^H$ and $f_{2,t}^s = \bar{g}_t^L$, and loadings matrices $\Lambda_1 = [\ \Lambda_{HC} \vdots \Lambda_H\ ]$ and $\Lambda_2 = [\ \Lambda_{LC} \vdots \Lambda_L\ ]$. The cross-sectional dimensions are $N_1 = N_H$ and $N_2 = N_L$. Additionally, we make the following assumption:

Assumption D.2. *The variables $\lambda_{1,i}$ and $e_{m,t}^{H,i}$ are such that:*

$$\frac{1}{\sqrt{N_H}}\sum_{i=1}^{N_H}\lambda_{1,i}\left[\begin{array}{c} e_{m,t}^{H,i} \\ e_{m',t}^{H,i} \end{array}\right] \xrightarrow{d} N\left(0, \left[\begin{array}{cc} \Omega_{\Lambda,m,m,t}^H & \Omega_{\Lambda,m,m',t}^H \\ \Omega_{\Lambda,m',m,t}^H & \Omega_{\Lambda,m',m',t}^H \end{array}\right]\right), \qquad (\mathcal{F}_t - stably),$$

as $N_H \to \infty$, where

$$\Omega^H_{\Lambda,m,m',t} = \underset{N_H\to\infty}{plim} \frac{1}{N_H}\sum_{i=1}^{N_H}\sum_{\ell=1}^{N_H} \lambda_{1,i}\lambda'_{1,\ell} Cov(e^{H,i}_{m,t}, e^{H,\ell}_{m',t}|\mathcal{F}_t), \qquad m,m' = 1,...,M.$$

Assumption D.2 is analogous to Assumption A.5 *a)* expressed for the high frequency DGP of the idiosyncratic innovation terms $e^{i,H}_{m,t}$.

PROPOSITION D.7. *Under the asymptotics in Assumption A.1 with $\mu > 0$, and Assumptions A.2 - A.8, D.2, the estimators $\hat{g}^C_t$, $\hat{g}^H_{m,t}$, $\hat{\tilde{g}}^L_t$ in Section 5 are such that:*

$$\sqrt{N_H}\begin{bmatrix} \hat{\mathcal{H}}_c \hat{g}^C_{m,t} - g^C_{m,t} - \frac{1}{T}\beta^C_{m,t} \\ \hat{\mathcal{H}}_{1,s}\hat{g}^H_{m,t} - [g^H_{m,t} - (\bar{g}^{H\prime}\bar{g}^C)(\bar{g}^{C\prime}\bar{g}^C)^{-1}g^C_{m,t}] - \frac{1}{T}\beta^H_{m,t}\end{bmatrix} \xrightarrow{d} N\left(0, \Sigma^{-1}_{\Lambda,1}\Omega^H_{\Lambda,m,m,t}\Sigma^{-1}_{\Lambda,1}\right),$$

$\mathcal{F}_t$-stably, the vectors for sub-periods m and m' are jointly asymptotically Gaussian ($\mathcal{F}_t$-stably) with covariance $\Sigma^{-1}_{\Lambda,1}\Omega^H_{\Lambda,m,m',t}\Sigma^{-1}_{\Lambda,1}$, and:

$$\sqrt{N_L}\left[\hat{\mathcal{H}}_{2,s}\hat{\tilde{g}}^L_t - [\bar{g}^L_t - (\bar{g}^{L\prime}\bar{g}^C)(\bar{g}^{C\prime}\bar{g}^C)^{-1}\bar{g}^C_t] - \frac{1}{T}\beta^L_t\right] \xrightarrow{d} N\left(0, \left(\Sigma^{-1}_{\Lambda,2}\Omega^L_{\Lambda,t}\Sigma^{-1}_{\Lambda,2}\right)^{(LL)}\right),$$

$\mathcal{F}_t$-stably, for any m, m', t, where matrices $\hat{\mathcal{H}}_c$, $\hat{\mathcal{H}}_{1,s}$ and $\hat{\mathcal{H}}_{2,s}$ are such that $\hat{\mathcal{H}}_c\hat{\mathcal{H}}'_c = (\frac{1}{T}\bar{g}^{C\prime}\bar{g}^C)^{-1} + o_p(\bar{N}^{-1/2})$, $\hat{\mathcal{H}}_{1,s}\hat{\mathcal{H}}'_{1,s} = (\frac{1}{T}\tilde{g}^{H\prime}\tilde{g}^H)^{-1} + o_p(\bar{N}^{-1/2})$ and $\hat{\mathcal{H}}_{2,s}\hat{\mathcal{H}}'_{2,s} = (\frac{1}{T}\tilde{g}^{L\prime}\tilde{g}^L)^{-1} + o_p(\bar{N}^{-1/2})$ with $\bar{N} := \max\{N_H, T\}$, we define $\bar{g}^U = [\bar{g}^U_1, ..., \bar{g}^U_T]'$, for $U = C, H, L$, $\tilde{g}^U = \bar{g}^U - \bar{g}^C(\bar{g}^{C\prime}\bar{g}^C)^{-1}(\bar{g}^{C\prime}\bar{g}^U)$, for $U = H, L$, $\Sigma_{\Lambda,1} = \lim_{N_H\to\infty}\frac{1}{N_H}\sum_{i=1}^{N_H}\lambda_{1,i}\lambda'_{1,i}$ and $\Sigma_{\Lambda,2} = \lim_{N_L\to\infty}\frac{1}{N_L}\sum_{i=1}^{N_L}\lambda_{2,i}\lambda'_{2,i}$, the bias terms are:

$$\begin{aligned}
\begin{bmatrix}\beta^C_{m,t}\\ \beta^H_{m,t}\end{bmatrix} &= \Sigma^{-1}_{\Lambda,1}\left(\bar{\eta}^2_{1,m,t}h_{1,t} - \bar{\Phi}_H h_{1,m,t}\right) - B_{\Lambda,1}h_{1,m,t}, \qquad B_{\Lambda,1} = \begin{bmatrix} -E[\beta^C_t\bar{g}^{C\prime}_t] & -E[\beta^C_t\bar{g}^{H\prime}_t] \\ E[\bar{g}^H_t\beta^{C\prime}_t] & -E[\beta^H_t\bar{g}^{H\prime}_t]\end{bmatrix},\\
\beta^C_t &= \bar{b}^{(C)}_{1,t} - E[(\bar{b}^{(C)}_{1,t} - \bar{b}^{(C)}_{2,t})(\bar{g}^H_t - \Phi\bar{g}^L_t)'](I_{k^H} - \Phi\Phi')^{-1}\bar{g}^H_t,\\
\beta^H_t &= (\Sigma^{(HH)}_{\Lambda,1})^{-1}\left(\bar{\eta}^2_{1,t}\bar{g}^H_t - \Sigma^{(HH)}_{\Lambda,1}E[\bar{g}^H_t\beta^{C\prime}_t]\bar{g}^C_t - \Sigma^{(HC)}_{\Lambda,1}\tilde{\beta}^C_t\right), \qquad \tilde{\beta}^C_t = \beta^C_t - E[\beta^C_t\bar{g}^{C\prime}_t]\bar{g}^C_t,\\
\beta^L_t &= (\Sigma^{(LL)}_{\Lambda,2})^{-1}\left(\bar{\eta}^2_{2,t}\bar{g}^L_t - \Sigma^{(LL)}_{\Lambda,2}E[\bar{g}^L_t\beta^{C\prime}_t]\bar{g}^C_t - \Sigma^{(LC)}_{\Lambda,2}\tilde{\beta}^C_t\right),
\end{aligned}$$

with $\bar{b}_{1,t} = \Sigma^{-1}_{\Lambda,1}\bar{\eta}^2_{1,t}h_{1,t}$, and the upper index (LL) denotes the lower-right (k^L, k^L) block of a matrix, and similarly for (HH), (LC).

From Proposition D.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 low-frequency sample projection of the high-frequency factor on the common factor. [2] The estimation rate is $\sqrt{N_H}$. As an oracle property, the asymptotic variance $\Sigma^{-1}_{\Lambda,1}\Omega^H_{\Lambda,m,m,t}\Sigma^{-1}_{\Lambda,1}$ equals the asymptotic variance of the infeasible estimator obtained by principal components method applied on the HF panel and rotated with a known matrix to separate the common and frequency-specific factors (see Theorem 1 and Proposition 1 in Bai (2003) for the asymptotic distribution of principal component estimators when $N \ll$

[2] Matrices $\hat{\mathcal{H}}_c$ and $\hat{\mathcal{H}}_{1,s}$ are not feasible estimators, and thus $\hat{\mathcal{H}}_c\hat{g}^C_{m,t}$ and $\hat{\mathcal{H}}_{1,s}\hat{g}^H_{m,t}$ are not empirically computable quantities. From Proposition D.7, vector $\hat{g}^C_{m,t}$ itself estimates a linear transformation of $g^C_{m,t}$ for which the flow-sampled values have identity sample variance-covariance matrix, and similarly for $\hat{g}^H_{m,t}$.

T^2). It differs from the asymptotic variance in Theorem 2 in Wang (2012) since the number of groups is constant in our asymptotic scheme. An error-in-variable problem originates both from estimation uncertainty, and asymptotic bias at order T^{-1}, of factor loadings used in the cross-sectional regression. This error-in-variable problem does not prevent consistency and asymptotic normality of the factor estimates in our double asymptotics, but leads to a bias term at order $1/T$. Vectors $\frac{1}{T}\beta_t^C$ and $\frac{1}{T}\beta_t^H$ are the biases at order T^{-1} of the estimators $\hat{\bar{g}}_t^C$ and $\hat{\bar{g}}_t^H$ of the flow-sampled common and high-frequency-specific factors.[3] Such biases induce biases at order T^{-1} in factor loadings, which in turn yield term $-\frac{1}{T}B_{\Lambda,1}h_{1,m,t}$ in the asymptotic bias of high-frequency factor estimates. Similarly, a linear transformation of vector $\hat{\bar{g}}_t^L$ estimates the projection residual of the flow-sampled low-frequency factor with convergence rate $\sqrt{N_L}$ and asymptotic bias at order $1/T$.

D.6.1 Proof of Proposition D.7

Let us first establish the asymptotic distribution of estimator $\hat{\bar{g}}_t^L$. After replacing $\varepsilon_{2,i,t} = \bar{e}_t^{i,L}$ and $F^c = \bar{g}^C$, $F_1^s = \bar{g}^H$, $F_2^s = \bar{g}^L$, from the asymptotic expansion (D.20) in Proposition D.4 i) we get:

$$\hat{\bar{g}}_t^L = \hat{\mathcal{H}}_{2,s}^{-1}\left[\bar{g}_t^L - (\bar{g}^{L\prime}\bar{g}^C)(\bar{g}^{C\prime}\bar{g}^C)^{-1}\bar{g}_t^C + \frac{1}{\sqrt{N_L}}u_{2,t}^{(L)} + \frac{1}{T}\beta_t^L\right] + o_p(N_L^{-1/2}),$$

where $u_{2,t}^{(L)}$ denotes the lower $(k^L,1)$ block of vector $u_{2,t} = \left(\frac{1}{N_L}\sum_{i=1}^{N_L}\lambda_{2,i}\lambda_{2,i}'\right)^{-1}\frac{1}{\sqrt{N_L}}\sum_{i=1}^{N_L}\lambda_{2,i}\bar{e}_t^{Li}$, and vector β_t^L is given in the statement of Proposition D.7. Moreover, we have $u_{2,t} \xrightarrow{d} N\left(0, \Sigma_{\Lambda,2}^{-1}\Omega_{\Lambda,t}^L\Sigma_{\Lambda,2}^{-1}\right)$, $\mathcal{F}_t$-stably, from Assumptions A.3 and A.5 *a)*. By rearranging terms, the asymptotic Gaussian distribution of estimator $\hat{\bar{g}}_t^L$ follows.

Let us now derive the asymptotic distribution of estimators $\hat{g}_{m,t}^C$ and $\hat{g}_{m,t}^H$. For this purpose, 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 (5.1) as:

$$\begin{aligned} x_{m,t}^H &= \Lambda_{HC}g_{m,t}^C + \Lambda_H g_{m,t}^H + e_{m,t}^H \;=\; \Lambda_1 g_{m,t} + e_{m,t}^H \\ &= \hat{\Lambda}_1\hat{\mathcal{U}}_1^{-1}g_{m,t} - \left(\hat{\Lambda}_1\hat{\mathcal{U}}_1^{-1} - \Lambda_1\right)g_{m,t} + e_{m,t}^H, \end{aligned} \tag{D.49}$$

where $g_{m,t} = [\,g_{m,t}^{C\,\prime}, g_{m,t}^{H\,\prime}\,]'$, $\Lambda_1 = [\Lambda_{HC} \vdots \Lambda_H] = [\Lambda_1^c \vdots \Lambda_1^s]$, $\hat{\Lambda}_1 = [\hat{\Lambda}_{HC} \vdots \hat{\Lambda}_H] = [\hat{\Lambda}_1^c \vdots \hat{\Lambda}_1^s]$, and matrix $\hat{\mathcal{U}}_1$ is defined in (C.103). By substituting equation (D.49) into estimator $\hat{g}_{m,t} = [\,\hat{g}_{m,t}^{C\,\prime}, \hat{g}_{m,t}^{H\,\prime}\,]' = \left(\hat{\Lambda}_1'\hat{\Lambda}_1\right)^{-1}\hat{\Lambda}_1'x_{m,t}^H$, and rearranging terms, we get:

$$\hat{g}_{m,t} \;=\; \hat{\mathcal{U}}_1^{-1}g_{m,t} - \left(\frac{\hat{\Lambda}_1'\hat{\Lambda}_1}{N_H}\right)^{-1}\frac{1}{N_H}\hat{\Lambda}_1'\left(\hat{\Lambda}_1\hat{\mathcal{U}}_1^{-1} - \Lambda_1\right)g_{m,t} + \left(\frac{\hat{\Lambda}_1'\hat{\Lambda}_1}{N_H}\right)^{-1}\frac{1}{N_H}\hat{\Lambda}_1'e_{m,t}^H. \tag{D.50}$$

From Lemma B.8 i) we have $\hat{\Lambda}_1'\hat{\Lambda}_1/N_H = \hat{\mathcal{U}}_1'\Sigma_{\Lambda,1}\hat{\mathcal{U}}_1 + o_p(1)$, which implies:

$$\left(\frac{\hat{\Lambda}_1'\hat{\Lambda}_1}{N_H}\right)^{-1} \;=\; \hat{\mathcal{U}}_1^{-1}\Sigma_{\Lambda,1}^{-1}\left(\hat{\mathcal{U}}_1'\right)^{-1} + o_p(1). \tag{D.51}$$

By plugging (D.51) into the third term in the r.h.s. of (D.50), using the equation $\hat{\Lambda}_1'e_{m,t}^H = \hat{\mathcal{U}}_1'\left(\hat{\Lambda}_1\hat{\mathcal{U}}_1^{-1} - \Lambda_1\right)'e_{m,t}^H +$

[3] The asymptotic distributions of such factors and factor loadings estimators, can be obtained from Proposition D.5.

$\hat{\mathcal{U}}_1'\Lambda_1' e^H_{m,t}$, and rearranging terms, we get the asymptotic expansion:

$$\begin{aligned} \hat{\mathcal{U}}_1\hat{g}_{m,t} &= g_{m,t} + \left[\Sigma_{\Lambda,1}^{-1} + o_p(1)\right]\frac{1}{N_H}\Lambda_1' e^H_{m,t} \\ &+ \left[\hat{\mathcal{U}}_1(\hat{\Lambda}_1'\hat{\Lambda}_1/N_H)^{-1}\hat{\mathcal{U}}_1'\right]\left[-\frac{1}{N_H}\left(\hat{\Lambda}_1\hat{\mathcal{U}}_1^{-1}\right)'\left(\hat{\Lambda}_1\hat{\mathcal{U}}_1^{-1} - \Lambda_1\right)g_{m,t} + \frac{1}{N_H}\left(\hat{\Lambda}_1\hat{\mathcal{U}}_1^{-1} - \Lambda_1\right)' e^H_{m,t}\right]. \end{aligned} \tag{D.52}$$

The terms in the first line on the r.h.s. of (D.52) correspond to the (infeasible) cross-sectional regression of observables on true factor loadings. The terms in the second line account for replacing the true factor loadings with the estimated ones in the feasible regression. We control the latter terms by using the uniform asymptotic expansion of the estimated factor loadings derived in Proposition D.4 i). We need to bound the remainder terms up to order $o_p(N_H^{-1/2})$ and take into account the bias terms of order T^{-1}. This yields the next lemma.

LEMMA D.8. *Under Assumption A.1 with $\mu > 0$, Assumptions A.2 - A.8 we have:*

$$\begin{aligned} &\left[\hat{\mathcal{U}}_1(\hat{\Lambda}_1'\hat{\Lambda}_1/N_H)^{-1}\hat{\mathcal{U}}_1'\right]\frac{1}{N_H}\left(\hat{\Lambda}_1\hat{\mathcal{U}}_1^{-1}\right)'\left(\hat{\Lambda}_1\hat{\mathcal{U}}_1^{-1} - \Lambda_1\right)g_{m,t} \\ &= \begin{bmatrix} 0_{(k^C\times 1)} \\ (\bar{g}^{H\prime}\bar{g}^C)(\bar{g}^{C\prime}\bar{g}^C)^{-1}g^C_{m,t} \end{bmatrix} + \frac{1}{T}\left(\Sigma_{\Lambda,1}^{-1}\bar{\Phi}_H + B_{\Lambda,1}\right)g_{m,t} + o_p(N_H^{-1/2}), \end{aligned} \tag{D.53}$$

and:

$$\left[\hat{\mathcal{U}}_1(\hat{\Lambda}_1'\hat{\Lambda}_1/N_H)^{-1}\hat{\mathcal{U}}_1'\right]\frac{1}{N_H}\left(\hat{\Lambda}_1\hat{\mathcal{U}}_1^{-1} - \Lambda_1\right)' e^H_{m,t} = \frac{1}{T}\Sigma_{\Lambda,1}^{-1}\bar{\eta}^2_{1,m,t}\bar{g}_t + o_p(N_H^{-1/2}), \tag{D.54}$$

where $\bar{\Phi}_H$, $B_{\Lambda,1}$ and $\bar{\eta}^2_{1,m,t}$ are defined in Proposition D.7.

From (D.52) and Lemma D.8 we get:

$$\hat{\mathcal{U}}_1\hat{g}_{m,t} = g_{m,t} + \left[\Sigma_{\Lambda,1}^{-1} + o_p(1)\right]\left[\frac{1}{N_H}\Lambda_1' e^H_{m,t}\right] - \begin{bmatrix} 0_{(k^C\times 1)} \\ (\bar{g}^{H\prime}\bar{g}^C)(\bar{g}^{C\prime}\bar{g}^C)^{-1}g^C_{m,t} \end{bmatrix} + \frac{1}{T}\beta_{m,t} + o_p(N_H^{-1/2}),$$

i.e.,

$$\sqrt{N_H}\begin{bmatrix} \hat{\mathcal{H}}_c\hat{g}^C_{m,t} - g^C_{m,t} - \frac{1}{T}\beta^C_{m,t} \\ \hat{\mathcal{H}}_{1,s}\hat{g}^H_{m,t} - \left(g^H_{m,t} - (\bar{g}^{H\prime}\bar{g}^C)(\bar{g}^{C\prime}\bar{g}^C)^{-1}g^C_{m,t}\right) - \frac{1}{T}\beta^H_{m,t} \end{bmatrix} = \left[\Sigma_{\Lambda,1}^{-1} + o_p(1)\right]\left[\frac{1}{\sqrt{N_H}}\Lambda_1' e^H_{m,t}\right] + o_p(1),$$

where $\beta_{m,t} = [\beta^{C\prime}_{m,t}, \beta^{H\prime}_{m,t}]' = -\left(\Sigma_{\Lambda,1}^{-1}\bar{\Phi}_H + B_{\Lambda,1}\right)g_{m,t} + \Sigma_{\Lambda,1}^{-1}\bar{\eta}^2_{1,m,t}\bar{g}_t$. From Assumption D.2 we have $\frac{1}{\sqrt{N_H}}\Lambda_1' e^H_{m,t} \xrightarrow{d} N(0, \Omega^H_{\Lambda,m,m,t})$, ($\mathcal{F}_t$-stably), as well as the joint asymptotic normality of $\frac{1}{\sqrt{N_H}}\Lambda_1' e^H_{m,t}$ and $\frac{1}{\sqrt{N_H}}\Lambda_1' e^H_{m',t}$ for $m \neq m'$. The conclusion follows. ■

D.6.2 Proof of Lemma D.8

Let us first show equation (D.53). We use the asymptotic expansions in Proposition D.4 i) with $f^c_t = \bar{g}^C_t$, $f^s_{1,t} = \bar{g}^H_t$, $f^s_{2,t} = \bar{g}^L_t$, $\beta^c_{1,t} = \beta^C_t$, $\beta^s_{1,t} = \beta^H_t$ and $\varepsilon_{1,i,t} = \bar{e}^{H,i}_t$. From equation (D.21) for $j = 1$ we have in matrix notation:

$$\hat{\Lambda}_1^c = \left\{\Lambda_1^c + \Lambda_1^s(\bar{g}^{H\prime}\bar{g}^C)(\bar{g}^{C\prime}\bar{g}^C)^{-1} + \frac{1}{\sqrt{T}}W_1^c + \frac{1}{T}\left(\Lambda_1^s E[\bar{g}^H_t\beta^{C\prime}_t] - \Lambda_1^c E[\beta^C_t\bar{g}^{C\prime}_t]\right)\right\}\hat{\mathcal{H}}_c + o_p(N_H^{-1/2}),$$

and from (D.22) for $j = 1$ we get:

$$\hat{\Lambda}_1^s = \left\{ \Lambda_1^s + \frac{1}{\sqrt{T}} W_1^s + \frac{1}{T}\left(-\Lambda_1^c E[\beta_t^C \bar{g}_t^{H\prime}] - \Lambda_1^s E[\beta_t^H \bar{g}_t^{H\prime}] \right) \right\} \hat{\mathcal{H}}_{s,1} + o_p(N_H^{-1/2}),$$

where:

$$W_1^c = \frac{1}{\sqrt{T}} \bar{e}^{H\prime} \bar{g}^C (\bar{g}^{C\,\prime} \bar{g}^C / T)^{-1}, \qquad W_1^s = \frac{1}{\sqrt{T}} \bar{e}^{H\prime} \tilde{g}^H (\tilde{g}^{H\,\prime} \tilde{g}^H / T)^{-1},$$

$\bar{e}^H = \left[\bar{e}_1^H, ..., \bar{e}_T^H \right]'$, $\bar{g}^C = \left[\bar{g}_1^C, ..., \bar{g}_T^C \right]'$, $\bar{g}^H = \left[\bar{g}_1^H, ..., \bar{g}_T^H \right]'$ and $\tilde{g}^H = \bar{g}^H - \bar{g}^C (\bar{g}^{C\,\prime} \bar{g}^C)^{-1} (\bar{g}^{C\,\prime} \bar{g}^H)$. Note that $\bar{g}^C$ and $\tilde{g}^H$ are mutually orthogonal in-sample. Thus, we have the expansion:

$$\hat{\Lambda}_1 \hat{\mathcal{U}}_1^{-1} - \Lambda_1 = \frac{1}{\sqrt{T}} G_1 + \Lambda_1 \left(\frac{1}{\sqrt{T}} Q_1 + \frac{1}{T} B_{\Lambda,1} \right) + o_p(N_H^{-1/2}), \tag{D.55}$$

where

$$G_1 = \left[W_1^c \vdots W_1^s \right] = \frac{1}{\sqrt{T}} \bar{e}^{H\prime} \tilde{g} (\tilde{g}' \tilde{g} / T)^{-1}, \qquad \tilde{g} = \left[\bar{g}^C \vdots \tilde{g}^H \right], \tag{D.56}$$

$$Q_1 = \begin{bmatrix} 0_{k^C \times k^C} & 0_{k^C \times k^H} \\ \frac{1}{\sqrt{T}} \tilde{g}^{H\,\prime} \bar{g}^C (\bar{g}^{C\,\prime} \bar{g}^C / T)^{-1} & 0_{k^H \times k^H} \end{bmatrix}, \tag{D.57}$$

$$B_{\Lambda,1} = \begin{bmatrix} -E[\beta_t^C \bar{g}_t^{C\prime}] & -E[\beta_t^C \bar{g}_t^{H\prime}] \\ E[\bar{g}_t^H \beta_t^{C\prime}] & -E[\beta_t^H \bar{g}_t^{H\prime}] \end{bmatrix}, \tag{D.58}$$

$$\mathcal{U}_1 = \begin{bmatrix} \hat{\mathcal{H}}_c & 0 \\ 0 & \hat{\mathcal{H}}_{s,1} \end{bmatrix}. \tag{D.59}$$

From equation (D.55) it follows:

$$\frac{1}{N_H} (\hat{\Lambda}_1 \hat{\mathcal{U}}_1^{-1})' \left(\hat{\Lambda}_1 \hat{\mathcal{U}}_1^{-1} - \Lambda_1 \right) \tag{D.60}$$

$$\begin{aligned} &= \frac{1}{N_H} \left[\Lambda_1 + \frac{1}{\sqrt{T}} G_1 + \Lambda_1 \left(\frac{1}{\sqrt{T}} Q_1 + \frac{1}{T} B_{\Lambda,1} \right) \right]' \left[\frac{1}{\sqrt{T}} G_1 + \Lambda_1 \left(\frac{1}{\sqrt{T}} Q_1 + \frac{1}{T} B_{\Lambda,1} \right) \right] + o_p(N_H^{-1/2}) \\ &= \frac{1}{N_H \sqrt{T}} \Lambda_1' G_1 + \frac{1}{N_H T} G_1' G_1 + \frac{1}{N_H T} (Q_1' \Lambda_1' G_1 + G_1' \Lambda_1 Q_1) + \frac{1}{N_H T^{3/2}} (B_{\Lambda,1}' \Lambda_1' G_1 + G_1' \Lambda_1 B_{\Lambda,1}) \\ &\quad + \frac{1}{\sqrt{T}} \left(\frac{1}{N_H} \Lambda_1' \Lambda_1 \right) Q_1 + \frac{1}{T} \left(\frac{1}{N_H} \Lambda_1' \Lambda_1 \right) B_{\Lambda,1} + \frac{1}{T} Q_1' \left(\frac{1}{N_H} \Lambda_1' \Lambda_1 \right) Q_1 + o_p(N_H^{-1/2}). \end{aligned} \tag{D.61}$$

By using $\Lambda_1' G_1 / \sqrt{N_H} = O_p(1)$, $Q_1 = O_p(1)$ and $\Lambda_1' \Lambda_1 / N_H = \Sigma_{\Lambda,1} + O(N_H^{-1/2})$, we get:

$$\frac{1}{N_H} (\hat{\Lambda}_1 \hat{\mathcal{U}}_1^{-1})' \left(\hat{\Lambda}_1 \hat{\mathcal{U}}_1^{-1} - \Lambda_1 \right) = \frac{1}{\sqrt{T}} \Sigma_{\Lambda,1} Q_1 + \frac{1}{T} \Sigma_{\Lambda,1} B_{\Lambda,1} + \frac{1}{T N_H} G_1' G_1 + \frac{1}{T} Q_1' \Sigma_{\Lambda,1} Q_1 + o_p(N_H^{-1/2}).$$

Let us consider the matrix $\frac{1}{TN_H} G_1' G_1$. The i-th row of matrix G_1 is $\frac{1}{\sqrt{T}} \sum_{t=1}^T \bar{e}_t^{H,i} \tilde{g}_t' (\tilde{g}' \tilde{g} / T)^{-1}$, where $\tilde{g}_t =$

$[\bar{g}_t^{C\,\prime}, \tilde{g}_t^{H\,\prime}]'$. Thus, we have:

$$\begin{aligned}
\frac{1}{N_H} G_1' G_1 &= (\tilde{g}'\tilde{g}/T)^{-1} \frac{1}{N_H} \sum_{i=1}^{N_H} \left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \bar{e}_t^{H,i} \tilde{g}_t \right) \left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \bar{e}_t^{H,i} \tilde{g}_t \right)' (\tilde{g}'\tilde{g}/T)^{-1} \\
&= \frac{1}{N_H} \sum_{i=1}^{N_H} E \left[\left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \bar{e}_t^{H,i} \bar{g}_t \right) \left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \bar{e}_t^{H,i} \bar{g}_t \right)' \right] + O_p(N_H^{-1/2} + T^{-1/2}) \\
&= \bar{\Phi}_H + O_p(N_H^{-1/2} + T^{-1/2}),
\end{aligned}$$

under a summability condition on the covariances. Hence, $\frac{1}{TN_H} G_1' G_1 = \frac{1}{T} \bar{\Phi}_H + o_p(N_H^{-1/2})$ since $T^{-3/2} = o(N_H^{-1/2})$. This yields:

$$\frac{1}{N_H} (\hat{\Lambda}_1 \hat{\mathcal{U}}_1^{-1})' \left(\hat{\Lambda}_1 \hat{\mathcal{U}}_1^{-1} - \Lambda_1 \right) = \frac{1}{\sqrt{T}} \Sigma_{\Lambda,1} Q_1 + \frac{1}{T} \Sigma_{\Lambda,1} B_{\Lambda,1} + \frac{1}{T} \bar{\Phi}_H + \frac{1}{T} Q_1' \Sigma_{\Lambda,1} Q_1 + o_p(N_H^{-1/2}).$$

Let us now consider term $\hat{\mathcal{U}}_1 \left(\frac{\hat{\Lambda}_1' \hat{\Lambda}_1}{N_H} \right)^{-1} \hat{\mathcal{U}}_1'$. From equation (D.55), using arguments similar to Lemma B.8 and equation (B.27), we have:

$$\hat{\mathcal{U}}_1 \left(\frac{\hat{\Lambda}_1' \hat{\Lambda}_1}{N_H} \right)^{-1} \hat{\mathcal{U}}_1' = \Sigma_{\Lambda,1}^{-1} - \frac{1}{\sqrt{T}} Q_1 \Sigma_{\Lambda,1}^{-1} - \frac{1}{\sqrt{T}} \Sigma_{\Lambda,1}^{-1} Q_1' + O_p(N_H^{-1/2} + T^{-1}). \tag{D.62}$$

Then, we get:

$$\begin{aligned}
&\left[\hat{\mathcal{U}}_1 \left(\frac{\hat{\Lambda}_1' \hat{\Lambda}_1}{N_H} \right)^{-1} \hat{\mathcal{U}}_1' \right] \frac{1}{N_H} (\hat{\Lambda}_1 \hat{\mathcal{U}}_1^{-1})' \left(\hat{\Lambda}_1 \hat{\mathcal{U}}_1^{-1} - \Lambda_1 \right) \\
&= \left[\Sigma_{\Lambda,1}^{-1} - \frac{1}{\sqrt{T}} Q_1 \Sigma_{\Lambda,1}^{-1} - \frac{1}{\sqrt{T}} \Sigma_{\Lambda,1}^{-1} Q_1' \right] \left[\frac{1}{\sqrt{T}} \Sigma_{\Lambda,1} Q_1 + \frac{1}{T} \Sigma_{\Lambda,1} B_{\Lambda,1} + \frac{1}{T} \bar{\Phi}_H + \frac{1}{T} Q_1' \Sigma_{\Lambda,1} Q_1 \right] + o_p(N_H^{-1/2}) \\
&= \frac{1}{\sqrt{T}} Q_1 + \frac{1}{T} [\Sigma_{\Lambda,1}^{-1} \bar{\Phi}_H + B_{\Lambda,1}] + o_p(N_H^{-1/2}),
\end{aligned}$$

where we use that N_H and N_L grow at the same rate with $N_H \ll T^3$, and matrix Q_1 is nilpotent. By post-multiplying times $g_{m,t}$, we get equation (D.53).

Let us now show equation (D.54). From (D.55), we have:

$$\begin{aligned}
\frac{1}{N_H} \left(\hat{\Lambda}_1 \hat{\mathcal{U}}_1^{-1} - \Lambda_1 \right)' e_{m,t}^H &= \frac{1}{N_H \sqrt{T}} G_1' e_{m,t}^H + \frac{1}{N_H \sqrt{T}} Q_1' \Lambda_1' e_{m,t}^H + \frac{1}{N_H T} B_{\Lambda,1}' \Lambda_1' e_{m,t}^H \\
&= \frac{1}{N_H \sqrt{T}} G_1' e_{m,t}^H + O_p\left(\frac{1}{\sqrt{N_H T}} \right).
\end{aligned}$$

Let us consider the first term in the r.h.s. We have:

$$\begin{aligned}
\frac{1}{N_H\sqrt{T}}G_1' e_{m,t}^H &= (\tilde{g}'\tilde{g}/T)^{-1}\frac{1}{N_H T}\sum_{i=1}^{N_H}\sum_{s=1}^{T}\bar{e}_s^{H,i}\tilde{g}_s e_{m,t}^{H,i} \\
&= (\tilde{g}'\tilde{g}/T)^{-1}\frac{1}{N_H T}\sum_{i=1}^{N_H}\bar{e}_t^{H,i}\tilde{g}_t e_{m,t}^{H,i} + (\tilde{g}'\tilde{g}/T)^{-1}\frac{1}{N_H T}\sum_{i=1}^{N_H}\sum_{s=1,s\neq t}^{T}\bar{e}_s^{H,i}\tilde{g}_s e_{m,t}^{H,i}.
\end{aligned}$$

Since $\frac{1}{\sqrt{N_H T}}\sum_{i=1}^{N_H}\sum_{s=1,s\neq t}^{T}\bar{e}_s^{H,i}\tilde{g}_s e_{m,t}^{H,i} = O_p(1)$, the second term in the r.h.s. is $O_p(1/\sqrt{N_H T})$. Moreover:

$$\frac{1}{N_H}\sum_{i=1}^{N_H}\bar{e}_t^{H,i}\tilde{g}_t e_{m,t}^{H,i} = \frac{1}{N_H}\sum_{i=1}^{N_H}E[\bar{e}_t^{H,i}e_{m,t}^{H,i}|\mathcal{F}_t]\bar{g}_t + O_p(N_H^{-1/2}+T^{-1/2}) = \bar{\eta}_{1,m,t}^2\bar{g}_t + O_p(N_H^{-1/2}+T^{-1/2}),$$

since $N_H \ll T^3$, and $(\tilde{g}'\tilde{g}/T)^{-1} = I_{k_1} + O_p(T^{-1/2})$. Thus, we get:

$$\frac{1}{N_H}\left(\hat{\Lambda}_1\hat{\mathcal{U}}_1^{-1} - \Lambda_1\right)' e_{m,t}^H = \frac{1}{T}\bar{\eta}_{1,m,t}^2\bar{g}_t + o_p(N_H^{-1/2}).$$

From equation (D.62) we get:

$$\begin{aligned}
&\left[\hat{\mathcal{U}}_1\left(\frac{\hat{\Lambda}_1'\hat{\Lambda}_1}{N_H}\right)^{-1}\hat{\mathcal{U}}_1'\right]\frac{1}{N_H}\left(\hat{\Lambda}_1\hat{\mathcal{U}}_1^{-1} - \Lambda_1\right)' e_{m,t}^H \\
=\ &\left[\Sigma_{\Lambda,1}^{-1} + O_p(N_H^{-1/2}+T^{-1/2})\right]\left[\frac{1}{T}\bar{\eta}_{m,t}^2\bar{g}_t + o_p(N_H^{-1/2})\right] = \frac{1}{T}\Sigma_{\Lambda,1}^{-1}\bar{\eta}_{1,m,t}^2\bar{g}_t + o_p(N_H^{-1/2}),
\end{aligned}$$

since $N_H \ll T^3$. This yields equation (D.54). ■

D.7 Digression on Assumption A.7

In this section we want to show that the conditions in Assumption A.7 hold under mild primitive conditions on the weak serial and cross-sectional dependence of the error terms and factors. We focus especially on cross-sectional dependence. We omit the group index j since it is immaterial for the arguments in this section. Let us denote $\varepsilon_t = (\varepsilon_{1,t}, ..., \varepsilon_{N,t})$ and assume the ε_t are $i.i.d.(0,\Omega)$ across t, with finite fourth-order moments. To simplify the argument, take $h_t = 1$ and $\lambda_i = 1$, so that $\xi_t = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\varepsilon_{i,t}$ and $\alpha_t = \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{s=1,s\neq t}^{T}\varepsilon_{i,t}\varepsilon_{i,s}$ are both scalars, and $\eta_t^2 = \eta^2$ is a constant. [4]

[4] The arguments could be generalized to the case where the errors and factors feature strong mixing serial dependence.

D.7.1 Check of Assumption A.7 a)

Under the above conditions, we have $\eta_{ts}^2 = 0$. Hence we have to show that $E\left[\left(\frac{1}{\sqrt{N}}\sum_{i=1}^N \varepsilon_{i,t}\varepsilon_{i,s}\right)^2\right] \leq M$, for any $s < t$ and a constant M. We have for $s < t$:

$$\begin{aligned}\left[\left(\frac{1}{\sqrt{N}}\sum_{i=1}^N \varepsilon_{i,t}\varepsilon_{i,s}\right)^2\right] &= \frac{1}{N}E\left[(\varepsilon_t'\varepsilon_s)^2\right] = \frac{1}{N}E\left[tr(\varepsilon_t\varepsilon_t'\varepsilon_s\varepsilon_s')\right] \\ &= \frac{1}{N}tr\{\Omega^2\},\end{aligned}$$

by the serial independence condition. Now, $tr\left\{\Omega^2\right\} = \sum_i\sum_j \sigma_{i,j}^4$. Thus, under the weak cross-sectional dependence condition $\frac{1}{N}\sum_i\sum_j \sigma_{i,j}^4 = \frac{1}{N}tr\left\{\Omega^2\right\} = O(1)$, Assumption A.7 a) is met.

D.7.2 Check of Assumption A.7 b)

Let us first show the validity of the bound $E[\alpha_t^2] = O(1)$. We have:

$$\begin{aligned}E[\alpha_t^2] &= \frac{1}{NT}\sum_{r=1,r\neq t}^T\sum_{u=1,u\neq s}^T E[\varepsilon_t'\varepsilon_r\varepsilon_t'\varepsilon_u] = \frac{1}{NT}\sum_{r=1,r\neq t}^T\sum_{u=1,u\neq t}^T trE[\varepsilon_t\varepsilon_t'\varepsilon_r\varepsilon_u'] \qquad \text{(D.63)}\\ &= \frac{1}{NT}\sum_{r=1,r\neq t}^T\sum_{u=1,u\neq t}^T tr(\Omega E[\varepsilon_r\varepsilon_u']) = \frac{1}{NT}\sum_{r=1,r\neq t}^T tr(\Omega^2)\\ &= O\left(\frac{1}{N}tr(\Omega^2)\right). \qquad \text{(D.64)}\end{aligned}$$

Hence, $E[\alpha_t^2] = O(1)$ holds under the weak cross-sectional dependence condition $\frac{1}{N}tr\left\{\Omega^2\right\} = O(1)$.
Let us now show $\frac{1}{T}\sum_{t=1}^T \xi_t\alpha_t = o_p(1)$. Since $E[\xi_t\alpha_t] = 0$, the conclusion follows if $E\left[(\frac{1}{T}\sum_{t=1}^T \xi_t\alpha_t)^2\right] = o(1)$. It is enough to prove $E[(\xi_t\alpha_t)^2] = O(1)$ and $E[\xi_t\alpha_t\xi_s\alpha_s] = o(1)$ for any $t \neq s$, uniformly in t, s. We focus on the second bound (the first one is proved by similar arguments). Let us write $\xi_t = \frac{1}{\sqrt{N}}\iota'\varepsilon_t$ and

$\alpha_t = \frac{1}{\sqrt{NT}} \sum_{r=1,r\neq t}^T \varepsilon_t' \varepsilon_r$, where ι is a $(N,1)$ vector of ones. We have for $t \neq s$:

$$\begin{aligned}
E[\xi_t \alpha_t \xi_s \alpha_s] &= \frac{1}{N^2 T} \sum_{r=1,r\neq t}^T \sum_{u=1,u\neq s}^T E[\varepsilon_t' E(\varepsilon_r \varepsilon_u' | \varepsilon_t, \varepsilon_s) \varepsilon_s \iota' \varepsilon_t \iota' \varepsilon_s] \\
&= \frac{1}{N^2 T} \sum_{r=1,r\neq t,s}^T \sum_{u=1,u\neq t,s}^T E[\varepsilon_t' E(\varepsilon_r \varepsilon_u' | \varepsilon_t, \varepsilon_s) \varepsilon_s \iota' \varepsilon_t \iota' \varepsilon_s] \\
&\quad + \frac{1}{N^2 T} \sum_{u=1,u\neq s}^T E[\varepsilon_t' \varepsilon_s E(\varepsilon_u' | \varepsilon_t, \varepsilon_s) \varepsilon_s \iota' \varepsilon_t \iota' \varepsilon_s] \\
&\quad + \frac{1}{N^2 T} \sum_{r=1,r\neq t}^T E[\varepsilon_t' E(\varepsilon_r | \varepsilon_t, \varepsilon_s) \varepsilon_t' \varepsilon_s \iota' \varepsilon_t \iota' \varepsilon_s] \\
&= \frac{T-2}{N^2 T} E[\varepsilon_t' \Omega \varepsilon_s \iota' \varepsilon_t \iota' \varepsilon_s] + \frac{2}{N^2 T} E[\varepsilon_t' \varepsilon_s \varepsilon_t' \varepsilon_s \iota' \varepsilon_t \iota' \varepsilon_s] \\
&= \frac{T-2}{N^2 T} \iota' \Omega^3 \iota + \frac{2}{N^2 T} tr\left\{ (E[\varepsilon_t \varepsilon_t' (\iota' \varepsilon_t)])^2 \right\}.
\end{aligned}$$

Under a weak cross-sectional dependence condition, namely $\frac{1}{N} \iota' \Omega^3 \iota = O(1)$, the first term in the r.h.s. is $O(N^{-1})$. Under the condition $\frac{1}{N^2} tr\left\{ (E[\varepsilon_t \varepsilon_t' (\iota' \varepsilon_t)])^2 \right\} = O(1)$, the second term is $O(T^{-1})$ (with the latter term vanishing if the distribution of the error terms is symmetric). Hence, $E[\xi_t \alpha_t \xi_s \alpha_s] = o(1)$ for any $t \neq s$, uniformly in t, s.
It remains to prove that $\frac{1}{\sqrt{T}} \sum_{t=1}^T \alpha_t = O_p(1)$. We have:

$$E\left[\left(\frac{1}{\sqrt{T}} \sum_{t=1}^T \alpha_t \right)^2 \right] = \frac{1}{T} E\left[\alpha_t^2\right] + \frac{1}{T} \sum_{t=1}^T \sum_{s=1,s\neq t}^T E\left[\alpha_t \alpha_s\right].$$

The first term in the r.h.s. is $O(T^{-1})$ if $\frac{1}{N} tr\left\{\Omega^2\right\} = O(1)$ (see (D.64)). For the second term, we have:

$$E[\alpha_t \alpha_s] = \frac{1}{TN} \sum_{r=1,r\neq t}^T \sum_{u=1,u\neq s}^T E[\varepsilon_t' \varepsilon_r \varepsilon_s' \varepsilon_u] = \frac{1}{NT} E[\varepsilon_t' \varepsilon_s \varepsilon_s' \varepsilon_t] = \frac{1}{NT} tr(\Omega^2),$$

for $t \neq s$. Hence $E\left[\left(\frac{1}{\sqrt{T}} \sum_{t=1}^T \alpha_t \right)^2 \right] = O(1)$ under the weak cross-sectional dependence condition $\frac{1}{N} tr\left\{\Omega^2\right\} = O(1)$.

D.7.3 Check of Assumption A.7 c)

We have to prove the bounds $E[\|\beta_t\|^2] = O(1)$ and $E[\|\bar{\beta}_t\|^2] = O(1)$, where $\beta_t = \frac{1}{\sqrt{NT}} \sum_{i=1}^N \sum_{s=1,s\neq t}^T \varepsilon_{i,t} (\varepsilon_{i,s} \zeta_{i,s} - E[\varepsilon_{i,s} \zeta_s])$ and $\bar{\beta}_t = \frac{1}{T} \sum_{i=1}^N \sum_{s=1,s\neq t}^T \varepsilon_{i,t} E[\varepsilon_{i,s} \zeta_s]$, for $\zeta_t = (\kappa_t, \xi_t, \alpha_t)'$. We focus on the second bound (the first one is proved using similar arguments). In the above framework we have:

$$\bar{\beta}_t = (1 - T^{-1}) \frac{1}{\sqrt{N}} \sum_{i=1}^N \sum_{j=1}^N \varepsilon_{i,t} \begin{bmatrix} \psi_{i,j} \\ \sigma_{i,j}^2 \\ 0 \end{bmatrix},$$

where $\psi_{i,j} := E[\varepsilon_{i,t}\varepsilon_{j,t}^2]$. Then we have:

$$\begin{aligned} E[\|\bar{\beta}_t\|^2] &= (1-T^{-1})^2\frac{1}{N}\sum_{i=1}^N\sum_{j=1}^N\sum_{k=1}^N\sum_{l=1}^N \sigma_{i,k}^2\psi_{ij}\psi_{kl} + (1-T^{-1})^2\frac{1}{N}\sum_{i=1}^N\sum_{j=1}^N\sum_{k=1}^N\sum_{l=1}^N \sigma_{i,k}^2\sigma_{ij}^2\sigma_{kl}^2 \\ &= (1-T^{-1})^2\left[\frac{1}{N}\iota'(\Psi'\Omega\Psi)\iota + \frac{1}{N}\iota'\Omega^3\iota\right], \end{aligned}$$

where Ψ denotes the matrix with elements $\psi_{i,j}$. Hence, we have $E[\|\bar{\beta}_t\|^2] = O(1)$ under the weak cross-sectional dependence conditions $\frac{1}{N}\iota'(\Psi'\Omega\Psi)\iota = O(1)$ and $\frac{1}{N}\iota'\Omega^3\iota = O(1)$. Note that $\Psi = 0$ if the distribution of $\varepsilon_{i,t}$ is symmetric.

D.8 Estimators based on fixed point iteration

In this Appendix we consider the estimator for group factor models based on the Least Squares (LS) method suggested by Wang (2012). The estimator uses fixed point iteration to solve the first-order conditions. We discuss here some issues concerning the uniqueness of the fixed point.
The group factor model is:

$$Y_1 = F^c\Lambda_1^{c\,\prime} + F_1^s\Lambda_1^{s\,\prime} + \varepsilon_1, \tag{D.65}$$
$$Y_2 = F^c\Lambda_2^{c\,\prime} + F_2^s\Lambda_2^{s\,\prime} + \varepsilon_2. \tag{D.66}$$

For expository purpose, we assume $N_1 = N_2 = N$. The estimators of factor values and factor loadings are defined by minimizing the LS criterion

$$Q = \sum_{j=1}^2 Tr[(Y_j - F^c\Lambda_j^{c\,\prime} - F_j^s\Lambda_j^{s\,\prime})'(Y_j - F^c\Lambda_j^{c\,\prime} + F_j^s\Lambda_j^{s\,\prime})], \tag{D.67}$$

w.r.t. arguments F^c, F_j^s, Λ_j^c, Λ_j^s, $j = 1, 2$, subject to the constraints:

$$F^{c\,\prime}F^c/T = I_{k^c}, \quad F_j^{s\,\prime}F_j^s/T = I_{k_j^s}, \quad F^{c\,\prime}F_j^s = 0, \quad j = 1, 2. \tag{D.68}$$

The first-order conditions (FOC) for this constrained minimization problem yield the following eigenvalue-eigenvector problems (see the proof at the end of this section):

- F_j^s is the $T \times k_j^s$ matrix of standardized eigenvectors of matrix

$$M_{F^c}(Y_jY_j'/N)M_{F^c} \tag{D.69}$$

  associated with the k_j^s largest eigenvalues, for $j = 1, 2$,

- F^c is the $T \times k^c$ matrix of standardized eigenvectors of matrix

$$M_{F^s}(Y_1Y_1'/N + Y_2Y_2'/N)M_{F^s} \tag{D.70}$$

  associated with the k^c largest eigenvalues,

where $M_{F^c} = I_T - F^c(F^{c\,\prime}F^c)^{-1}F^{c\,\prime}$ and $M_{F^s} = I_T - F^s(F^{s\prime}F^s)^{-1}F^{s\,\prime}$, with $F^s = [F_1^s \;\; F_2^s]$. The eigenvectors are normalized such that $F^{c\,\prime}F^c/T = I_{k^c}$, $F_j^{s\,\prime}F_j^s/T = I_{k_j^s}$, for $j = 1, 2$, and satisfy automatically the identification restrictions $F^{c\,\prime}F_j^s = 0$, for $j = 1, 2$.

Wang (2012) suggests to solve the FOC by an iterative procedure. Given an estimate $\tilde{F}^c$, the estimate $\hat{F}_j^s$ is computed by the spectral decomposition of the matrix in (D.69) with $F^c = \tilde{F}^c$, for $j = 1, 2$. The estimate $\hat{F}^s = [\hat{F}_1^s \; \hat{F}_2^s]$ is used to compute the matrix in (D.70), whose spectral decomposition yields a new estimate $\hat{F}^c$. This procedure defines the (stochastic) mapping $\tilde{F}^c \rightarrow \hat{F}^c = \Psi(\tilde{F}^c)$.
Let us now investigate the properties of the mapping Ψ. For this purpose we consider the setting with scalar factors, i.e. $k^c = k_1^s = k_2^s = 1$, and the next assumption.

Assumption D.3. *a) The errors are $\varepsilon_1 = \varepsilon_2 = 0$, and the true factor values are such that $F^{c\prime}F^c/T = F_j^{s\prime}F_j^s/T = 1$, $F^{c\prime}F_j^s = 0$, for $j = 1, 2$. b) $F_1^{s\prime}F_2^s = 0$, $j = 1, 2$.*

Assumption D.3 defines a specific realization of the errors and the factors. In part a), we shut down the errors to mimic the large N, T, setting where the impact of the idiosyncratic shocks vanishes. The factor values match in-sample the theoretical normalization restrictions. For expository purpose, we assume that the group-specific factors are orthogonal, and part b) matches this condition in-sample.

PROPOSITION D.9. *Under Assumption D.3, any vector $\tilde{F}^c$, that is a linear combination of F^c, F_1^s, F_2^s (true factor values), is a fixed point of the mapping Ψ (up to a sign change).*

Thus, the set of fixed points of Ψ includes the three-dimensional linear space spanned by vectors F^c, F_1^s, F_2^s.

Proof of Proposition D.9: Define the $T \times 3$ matrix $H = [F^c \;\; F_1^s \;\; F_2^s]$. Under Assumption D.3 we have $H'H/T = I_3$, and the data can be written as $Y_1 = H[\Lambda_1^c \;\; \Lambda_1^s \;\; 0]'$ and $Y_2 = H[\Lambda_2^c \;\; 0 \;\; \Lambda_2^s]'$. Then, we get:

$$Y_1 Y_1'/N = H \begin{bmatrix} \Lambda_1^{c\prime}\Lambda_1^c/N & \Lambda_1^{c\prime}\Lambda_1^s/N & 0 \\ \Lambda_1^{s\prime}\Lambda_1^c/N & \Lambda_1^{s\prime}\Lambda_1^s/N & 0 \\ 0 & 0 & 0 \end{bmatrix} H' \equiv H\Pi_1 H'.$$

Similarly, we have $Y_2 Y_2'/N = H\Pi_2 H'$ for a suitable 3×3 matrix Π_2, and $Y_1 Y_1'/N + Y_2 Y_2'/N = H\Pi H'$ with $\Pi = \Pi_1 + \Pi_2$.
Now, let

$$\tilde{F}^c = F^c \beta_1 + F_1^s \beta_2 + F_2^s \beta_3 = H\beta,$$

where the 3×1 vector $\beta = (\beta_1, \beta_2, \beta_3)'$ is such that $\beta'\beta = 1$. Then:

$$\begin{aligned} M_{\tilde{F}^c} &= I_T - \frac{1}{T}\tilde{F}^c \tilde{F}^{c\prime} \\ &= I_T - \frac{1}{T} H\beta\beta' H' = M_H + \frac{1}{T} H M_\beta H', \end{aligned}$$

where $M_\beta = I_3 - \beta(\beta'\beta)^{-1}\beta' = I_3 - \beta\beta'$. Therefore, the matrix in (D.69) corresponding to $\tilde{F}^c$ can be written as:

$$\begin{aligned} M_{\tilde{F}^c}(Y_j Y_j'/N) M_{\tilde{F}^c} &= (M_H + \frac{1}{T} H M_\beta H') H\Pi_j H' (M_H + \frac{1}{T} H M_\beta H') \\ &= H M_\beta \Pi_j M_\beta H'. \end{aligned}$$

The eigenvector associated with the largest eigenvalue of matrix $H M_\beta \Pi_j M_\beta H'$ is in the column space of H:

$$\hat{F}_j^s = H\alpha_j,$$

where the 3×1 vector α_j is the normalized eigenvector of matrix $M_\beta \Pi_j M_\beta$ associated with the largest eigenvalue, $j = 1, 2$. In particular, α_j is orthogonal to β, $j = 1, 2$. The vectors α_1 and α_2 are not collinear.

Let $\hat{F}^s = [\hat{F}_1^s \ \hat{F}_2^s] = H\alpha$, where $\alpha = [\alpha_1 \ \alpha_2]$. The matrix in (D.70) corresponding to $\hat{F}^s$ is

$$M_{\hat{F}^s}(Y_1Y_1'/N + Y_2Y_2'/N)M_{\hat{F}^s} = HM_\alpha \Pi M_\alpha H'.$$

The eigenvector $\hat{F}^c$ of this matrix associated with the largest eigenvalue is $\hat{F}^c = H\gamma$, where γ is the eigenvector of matrix $M_\alpha \Pi M_\alpha$ associated with the largest eigenvalue. This implies that γ is orthogonal to α_1 and α_2, and thus is collinear to β. By normalization we have either $\gamma = \beta$, or $\gamma = -\beta$. Therefore, either $\hat{F}^c = \tilde{F}^c$, or $\hat{F}^c = -\tilde{F}^c$. ■

Proof of the FOC for the constrained minimization (D.67)-(D.68): The Lagrange multipliers for the identification restrictions (D.68) are zero. The FOC for the factor loadings under the constraints yield:

$$\begin{aligned} \Lambda_j^c &= Y_j'F^c(F^{c\,\prime}F^c)^{-1}, \\ \Lambda_j^s &= Y_j'F_j^s(F_j^{s\,\prime}F_j^s)^{-1}, \quad j = 1, 2. \end{aligned}$$

From these equations, the residuals are

$$Y_j - F^c\Lambda_j^{c\,\prime} - F_j^s\Lambda_j^{s\,\prime} = (I_T - P_{F^c} - P_{F_j^s})Y_j, \quad j = 1, 2,$$

where $P_{F^c} = F^c(F^{c\,\prime}F^c)^{-1}F^{c\,\prime} = I_T - M_{F^c}$ and $P_{F_j^s} = F_j^s(F_j^{s\,\prime}F_j^s)^{-1}F_j^{s\,\prime} = I_T - M_{F_j^s}$. From the orthogonality $F^{c\,\prime}F_j^s = 0$ in (D.68), matrices M_{F^c} and $M_{F_j^s}$ commute, and matrices

$$I_T - P_{F^c} - P_{F_j^s} = M_{F_j^s}M_{F^c} = M_{F^c}M_{F_j^s}, \quad j = 1, 2,$$

are idempotent. Therefore, the concentrated LS criterion becomes:

$$Q = \sum_{j=1}^{2} Tr[Y_j'M_{F^c}M_{F_j^s}Y_j]. \tag{D.71}$$

From the constraints (D.68) and the commutative property of the trace, the concentrated LS criterion can be rewritten as:

$$\begin{aligned} Q &= \sum_{j=1}^{2} Tr[M_{F_j^s}M_{F^c}Y_jY_j'M_{F^c}] \\ &= \sum_{j=1}^{2} Tr[M_{F^c}Y_jY_j'M_{F^c}] - \sum_{j=1}^{2}\frac{1}{T}Tr[F_j^{s\,\prime}M_{F^c}Y_jY_j'M_{F^c}F_j^s]. \end{aligned}$$

For $j = 1, 2$, the minimization of this concentrated criterion w.r.t. F_j^s is equivalent to the maximization of $Tr[F_j^{s\,\prime}M_{F^c}Y_jY_j'M_{F^c}F_j^s]$. Under the constraint $F^{s\prime}F_j^s/T = I_{k_j^s}$, this problem is solved by the matrix of normalized eigenvectors of matrix $M_{F^c}Y_jY_j'M_{F^c}$ associated with the k_j^s largest eigenvalues.
Similarly, from the constraints (D.68) and the commutative property of the trace, the concentrated LS criterion

(D.71) can be rewritten as:

$$\begin{aligned}
Q &= \sum_{j=1}^{2} Tr[M_{F^c} M_{F_j^s} Y_j Y_j' M_{F_j^s}] \\
&= \sum_{j=1}^{2} Tr[M_{F_j^s} Y_j Y_j' M_{F_j^s}] - \sum_{j=1}^{2} \frac{1}{T} Tr[F^{c\prime} M_{F_j^s} Y_j Y_j' M_{F_j^s} F^c] \\
&= \sum_{j=1}^{2} Tr[M_{F_j^s} Y_j Y_j' M_{F_j^s}] - \frac{1}{T} Tr[F^{c\prime} M_{F^s} (\sum_{j=1}^{2} Y_j Y_j') M_{F^s} F^c].
\end{aligned}$$

The minimization of this concentrated criterion w.r.t. F^c is equivalent to the maximization of $Tr[F^{c\prime} M_{F^s}(\sum_{j=1}^{2} Y_j Y_j') M_{F^s} F^c]$. Under the constraint $F^{c\prime} F^c/T = I_{k^c}$, this problem is solved by the matrix of normalized eigenvectors of matrix $M_{F^s}(\sum_{j=1}^{2} Y_j Y_j') M_{F^s}$ associated with the k^c largest eigenvalues. ■

D.9 Practical implementation of the procedure

Let us first assume that k^C, k^H, k^L, i.e. the number of respectively common, high and low frequency factors in equation (5.1), are known and are all strictly larger than zero. The identification strategy presented in Section 2 directly implies a simple three-step estimation procedure for the factor values and the factor loadings (see Section 3), which is summarized here for practical implementation purposes:

1. *PCA performed on the HF and LF panels separately*
   Define the (T, N_H) matrix of temporally aggregated (in our application flow-sampled) demeaned HF observables as $X^H = [x_1^H, ..., x_T^H]'$, and the (T, N_L) matrix of demeaned LF observables as $X^L = [x_1^L, ..., x_T^L]'$. 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:

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

   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:

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

   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 $\hat{W}_H^C$ be the $(k^C + k^H, k^C)$ matrix whose columns are the canonical directions for $\hat{h}_{H,t}$ associated with the k^C largest canonical correlations between $\hat{h}_H$ and $\hat{h}_L$. Then, an estimator of the (in our application flow-sampled) common factor is $\hat{\underline{g}}_t^C = \hat{W}_H^{C\prime} \hat{h}_{H,t}$, for $t = 1, ..., T$. Analogously, $\hat{\underline{g}}_t^{C*} = \hat{W}_L^{C\prime} \hat{h}_{L,t}$, for $t = 1, ..., T$, where $\hat{W}_L^C$ is the $(k^C + k^L, k^C)$ matrix of the canonical directions for $\hat{h}_{L,t}$.

   As explained in Subsection D.2, an alternative estimator of the flow-sampled common factor values $\hat{\underline{g}}_t^{C\star}$, $t = 1, ..., T$, is obtained from the eigenvectors associated to the k^C largest eigenvalues of matrix

$\frac{1}{T}(\hat{h}_H\hat{h}_H' + \hat{h}_L\hat{h}_L')$. [5] The rest of the estimation procedure can be performed replacing $\hat{\hat{g}}_t^C$ with $\hat{\hat{g}}_t^{C\star}$.

The estimated loadings matrices $\hat{\Lambda}_{HC}$ and $\hat{\Lambda}_{LC}$ are obtained from the least squares regressions of x_t^H and x_t^L on estimated factor $\hat{\hat{g}}_t^C$. Collect the residuals of these regressions:

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

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

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

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

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

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

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

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 $\hat{\Lambda}_L$ are obtained from the least squares regression of $\hat{\xi}_t^H$ and $\hat{\xi}_t^L$ on respectively the estimated factors $\hat{\hat{g}}_t^H$ and $\hat{\hat{g}}_t^L$.

3. *Reconstruction of the common and high frequency-specific factors*
   The estimates of the common and HF 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.

As discussed in Section 5, an alternative estimation approach consists in performing PCA prior to aggregation. In this case, the first step in the above procedure is modified as follows:

1.' *PCA performed on the HF and LF panels separately prior to aggregation*
   Define the (TM, N_H) matrix of HF observables as $X^{HF} = [x_{1,1}^H, ..., x_{M,1}^H, ..., x_{M,T}^H]'$. We perform PCA on the HF data:

$$\left(\frac{1}{TMN_H}X^{HF}X^{HF\prime}\right)\check{h}_{HF} = \check{h}_{HF}\check{V}_{HF},$$

   where $\check{h}_{HF}$ is the $(TM, k^C + k^H)$ matrix of eigenvectors, and $\check{V}_{HF}$ is the diagonal matrix of the eigenvalues, of matrix $(TMN_H)^{-1}X^{HF}X^{HF\prime}$. Then, the estimated flow-sampled values of the pervasive factors in the HF panel are collected in the $(T, k^C + k^H)$ matrix $\hat{h}_H = [\hat{h}_{H,1}, ..., \hat{h}_{H,T}]$ where $\hat{h}_{H,t} = \sum_{m=1}^M \check{h}_{HF,m,t}$. The estimated pervasive factors in the LF panel are obtained as in step 1 above.

The other steps 2 and 3 are unchanged.
Since the factors dimensions are unknown, the aforementioned procedure is implemented with estimated factors dimensions $\hat{k}^C$, $\hat{k}^H$, and $\hat{k}^L$. Inference on the number of common, low and high-frequency-specific factors proceeds as follows:

[5] As shown in Subsection D.2, we have $\hat{\hat{g}}_t^{C\star} = \frac{1}{2}(\hat{\hat{g}}_t^C + \hat{\hat{g}}_t^{C*})$.

1. Estimate $k_1 = k^C + k^H$ and $k_2 = k^C + k^L$, i.e. the numbers of pervasive factors in panels X^H and X^L, by some consistent estimators, as the IC_{p1} and IC_{p2} criteria of Bai and Ng (2002).

2. Let $\underline{k} := \min(\hat{k}_1, \hat{k}_2)$. Test sequentially:

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

for any given $r = \underline{k},\ \underline{k} - 1,\ \ldots,\ 1$. We use the statistic $\tilde{\xi}(r)$ defined in equation (4.6), which is based on $\hat{\xi}(r) = \sum_{\ell=1}^{r} \hat{\rho}_\ell$, where the $\hat{\rho}_\ell$, for $\ell = 1, \ldots, r$, are the r largest canonical correlations between $\hat{h}_{H,t}$ and $\hat{h}_{L,t}$. Here, $\hat{h}_{H,t}$ and $\hat{h}_{L,t}$ are the first $\hat{k}_{X^H}$ and $\hat{k}_{X^L}$ PCs extracted from the X^H and X^L panels, respectively, and the canonical correlations are the squared roots of the eigenvalues of matrix $\hat{R}$ defined in equation (3.1). We reject $H_0 = H(r)$ if $\tilde{\xi}(r) < z_{\alpha_{NT}}$, where critical value $z_{\alpha_{NT}}$ is set as in equation (4.7), with $\gamma = 0.1$ and constant $c = 0.95$ as in the Monte Carlo study. Estimate $\hat{k}^C$ is the largest dimension r such that H_0 is not rejected, or $\hat{k}^C = 0$ if H_0 is rejected for all r.

3. The dimensions of frequency-specific factors are obtained by difference: $\hat{k}^H = \hat{k}_1 - \hat{k}^C$, and $\hat{k}^L = \hat{k}_2 - \hat{k}^C$.

D.9.1 Implementation choices

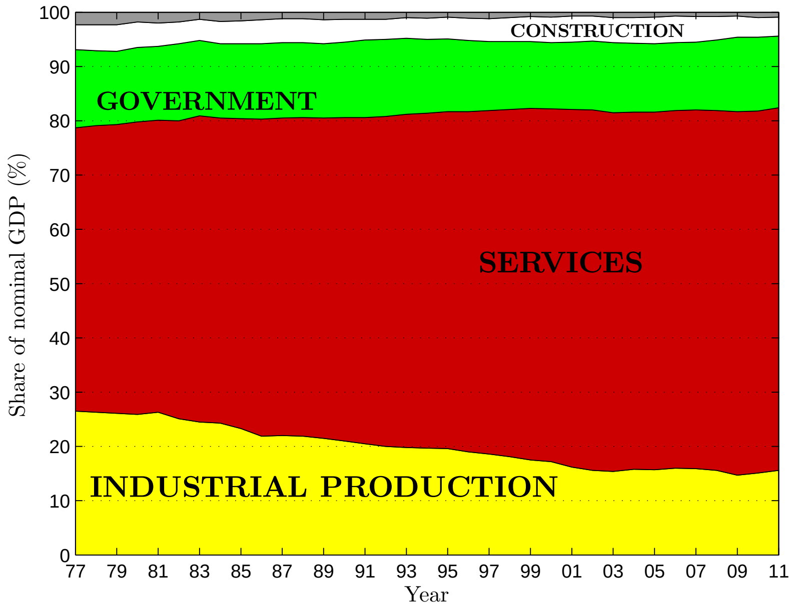
The PCA in step 1 (or step 1') in the estimation procedure is performed on demeaned and standardized data. Hence, for step 1 the demeaned observables x_t^{Hi} and x_t^{Li} are replaced by x_t^{Hi}/σ^{Hi} and x_t^{Li}/σ^{Li}, where σ^{Hi} and σ^{Li} are the sample standard deviations of the HF and LF series. For step 1' the HF observables are standardized at high-frequency.

In the empirical analysis in Section 7, we use $\hat{\hat{g}}_t^{C\star}$ to estimate the common factor values and perform PCA on the flow-sampled HF data. To compute the test statistics $\tilde{\xi}(r)$ in Theorem 2 we need the estimator $\hat{\Sigma}_U$ of matrices Σ_U and $\tilde{\Sigma}_U$. The estimator proposed in Theorem 2 is valid under the assumption of uncorrelated errors within and between each of the two panels. Another estimator which takes into account (contemporaneous) weak cross sectional correlations in the errors is $\hat{\Sigma}_U^* = (N_2/N_1)\hat{\Sigma}_{u,11}^{(cc)} + \hat{\Sigma}_{u,22}^{(cc)} - \sqrt{N_2/N_1}\hat{\Sigma}_{u,12}^{(cc)} - \sqrt{N_2/N_1}\hat{\Sigma}_{u,21}^{(cc)}$, where $\hat{\Sigma}_{u,ij} = \left(\frac{1}{N_i}\hat{\Lambda}_i'\hat{\Lambda}_i\right)^{-1}\left(\frac{1}{\sqrt{N_i N_j}}\hat{\Lambda}_i'\hat{\Gamma}_{ij}\hat{\Lambda}_j\right)\left(\frac{1}{N_j}\hat{\Lambda}_j'\hat{\Lambda}_j\right)^{-1}$ for $i, j = 1, 2$, with estimated loadings $\hat{\Lambda}_1 = [\hat{\Lambda}_{HC} \vdots \hat{\Lambda}_H]$ and $\hat{\Lambda}_2 = [\hat{\Lambda}_{LC} \vdots \hat{\Lambda}_L]$, and $N_1 = N_H$ and $N_2 = N_L$. Moreover, $\hat{\Gamma}_{ij}$ is an estimator of the cross sectional covariance matrix of the panels of residuals $\hat{\hat{e}}^i$ and $\hat{\hat{e}}^j$ with $i, j = 1, 2$, using thresholding, where $\hat{\hat{e}}^1$ (resp. $\hat{\hat{e}}^2$) is the (T, N_H) (resp. (T, N_L)) panel of residuals obtained by regressing each series x_t^{Hi} (resp. x_t^{Li}) in panel X^H(resp. X^L) on both $\hat{\hat{g}}_t^C$ and $\hat{\hat{g}}_t^H$ (resp. $\hat{\hat{g}}_t^L$) factors. In the empirical application we use the second estimator $\hat{\Sigma}_U^*$.

D.10 Dataset description

Figure D.1 shows that the share of the Industrial Production (IP) sector in the US economy has been in decline since the late 70's, which marks the beginning of our sample period. We use the class of mixed frequency group factor model as well as the test of common versus group specific factors introduced in this paper to shed light on the key question of interest, namely whether, despite the shrinking size of the IP sectors, the factors related to IP are still dominant determinants of US output fluctuations.

Figure D.1: Sectoral decomposition of US nominal GDP.



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).

D.10.1 High Frequency dataset: Industrial Production sectors

Our high frequency dataset includes the same 117 sectors constituting the aggregate Industrial Production index considered by Foerster, Sarte, and Watson (2011) for the years 1977-2011. This sample period coincides with the maximum number of years for which the data for the 42 non-Industrial Production sectors of our low frequency panel were available and therefore - differently from Foerster, Sarte, and Watson (2011) - we do not consider the entire time series available for IP data starting in 1972. We download the monthly level of the 117 IP indices from the Board of Governors of the Federal Reserve System (FED)[6]. From these raw data, which are indices of real output, we compute the corresponding quarterly growth rates.

The 117 sectors roughly correspond to a four-digit industry in the North American Industry Classification System (NAICS) for year 2002. The IP sectors are classified by the FED according in the following subsectors: *Manufacturing*, *Mining* and *Utilities*. *Manufacturing* comprises those industries included in NAICS definition of manufacturing plus the logging and newspaper, periodical, book, and directory publishing industries that have traditionally been considered manufacturing, and is divided in *Durable*, *Nondurable* and *Other manufacturing*. *Durable manufacturing* includes three-digit NAICS codes 321, 327, 331-337, and 339. *Nondurable manufacturing* includes three-digit NAICS codes 311-316 and 322-326. *Other manufacturing* includes NAICS codes 1133 and 5111. *Mining* includes three-digit NAICS codes 211-213. *Utilities* include electric utilities and natural gas distribution, corresponding to NAICS codes 2211 and 2212.[7] We refer to Foerster, Sarte, and Watson (2011),

[6] See `http://www.federalreserve.gov/releases/G17/default.htm`.

[7] For a detailed description of the IP constituents see `http://www.federalreserve.gov/releases/g17/About.htm` .

and especially to their Appendix A, for a list of the names of the 117 sectors, and the methodology used to approximate the missing data for some sectors.

Table D.1: List of non-Industrial Production sectors. (Source: BEA)

Sector	**NAICS 2002 codes**
Farms	111, 112
Forestry, fishing, and related activities	113, 114, 115
Construction	23
Wholesale trade	42
Retail trade	44, 45
Air transportation	481
Rail transportation	482
Water transportation	483
Truck transportation	484
Transit and ground passenger transportation	485
Pipeline transportation	486
Other transportation and support activities	487, 488, 492
Warehousing and storage	493
Publishing industries (includes software)	511, 516
Motion picture and sound recording industries	512
Broadcasting and telecommunications	515, 517
Information and data processing services	518, 519
Federal Reserve banks, credit intermediation, and related activities	521, 522
Securities, commodity contracts, and investments	523
Insurance carriers and related activities	524
Funds, trusts, and other financial vehicles	525
Real estate	531
Rental and leasing services and lessors of intangible assets	532, 533
Legal services	5411
Computer systems design and related services	5415
Miscellaneous professional, scientific, and technical services	5412-5414, 5416-5419
Management of companies and enterprises	55
Administrative and support services	561
Waste management and remediation services	562
Educational services	61
Ambulatory health care services	621
Hospitals and nursing and residential care facilities	622, 623
Social assistance	624
Performing arts, spectator sports, museums, and related activities	711, 712
Amusements, gambling, and recreation industries	713
Accommodation	721
Food services and drinking places	722
Other services, except government	81
Federal Government - General government	-
Federal Government - Government enterprises	- (includes 491)
State and Local Government - General government	-
State and Local Government - Government enterprises	-

D.10.2 Low Frequency dataset: non-Industrial Production sectors

The US Bureau of Economic Analysis (BEA) publishes at yearly frequency the growth rates for the real Gross Domestic Product and real Gross Output for all the sectors of the US economy, not only for the sectors included in the IP index. We use the *Release Date November 13, 2012* dataset as downloaded for the BEA website[8]. The period 1977-2011 coincides with the maximum number of years for which the data for the 42 non-Industrial

[8] See `http://www.bea.gov/industry/gdpbyind_data.htm.`

Production sectors in our low frequency panel were available at the date of download of the dataset. [9] Our original BEA dataset includes the time series for the output growth rates of 65 mutually exclusive sectors constituting the entire US economy, for the sample period 1977-2011. These sectors are aggregates of either 2 or 3 digits 2002 NAICS codes. Out of these 65 sectors, 19 are *Manufacturing* sectors (NAICS 2002 codes: 31-33), 3 are *Mining* sectors (NAICS 2002 codes: 211-213) and one is *Utilities* (NAICS 2002 code: 22). These 23 sectors are all included in the IP dataset, and therefore are excluded from our LF panel to avoid duplication of sectors in the two panels. The IP sectors *Logging*, *Newspaper Publishers* and *Periodical, Book, and Other Publishers* (NAICS 1133, 5111, 5112) are subsectors of the 2 BEA sectors *Publishing industries (includes software)* and *Forestry, fishing, and related activities*. We keep these 2 sectors in the low frequency panel. Therefore our non-IP low frequency panel includes the 42 sectors listed in Table D.1 together with the corresponding NAICS 2002 codes. In Table D.2 we report the names of the sectors corresponding to the aggregated version of the yearly indices used in Table III, and in the analogous tables in the subsample analysis (Section D.11.4), together with their corresponding first or first two NAICS 2002 codes. The yearly growth rates of these real aggregated indices are downloaded from the BEA website.

Table D.2: List of aggregates of non-Industrial Production sectors. (Source: BEA)

Sector	**NAICS 2002 codes**
GDP (all sectors)	all sectors included
Manufacturing	31, 32, 33
Agriculture, forestry, fishing, and hunting	11
Construction	23
Wholesale trade	42
Retail trade	44, 45
Transportation and warehousing	48, 49 (except 491)
Information	51
Finance, insurance, real estate, rental, and leasing	52, 53
Professional and business services	54
Educational services, health care, and social assistance	6
Arts, entertainment, recreation, accommodation, and food services	7
Government	- (includes 491)

D.11 Additional empirical results

This section collects supplemental empirical results to the ones presented in Section 7 of the paper. It is divided into four parts: Subsection D.11.1 gives the results of a "Granularity analysis" analogous to the one in Section II.B of Foerster, Sarte, and Watson (2011) - performed on our dataset. Subsection D.11.2 covers empirical results analogous to Section 7.2 obtained with an alternative estimation and inference method, in which PCs are extracted from the HF panel prior to aggregation. Subsection D.11.3 provides additional tables and figures to the ones in Section 7.2. Subsection D.11.4 reports the results of the empirical analysis performed on different sub-samples of the one considered in Section 7.

[9] Time series for 22 aggregates of our 42 sectors are also available from the BEA website since 1947, and time series for a more disaggregated version of our 42 indices, but only for Gross Output, is available only from 1997.

D.11.1 Granularity analysis

In this section we report the results of the same analysis, performed on our dataset, as the one in Section II.B of Foerster, Sarte, and Watson (2011). Our objective is to rule out the possibilities that *a)* sectoral weights in GDP and IP aggregate indexes are the major determinants in explaining the variability of the indexes themselves, and *b)* that their aggregate variability is driven mainly by sector-specific variability. Specifically, we replicate the analysis of Tables 1, 2, and 3 in Foerster, Sarte, and Watson (2011) for the growth rates of the sectoral components of the following four different indices:

1. The quarterly IP index, decomposed in the same 117 different sectoral indices as in Foerster, Sarte, and Watson (2011). The weights of each sector used in the share-weight decomposition correspond to the sectoral weights in the IP index. Results are displayed in Tables D.3 - D.5.

2. The annual GDP index, decomposed in all the 65 different sectoral indices produced by the BEA. The weights of each sector used in the share-weight decomposition correspond to the sectoral weights in the GDP index as produced by the BEA. Results are displayed in Tables D.6 - D.8.

3. A new synthetic annual Manufacturing index, which we created from the 19 different sectoral GDP indices labeled as "Manufacturing" (NAICS 2002 code: 22), produced by the BEA. The weights of each sector used in the share-weight decomposition correspond to the sectoral weights in the GDP index as produced by the BEA. Results are displayed in Tables D.9 - D.11.

4. A new synthetic annual non-IP index, which we created from the 42 different sectoral non-IP indices listed in Table D.1, produced by the BEA. The weights of each sector used in the share-weight decomposition correspond to the sectoral weights in the GDP index as produced by the BEA. Results are displayed in Tables D.12 - D.14.

The notation, and the formulas used to produce the results in Tables D.3 - D.14, are the same as those used in Tables 1, 2, and 3 in Foerster, Sarte, and Watson (2011). The time series of the four components of the share-weight decomposition of Tables D.3, D.6, D.9, and D.12, are displayed in Figures D.2, D.3, D.4, and D.5, respectively. Finally, Figures D.6 and D.7 display the histograms of the standard deviations of the growth rates of the 117 IP indices (quarterly), and 42 non-IP indices (annual), computed over the four different sample periods considered in the empirical analysis. In the captions of Figures D.6 and D.7 we report tables displaying the 25%, 50%, and 75% quantiles of the empirical distributions represented by the histograms.
Overall, the results support our objectives and provide evidence to rule out the aforementioned possibilities *a)* and *b)*, as mentioned above.

Quarterly IP Index

Table D.3: Share weight decomposition of quarterly Industrial Production index.

Series	1977-2011	1977-1983	1984-2007	2008-2011
$g_t = \sum w_{it} x_{it}$	5.7	7.9	3.6	9.7
$(1/N) \sum x_{it}$	7.3	9.9	4.2	13.7
$\sum(\bar{w}_i - (1/N)) x_{it}$	2.2	3.3	1.5	2.9
$\sum(w_{it} - \bar{w}_i) x_{it}$	1.0	1.0	0.5	2.1

Entries are the sample standard deviations of the quarterly growth rates of the quarterly Industrial Production index growth (g_t) and its components ($x_{i,t}$). Percentage points are at annual rates. The table corresponds to Table 1 in Foerster, Sarte, and Watson (2011).

Table D.4: Average pairwise correlations of sectoral Industrial Production indices.

1977-2011	1977-1983	1984-2007	2008-2011
0.21	0.25	0.12	0.34

Entries are the average pairwise sample correlations of the quarterly growth rates of the 117 Industrial Production indices considered in the paper. The table corresponds to Table 2 in Foerster, Sarte, and Watson (2011).

Table D.5: Standard deviation of aggregate Industrial Production indices constructed with and without sectoral covariance

	1977-2011	1977-1983	1984-2007	2008-2011
	A. Using Actual w_{it} Share Weights			
With sectoral covariation	5.7	7.9	3.6	9.7
Without sectoral covariation	1.9	2.5	1.6	2.5
	B. Using Equal $(1/N)$ Share Weights			
With sectoral covariation	7.3	9.9	4.2	13.7
Without sectoral covariation	1.9	2.7	1.4	2.4

The entries for rows labeled "with sectoral covariation" are sample standard deviations of $\sum w_{it} x_{it}$ (Panel A) and $N^{-1} \sum x_{it}$ (Panel B). The entries labeled "without sectoral covariation" are computed as: $\sqrt{T^{-1} \sum_t \sum_i h_{it}^2 (x_{it} - \bar{x}_i)^2}$, where $h_{it} = w_{it}$ in panel A and $h_{it} = N^{-1}$ in panel B. Percentage points are at annual rates. The table corresponds to Table 3 in Foerster, Sarte, and Watson (2011).

Annual GDP sectoral indices (all sectors)

Table D.6: Share weight decomposition of annual GDP index.

Series (GDP)	1977-2011	1977-1983	1984-2007	2008-2011
$g_t = \sum w_{it} x_{it}$	2.1	2.6	1.4	2.3
$(1/N) \sum x_{it}$	2.7	3.4	1.7	5.2
$\sum(\bar{w}_i - (1/N))x_{it}$	0.9	1.0	0.8	1.8
$\sum(w_{it} - \bar{w}_i)x_{it}$	0.4	0.2	0.2	1.2

Entries are the sample standard deviations of the annual growth rates of annual GDP index growth (g_t) and its components ($x_{i,t}$). The index is constructed using weights of nominal GDP. The table is the analogous of Table 1 in Foerster, Sarte, and Watson (2011) for GDP data.

Table D.7: Average pairwise correlations of sectoral GDP indices.

1977-2011	1977-1983	1984-2007	2008-2011
0.18	0.29	0.11	0.19

Entries are the average pairwise sample correlations of the annual growth rates of the 65 sectoral GDP indices. The table is the analogous of Table 2 in Foerster, Sarte, and Watson (2011) for GDP data.

Table D.8: Standard deviation of aggregate GDP indices constructed with and without sectoral covariance

	1977-2011	1977-1983	1984-2007	2008-2011
	A. Using Actual w_{it} Share Weights			
With sectoral covariation	2.1	2.6	1.4	2.3
Without sectoral covariation	0.8	0.9	0.7	1.0
	B. Using Equal $(1/N)$ Share Weights			
With sectoral covariation	2.7	3.4	1.7	5.2
Without sectoral covariation	1.2	1.1	1.0	1.9

The entries for rows labeled "with sectoral covariation" are sample standard deviations of $\sum w_{it} x_{it}$ (Panel A) and $N^{-1} \sum x_{it}$ (Panel B). The entries labeled "without sectoral covariation" are computed as: $\sqrt{T^{-1} \sum_t \sum_i h_{it}^2 (x_{it} - \bar{x}_i)^2}$, where $h_{it} = w_{it}$ in panel A and $h_{it} = N^{-1}$ in panel B. The table is the analogous of Table 3 in Foerster, Sarte, and Watson (2011) for GDP data.

Annual Manufacturing sectors in GDP index

Table D.9: Share weight decomposition of aggregate Manufacturing index.

Series(GDP)	1977-2011	1977-1983	1984-2007	2008-2011
$g_t = \sum w_{it} x_{it}$	4.5	5.8	3.5	7.6
$(1/N) \sum x_{it}$	5.3	6.9	3.4	10.9
$\sum(\bar{w}_i - (1/N))x_{it}$	1.2	0.9	1.1	2.2
$\sum(w_{it} - \bar{w}_i)x_{it}$	1.6	0.6	0.4	4.8

Entries are the sample standard deviations of the annual growth rate and the components of the annual index (g_t) created from the 19 Manufacturing sectors in the GDP index. The index is constructed using weights of nominal GDP. The table is the analogous of Table 1 in Foerster, Sarte, and Watson (2011) for Manufacturing data.

Table D.10: Average pairwise correlations of sectoral Manufacturing indices.

1977-2011	1977-1983	1984-2007	2008-2011
0.35	0.48	0.27	0.29

Entries are the average pairwise sample correlations of the annual growth rates of the 19 Manufacturing sectors in the GDP index. The table is the analogous of Table 2 in Foerster, Sarte, and Watson (2011) for Manufacturing data.

Table D.11: Standard deviation of aggregate Manufacturing indices constructed with and without sectoral covariance

	1977-2011	1977-1983	1984-2007	2008-2011
	A. Using Actual w_{it} Share Weights			
With sectoral covariation	4.5	5.8	3.5	7.6
Without sectoral covariation	2.7	2.6	2.3	3.9
	B. Using Equal $(1/N)$ Share Weights			
With sectoral covariation	5.3	6.9	3.4	10.9
Without sectoral covariation	2.8	2.9	1.9	5.3

The entries for rows labeled "with sectoral covariation" are sample standard deviations of $\sum w_{it} x_{it}$ (Panel A) and $N^{-1} \sum x_{it}$ (Panel B). The entries labeled "without sectoral covariation" are computed as: $\sqrt{T^{-1} \sum_t \sum_i h_{it}^2 (x_{it} - \bar{x}_i)^2}$, where $h_{it} = w_{it}$ in panel A and $h_{it} = N^{-1}$ in panel B. The table is the analogous of Table 3 in Foerster, Sarte, and Watson (2011) for Manufacturing data.

Annual non-IP sectors in GDP sectoral indices

Table D.12: Share weight decomposition of aggregate index of non-IP sectors.

Series(GDP)	1977-2011	1977-1983	1984-2007	2008-2011
$g_t = \sum w_{it}x_{it}$	1.7	2.1	1.3	1.9
$(1/N)\sum x_{it}$	2.1	2.6	1.3	3.4
$\sum(\bar{w}_i - (1/N))x_{it}$	0.9	0.5	1.0	1.5
$\sum(w_{it} - \bar{w}_i)x_{it}$	0.2	0.3	0.2	0.2

Entries are the sample standard deviations of the annual growth rate and the components of the annual index (g_t) created from the 42 non-IP sectors in the GDP index considered in our paper. The index is constructed using weights of nominal GDP. The table is the analogous of Table 1 in Foerster, Sarte, and Watson (2011) for non-IP data.

Table D.13: Average pairwise correlations of sectoral non-IP indices.

1977-2011	1977-1983	1984-2007	2008-2011
0.18	0.32	0.10	0.21

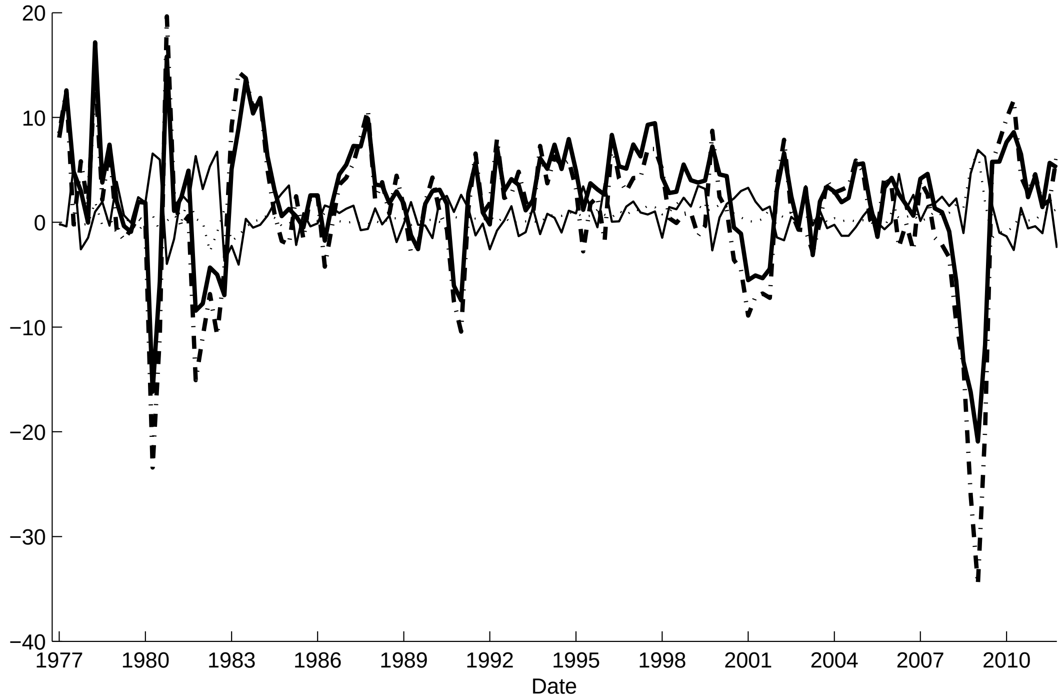
Entries are the average pairwise sample correlations of the annual growth rates of the 42 non-IP sectors in the GDP index considered in our paper. The table is the analogous of Table 2 in Foerster, Sarte, and Watson (2011) for non-IP data.

Table D.14: Standard deviation of aggregate indices of non-IP sectors constructed with and without sectoral covariance

	1977-2011	1977-1983	1984-2007	2008-2011
	A. Using Actual w_{it} Share Weights			
With sectoral covariation	1.7	2.1	1.3	1.9
Without sectoral covariation	0.9	0.9	0.8	0.9
	B. Using Equal $(1/N)$ Share Weights			
With sectoral covariation	2.1	2.6	1.3	3.4
Without sectoral covariation	1.2	0.9	1.1	1.4

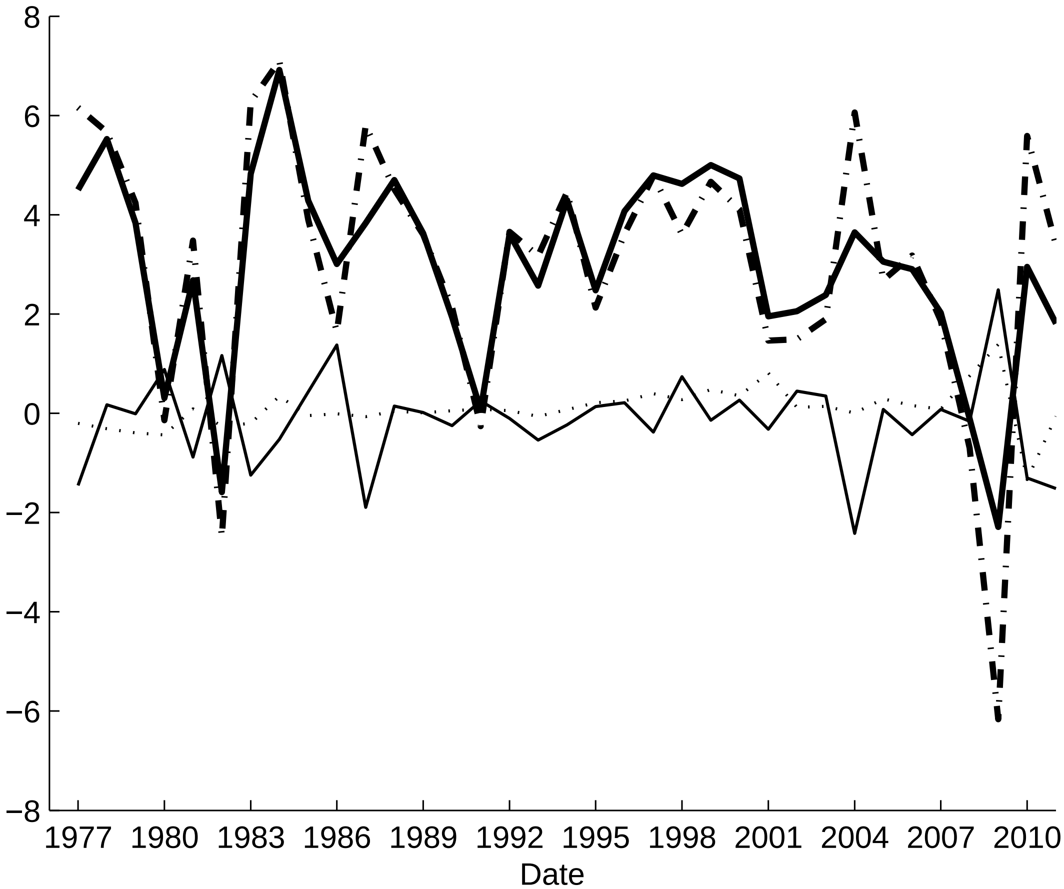
The entries for rows labeled "with sectoral covariation" are sample standard deviations of $\sum w_{it}x_{it}$ (Panel A) and $N^{-1}\sum x_{it}$ (Panel B). The entries labeled "without sectoral covariation" are computed as: $\sqrt{T^{-1}\sum_t\sum_i h_{it}^2(x_{it} - \bar{x}_i)^2}$, where $h_{it} = w_{it}$ in panel A and $h_{it} = N^{-1}$ in panel B. The table is the analogous of Table 3 in Foerster, Sarte, and Watson (2011) forn non-IP data.

Figure D.2: Share weight decomposition of quarterly IP index.



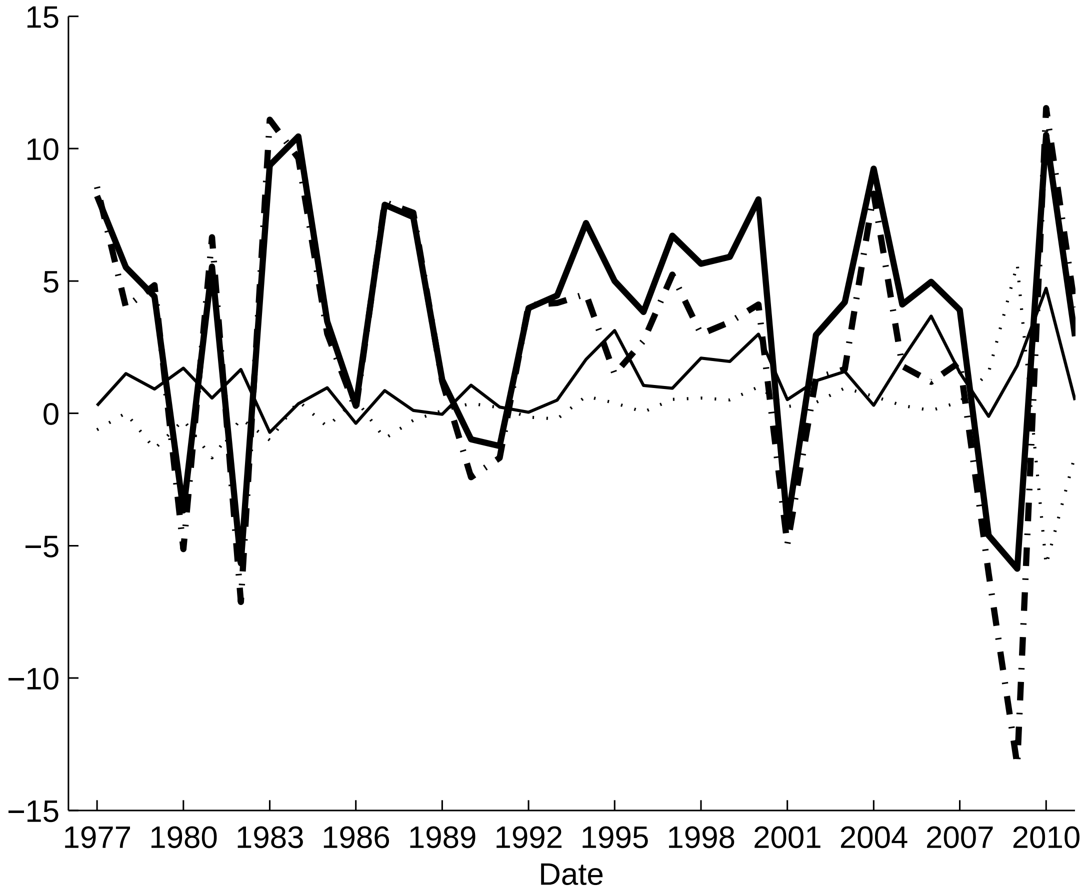
The figure displays the share weight decomposition of quarterly IP index growth rates. Percentage points are at annual rates. This figure corresponds to Figure 3 in Foerster, Sarte, and Watson (2011). The bold solid line corresponds to $\sum w_{it}x_{it}$ (i.e. the aggregate IP index). The bold dash-dotted line corresponds to $\sum(1/N)x_{it}$. The thin solid line corresponds to $\sum(\bar{w}_i - (1/N))x_{it}$. The thin dotted line corresponds to $\sum(w_{it} - \bar{w}_i)x_{it}$.

Figure D.3: Share weight decomposition of annual GDP index.



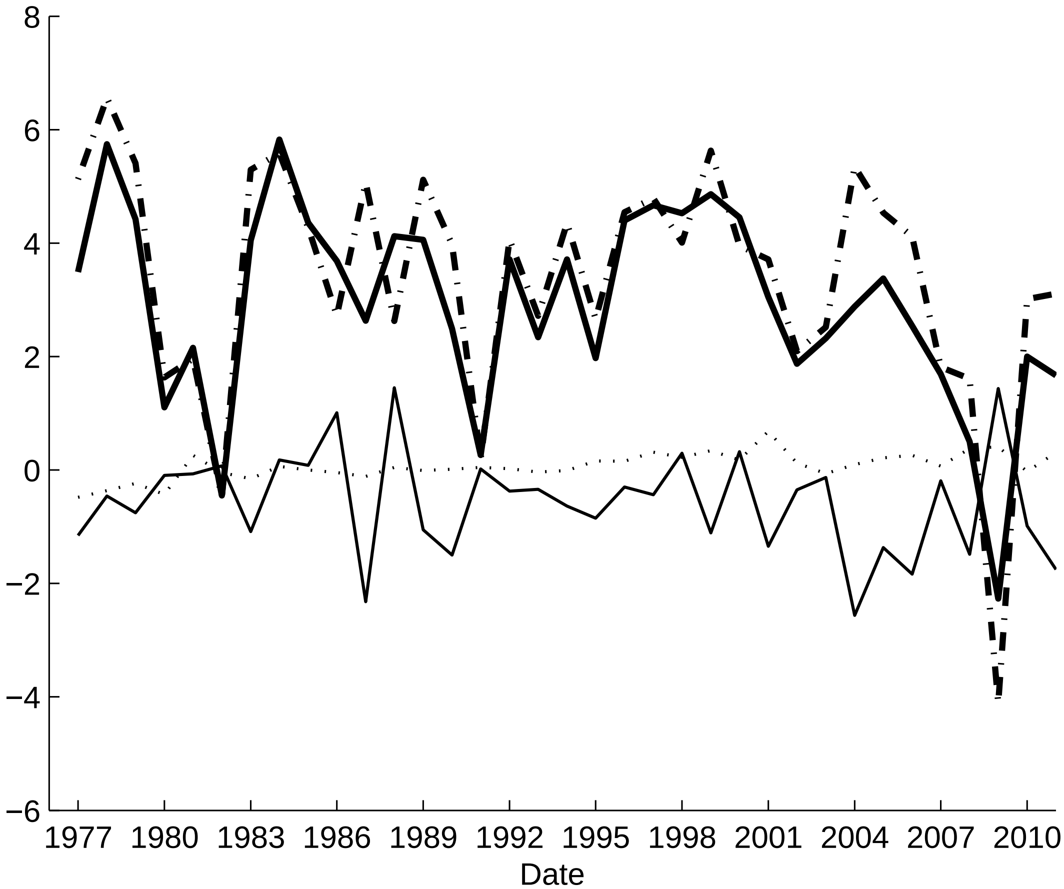
The figure displays the share weight decomposition of annual GDP index. The index is constructed using weights (w_{it}) of nominal GDP. This figure corresponds to Figure 3 in Foerster, Sarte, and Watson (2011). The bold solid line corresponds to $\sum w_{it}x_{it}$ (i.e. the aggregate real GDP index). The bold dash-dotted line corresponds to $\sum(1/N)x_{it}$. The thin solid line corresponds to $\sum(\bar{w}_i - (1/N))x_{it}$. The thin dotted line corresponds to $\sum(w_{it} - \bar{w}_i)x_{it}$.

Figure D.4: Share weight decomposition of annual Manufacturing sectors in GDP indices.



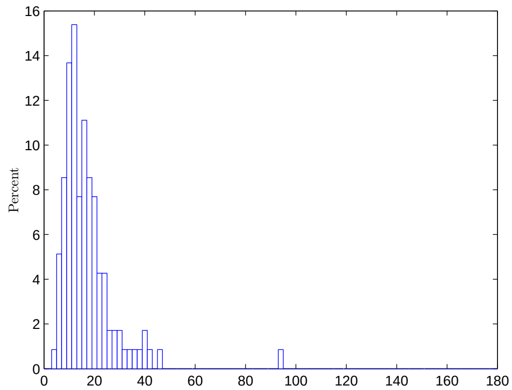
The figure displays the share weight decomposition of annual GDP index of Manufacturing sectors. The index is constructed using weights (w_{it}) of nominal GDP. This figure corresponds to Figure 3 in Foerster, Sarte, and Watson (2011). The bold solid line corresponds to $\sum w_{it}x_{it}$ (i.e. the aggregate real GDP index for IP sectors). The bold dash-dotted line corresponds to $\sum(1/N)x_{it}$. The thin solid line corresponds to $\sum(\bar{w}_i - (1/N))x_{it}$. The thin dotted line corresponds to $\sum(w_{it} - \bar{w}_i)x_{it}$.

Figure D.5: Share weight decomposition of annual non-IP sectors in GDP indices.

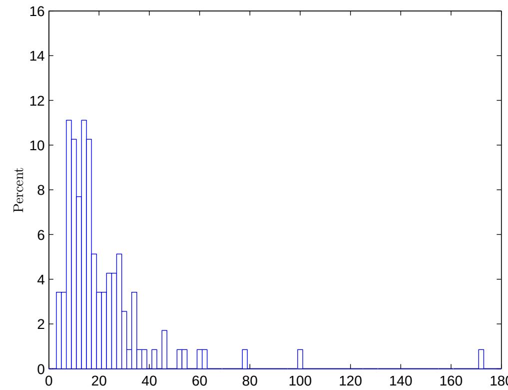


The figure displays the share weight decomposition of annual GDP index of non-IP sectors . The index is constructed using weights (w_{it}) of nominal GDP growth rates. This figure corresponds to Figure 3 in Foerster, Sarte, and Watson (2011). The bold solid line corresponds to $\sum w_{it}x_{it}$ (i.e. the aggregate real GDP index for non-IP sectors). The bold dash-dotted line corresponds to $\sum(1/N)x_{it}$. The thin solid line corresponds to $\sum(\bar{w}_i - (1/N))x_{it}$. The thin dotted line corresponds to $\sum(w_{it} - \bar{w}_i)x_{it}$.

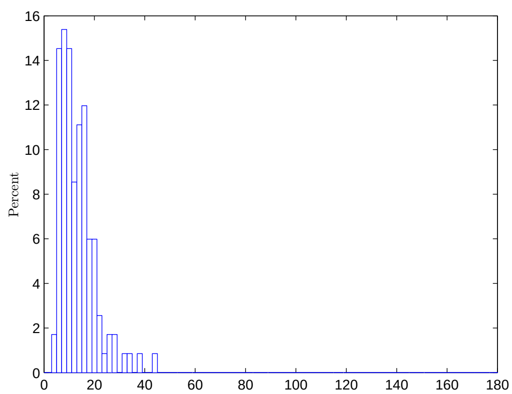
Figure D.6: Standard deviations of quarterly growth rates of sectoral Industrial Production indices.



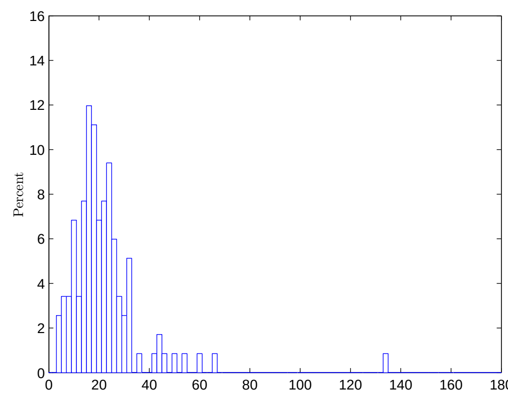
(a) 1977.Q1-2011.Q4



(b) 1977.Q1-1983.Q4



(c) 1984.Q1-2007.Q4

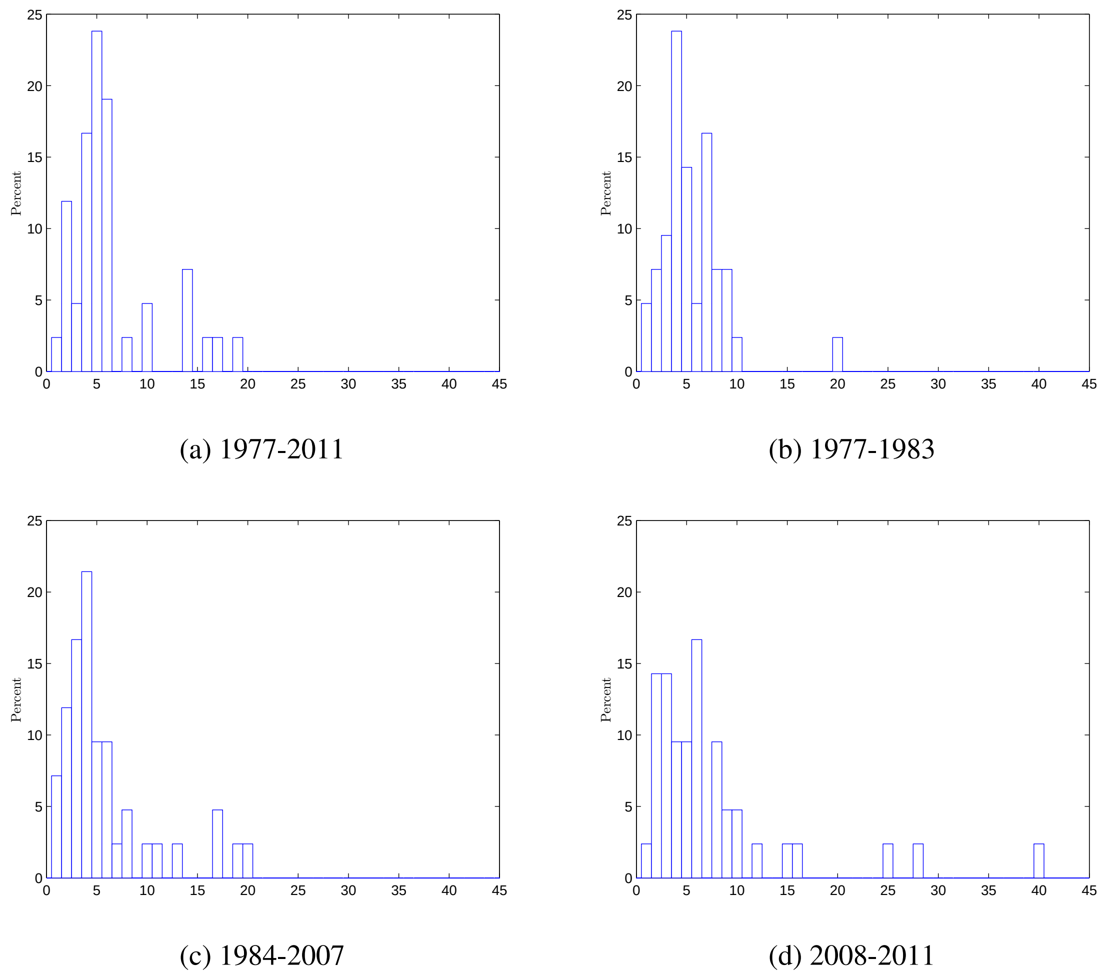


(d) 2008.Q1-2011.Q4

Each panel displays the histogram of the standard deviations of quarterly growth rates of sectoral IP indices. Percentage points are at annual rates. The graphs correspond to Figure 2 in Foerster, Sarte, and Watson (2011).

Panel	25th Percentile	Median	75th Percentile
(a) 1977.Q1-2011.Q4	10.77	14.19	19.71
(b) 1977.Q1-1983.Q4	10.60	15.92	25.29
(c) 1984.Q1-2007.Q4	8.32	11.48	16.71
(d) 2008.Q1-2011.Q4	14.40	18.91	25.65

Figure D.7: Standard deviations of annual growth rates of non-IP sectoral GDP indices.



(a) 1977-2011

(b) 1977-1983

(c) 1984-2007

(d) 2008-2011

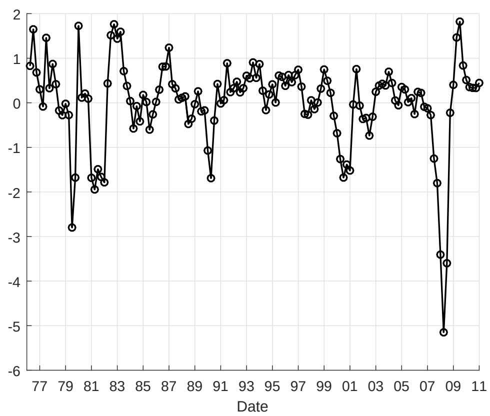
Each panel displays the histogram of the standard deviations of the annual growth rates of the 42 non-IP sectoral GDP indices. The graphs correspond to Figure 2 in Foerster, Sarte, and Watson (2011).

Panel	25th Percentile	Median	75th Percentile
(a) 1977-2011	3.92	4.91	6.39
(b) 1977-1983	3.65	4.89	7.05
(c) 1984-2007	3.22	4.31	6.19
(d) 2008-2011	3.22	5.47	8.39

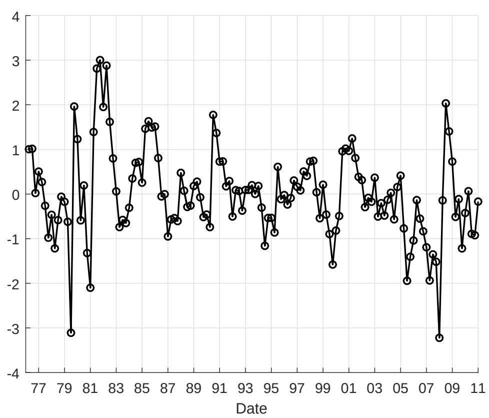
D.11.2 Tables and figures with PCA prior to aggregation

In this subsection we report the empirical results analogous to Section 7.2 obtained with an alternative estimation and inference method, in which PCs are extracted from the HF panel prior to aggregation (see Section D.9). The results are displayed in Figure D.8, Tables D.15, D.16 and D.17. By comparing with Figure 2, Tables I, II and III we see that the empirical results are substantially unchanged, whether aggregation is performed prior or post PCA on the HF data.

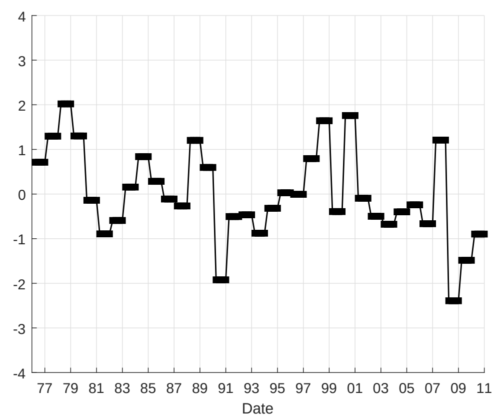
Figure D.8: Sample paths of the estimated factors, 1977 - 2011 (PCA first)



(a) Common factor



(b) High Frequency-specific factor



(c) Low Frequency-specific factor

Panel (a) displays the path of the estimated common factor. Panel (b) displays that of the HF-specific factor and Panel (c) that of the LF-specific factor. The factors are estimated from the panels of 42 annual non-IP GDP sectoral series and 117 quarterly IP indices using a mixed frequency group factor model with $k^C = k^H = k^L = 1$.

Table D.15: Adjusted R^2 and percentage values of BIC of the regressions with common and/or frequency-specific factors from economic activity indices growth rates (PCA first)

	$\bar{R}^2$: Quantiles					
Factors	10%	25%	50%	75%	90%	% BIC
Observables: Gross Domestic Product, 1977-2011						
common	-2.1	-0.5	11.6	29.0	43.0	38.1
common, LF-specific	0.3	8.9	25.5	34.7	60.3	28.6
LF-specific	-2.8	-2.2	5.5	15.5	22.9	33.3
Observables: IP, 1977.Q1-2011.Q4						
common	0.4	4.9	19.8	35.7	59.5	43.6
common, HF-specific	1.1	6.8	28.7	45.3	63.4	47.9
HF-specific	-0.6	-0.1	2.9	11.8	24.4	8.5

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 three 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. For both the IP and the non-IP panels, the sectoral growth rates are regressed on either the common factor only, or both the common and the frequency-specific factors, or the frequency-specific factor only.

Table D.16: Regression of yearly sectoral GDP growth on the common and LF-specific factors: adjusted R^2. (PCA first)

Panel A. Regressor: common factor

Sector	$\bar{R}^2$
Ten sectors with largest $\bar{R}^2$	
Truck transportation	62.78
Accommodation	62.23
Construction	44.00
Other transp. & support activ.	43.22
Administr. and support services	42.96
Other services, except government	42.52
Warehousing and storage	40.83
Air transportation	31.50
Retail trade	30.24
Amusem., gambling, & recr. ind.	29.31
Ten sectors with smallest $\bar{R}^2$	
Funds, trusts, & other finan. vehicles	-1.19
Motion picture & sound record. ind.	-1.74
Information and data processing services	-1.78
Pipeline transportation	-1.85
Transit and ground passenger transportation	-1.95
General government (States and Local)	-2.10
Forestry, fishing, and related activities	-2.27
Water transportation	-2.89
Securities, commodity contracts, and investm.	-3.01
Insurance carriers and related activities	-3.03

Panel B. Regressors: common and LF spec. factors

Sector	$\bar{R}^2$
Ten sectors with largest $\bar{R}^2$	
Misc. prof., scient., & tech. serv.	66.52
Administrative and support services	62.80
Truck transportation	62.15
Accommodation	61.26
Construction	59.96
Warehousing and storage	52.35
Government enterprises (States and Local)	45.69
Other services, except government	41.69
Other transportation and support activities	41.63
Government enterprises (FEDERAL)	37.64
Ten sectors with smallest $\bar{R}^2$	
Management of companies and enterprises	7.70
Ambulatory health care services	7.63
Funds, trusts, and other fin. vehicles	6.21
Information and data processing services	2.20
Educational services	1.22
Insurance carriers and related activities	0.47
Water transportation	-0.22
Farms	-2.07
Forestry, fishing, and related activities	-5.21
Securities, commodity contracts, and investm.	-5.95

Panel C. Increment in adjusted R^2

Sector	$\Delta\bar{R}^2$
Ten sectors with largest change in $\bar{R}^2$	
Misc. prof., scient., & tech. serv.	49.28
Government enterprises (States and Local)	34.40
Rental & leasing serv. & lessors of int. assets	29.48
General government (States and Local)	25.10
Legal services	24.63
Motion picture and sound recording industries	23.25
Fed. Reserve banks, credit interm., & rel. activ.	20.58
Social assistance	19.85
Administrative and support services	19.85
Real estate	17.94
Ten sectors with smallest change in $\bar{R}^2$	
Accommodation	-0.98
Rail transportation	-1.18
Other transportation and support activities	-1.59
Air transportation	-1.73
Retail trade	-2.17
Amusements, gambling, and recreation industries	-2.18
Educational services	-2.62
Farms	-2.86
Forestry, fishing, and related activities	-2.94
Securities, commodity contracts, & investm.	-2.94

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 cases 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*.

Table D.17: Adj. R^2 of aggregate IP and selected GDP indices growth rates on estimated factors (PCA first)

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)	BIC
Industrial Production	88.72	5.70	90.24	1.52	CH

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	60.65	8.81	74.38	13.72	CL
GDP - Manufacturing	82.24	-3.03	81.88	-0.36	C
GDP - Agriculture, forestry, fishing, and hunting	1.29	-2.59	-1.46	-2.75	C
GDP - Construction	44.00	11.62	59.96	15.96	CL
GDP - Wholesale trade	20.24	8.25	30.98	10.74	CL
GDP - Retail trade	30.24	-2.92	28.07	-2.17	C
GDP - Transportation and warehousing	62.41	-2.95	61.24	-1.17	C
GDP - Information	12.07	22.43	37.52	25.44	CL
GDP - Finance, insurance, real estate, rental, and leasing	-1.40	21.47	21.35	22.75	L
GDP - Professional and business services	30.12	30.48	65.77	35.65	CL
GDP - Educational services, health care, and social assistance	-1.32	17.88	17.69	19.01	L
GDP - Arts, entertainment, recreation, accommodation, and food serv.	53.12	-2.16	53.33	0.21	C
GDP - Government	-2.12	22.23	20.35	22.46	L

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 only (columns $\bar{R}^2(H)$ and $\bar{R}^2(L)$), 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.

D.11.3 Supplementary tables and figures to Section 7.2

In Table D.18 we report the estimated number of pervasive factors k_1 and k_2, selected in each of the panels of data considered in Section 7.2 according to the IC_{p1} and IC_{p2} information criteria of Bai and Ng (2002).

Table D.18: Estimated number of factors: results for IC_{p1} and IC_{p2} information criteria

	X_{HF}	X^H	X^L
IP data: 1977.Q1-2011.Q. Non-IP data: Gross Domestic Product, 1977-2011			
IC_{p1}	2	2	1
IC_{p2}	1	2	1

The number of latent pervasive factors selected by the IC_{p1} and 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 correspond to non-IP data. We use $k_{max} = 15$ as maximum number of factors when computing IC_p's criteria. In the first line the quarterly IP data are for sample period 1977.Q1-2011.Q4, and the annual non-IP data are GDP growth rates for the sample period 1977-2011.

In Table D.19 we report the four eigenvalues of the sample variance-covariance matrix of the stacked PC's estimated in each subpanel of IP data (X^H) and non-IP data (X^L). The two largest eigenvalues are equal to 1 plus the largest canonical correlations, as implied by Lemma D.2. We find an eigenvalue close to two, which is consistent with the presence of one common factor in each of the two different mixed frequency dataset considered. The asymptotic theory developed for the number of canonical correlations equal to one among the PC's extracted separately from the two panels, could be used to derive a test statistic for the number of common factors among the two panels, based on the number of eigenvalues equal to 2 of the sample variance-covariance matrix of the stacked PC's.

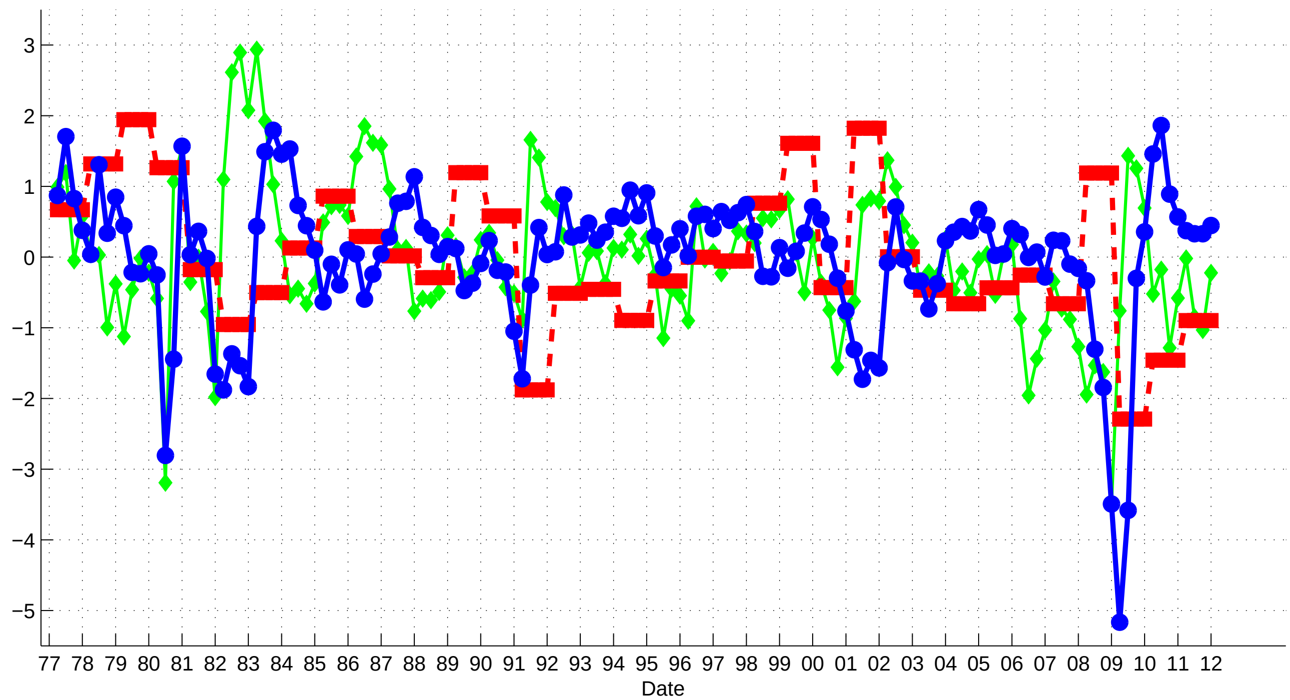
Table D.19: Eigenvalues of the variance-covariance matrix of the stacked PC's (Aggregation first)

1^{st} eig.	2^{nd} eig.	3^{rd} eig.	4^{th} eig.
IP data: 1977.Q1-2011.Q4. Non-IP data: Gross Domestic Product, 1977-2011			
1.84	1.06	0.94	0.16

In this table we report the eigenvalues of the sample variance-covariance matrix of the stacked PC's estimated in each subpanel of IP (X^H) and non-IP data (X^L). We extract the first 2 PC's in each subgroup, and compute the variance-covariance matrix of these 4 stacked PC's.

Figure D.9 provides an alternative representation of the estimates of the common, HF-specific and LF-specific factors displayed in Figure 2, Panels *(b)* - *(d)*, where the three factor paths are superposed. The values of the three factors are computed for the two mixed frequency panels of 42 GDP sectors and 117 IP indices.

Figure D.9: Trajectories of the estimated common, HF-specific and LF-specific factors (Aggregation first)



The Figure displays the time series of estimated values of the common factor (blue circles), the LF-specific factor (red squares) and the HF-specific factor (green diamonds). For each year we represent the LF factor as 4 squares corresponding to the 4 quarters, assuming the same value. The factors are estimated from the panel of real output growth rates of 42 GDP sectors and 117 Industrial Production indices, using a mixed frequency factor model with $k^C = k^H = k^L = 1$. The sample period for the estimation of the factor model is 1977.Q1-2011.Q4.

In Table D.20 we report the empirical correlation matrix of the estimated factors, computed at yearly frequency. In Table D.21 we display the sample correlations among the quarterly estimates of the common and HF-specific factors and the yearly estimates of the LF-specific factor.

Table D.20: Correlation matrix of the estimated factors, computed at annual frequency. (Aggregation first)

	$\hat{\bar{g}}_t^C$	$\hat{\bar{g}}_t^H$	$\hat{\bar{g}}_t^L$
$\hat{\bar{g}}_t^C$	1.00		
$\hat{\bar{g}}_t^H$	0.00	1.00	
$\hat{\bar{g}}_t^L$	0.00	-0.18	1.00

In the table we display the sample correlation matrix of the stacked vector of estimated factors $(\hat{\bar{g}}_t^C, \hat{\bar{g}}_t^H, \hat{\bar{g}}_t^L)$. 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 the factor model and the regressions is 1977.Q1-2011.Q4.

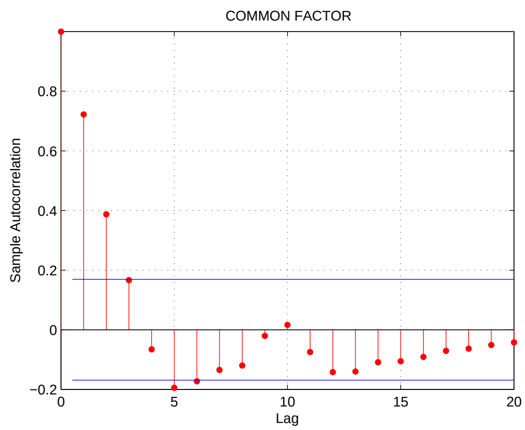
Table D.21: Correlation matrix of the estimated common, HF-specific and LF-specific factors: quarterly observations (Aggregation first)

	$\hat{g}_{1,t}^C$	$\hat{g}_{2,t}^C$	$\hat{g}_{3,t}^C$	$\hat{g}_{4,t}^C$	$\hat{g}_{1,t}^H$	$\hat{g}_{2,t}^H$	$\hat{g}_{3,t}^H$	$\hat{g}_{4,t}^H$	$\hat{\bar{g}}_t^L$
$\hat{g}_{1,t}^C$	1.00	0.75	0.41	0.20	0.12	-0.38	-0.53	-0.29	0.28
$\hat{g}_{2,t}^C$	0.75	1.00	0.77	0.30	0.35	0.11	-0.41	-0.26	-0.05
$\hat{g}_{3,t}^C$	0.41	0.77	1.00	0.66	0.40	0.23	-0.19	-0.03	-0.35
$\hat{g}_{4,t}^C$	0.20	0.30	0.66	1.00	0.34	-0.03	0.11	0.43	-0.16
$\hat{g}_{1,t}^H$	0.12	0.35	0.40	0.34	1.00	0.56	0.44	0.41	-0.01
$\hat{g}_{2,t}^H$	-0.38	0.11	0.23	-0.03	0.56	1.00	0.65	0.47	-0.27
$\hat{g}_{3,t}^H$	-0.53	-0.41	-0.19	0.11	0.44	0.65	1.00	0.79	-0.12
$\hat{g}_{4,t}^H$	-0.29	-0.26	-0.03	0.43	0.41	0.47	0.79	1.00	-0.06
$\hat{\bar{g}}_t^L$	0.28	-0.05	-0.35	-0.16	-0.01	-0.27	-0.12	-0.06	1.00

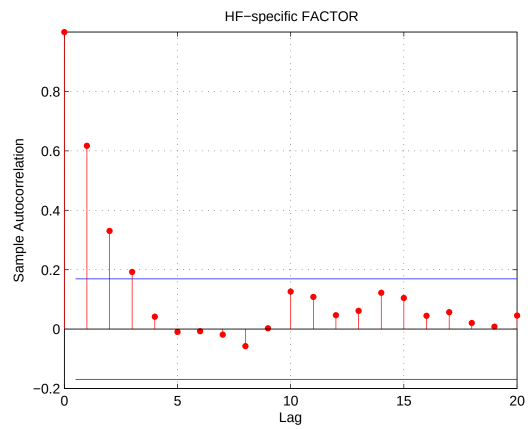
In the table we display the correlation matrix of the stacked vector of estimated factors $(\hat{g}_{1,t}^C, \hat{g}_{2,t}^C, \hat{g}_{3,t}^C, \hat{g}_{4,t}^C, \hat{g}_{1,t}^H, \hat{g}_{2,t}^H, \hat{g}_{3,t}^H, \hat{g}_{4,t}^H, \hat{\bar{g}}_t^L)$. The factors are estimated from the panel of 42 GDP sectors 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 the factor model and the regressions is 1977-2011.

Figure D.10 displays the sample autocorrelations functions for the estimated common, HF-specific and LF-specific factors.

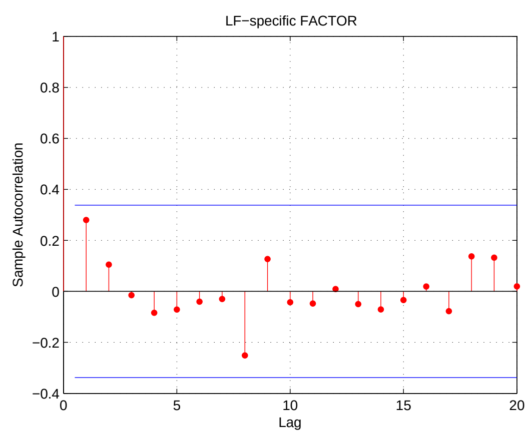
Figure D.10: Autocorrelation functions of the estimated common and specific factors. (Aggregation first)



(a) Common factor: autocorrelation function.



(b) HF factor: autocorrelation function.

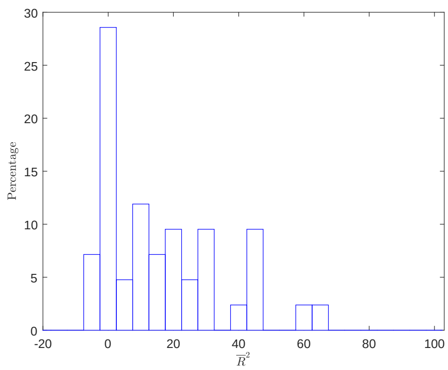


(c) LF factor: autocorrelation function.

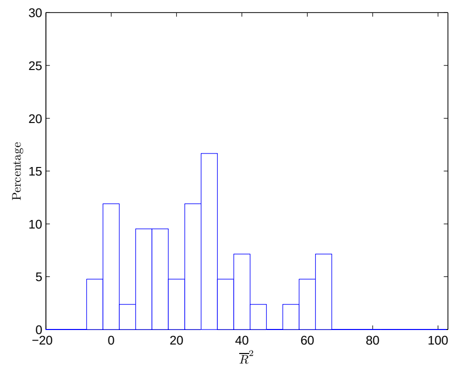
Panel (a) displays the sample autocorrelation function of the estimated values of the common factor at high frequency. Panel (b) displays the sample autocorrelation function of the estimated values of the HF factor at high frequency. Panel (c) displays the sample autocorrelation function of the estimated values of the LF factor at low frequency. The horizontal lines are asymptotic 95% confidence bands. The factor values are estimated from the panel of 42 GDP sectors and 117 Industrial Production indices using a mixed frequency factor model with $k^C = k^H = k^L = 1$. The sample period for the estimation of the factor model is 1977.Q1-2011.Q4.

The histograms in Figure D.11, Panels *(a)* and *(b)*, represent the empirical distribution of the $\bar{R}^2$ corresponding to the first and second lines of Table I, respectively. Moreover, the histograms in Panels *(c)* and *(d)*, represent the empirical distribution of the $\bar{R}^2$ corresponding the fourth and fifth lines of Table I.

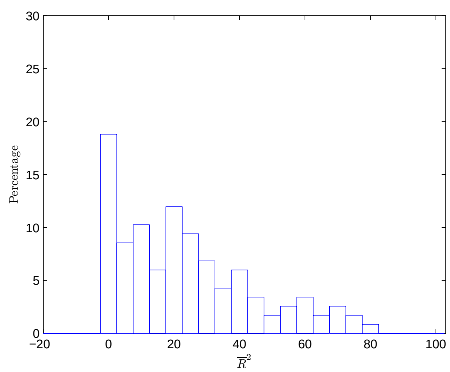
Figure D.11: Adj. R^2 of the regression of yearly sectoral GDP growth rates on estimated factors (Aggregation first)



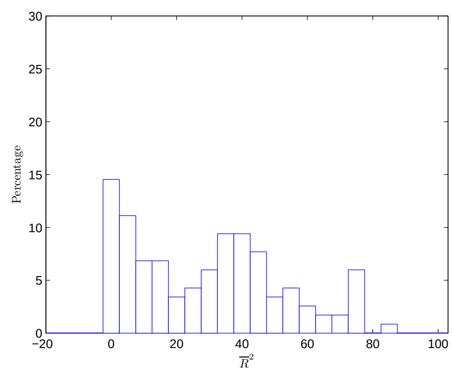
(a) GDP sectors vs. common factor



(b) GDP sectors vs. common & LF factors



(c) IP sectors vs. common factor



(d) IP sectors vs. common & HF factors

In Panel (a) we show the histogram of the adjusted R^2, denoted $\bar{R}^2$, of the regressions of the yearly growth rates of sectoral GDP indices on the estimated common factor. In Panel (b) we show the histogram of the adjusted R^2 of the regressions of the same growth rates on the estimated common and LF-specific factors. In Panel (c) we show the histogram of the adjusted R^2, of the regressions of the quarterly growth rates of the IP indices on the estimated common factor. In Panel (d) we show the histogram of the adjusted R^2 of the regressions of the same growth rates on the estimated common and HF-specific 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 the factor model and the regressions is 1977-2011.

In the following Table D.22 we list the top and bottom ten IP sectors in terms of $\bar{R}^2$ when regressed on the common factor only, and both the common and HF-specific factors. We also report the top and bottom ten IP sectors with the highest and lowest absolute increments in $\bar{R}^2$ when the HF-specific factor is added to the common one. The factors are estimated from the panel of 42 GDP sectors and 117 Industrial Production indices using a mixed frequency factor model with $k^C = k^H = k^L = 1$. PCA is performed on the flow-sampled HF data (PCA post aggregation). The sample period for the estimation of both the factor model and the regressions is 1977-2011. Table D.23 provides the same results when PCA is perfomed on the quarterly HF data (PCA prior to aggregation).

Tables D.24, D.25, D.26 display the entire lists of 42 non-IP sectors ranked by the three criteria considered in Table II, Panels A, B, and C, respectively. Specifically, Table D.24 displays the full list of non-IP sectors ranked according to the value of their $\bar{R}^2$ when regressed on the common factor only, Table D.25 displays the full list of non-IP sectors ranked according to the value of their $\bar{R}^2$ when regressed on the common and LF-specific factors, Table D.26 displays the full list of non-IPsectors ranked according to the value of the increment in $\bar{R}^2$ when the LF-specific factor is added to the common one.

Table D.22: Regression of quarterly sectoral IP growth on the common and HF-specific factors: adjusted R^2 (Aggregation first)

Panel a. Regressor: common factor

Sector	$\bar{R}^2$
Ten sectors with largest $\bar{R}^2$	
Forging and stamping	79.86
Other fabricated metal product"	74.67
Coating, engraving, heat treat., & allied activ.	74.28
Com. & serv. ind. machin. & other gen. purp. machin.	72.01
Machine shops, turned product, & screw, nut, & bolt	71.53
Foundries	69.78
Other electrical equipment	66.79
Metalworking machinery	66.00
Plastics product	61.36
Household and instit. furniture and kitchen cabinet	60.74
Ten sectors with smallest $\bar{R}^2$	
Other Food Except Coffee & Tea	-0.28
Pharmaceutical & medicine	-0.46
Grain & oilseed milling	-0.46
Aerospace product & parts	-0.50
Ice Cream & Frozen Desserts	-0.56
Oil & gas extraction	-0.56
Wineries & Distilleries	-0.56
Nonferrous metal (except aluminum) smelt. & refin.	-0.58
Dairy product (except frozen)	-0.66
Fruit & vegetable preserving & specialty food	-0.71

Panel b. Regressors: common and HF spec. factors

Sector	$\bar{R}^2$
Ten sectors with largest $\bar{R}^2$	
Forging & stamping	82.87
Com. & serv. ind. machin. & other gen. purp. machin.	76.77
Plastics product	76.72
Metalworking machinery	75.77
Other fabricated metal product	75.69
Coating, engraving, heat treating, & allied activities	74.68
Household & institutional furniture & kitchen cabinet	73.36
Machine shops, turned product, & screw, nut, & bolt	73.14
Millwork	71.19
Foundries	69.61
Ten sectors with smallest $\bar{R}^2$	
Other Food Except Coffee & Tea	0.79
Sugar & confectionery product	0.75
Ice Cream & Frozen Desserts	0.19
Wineries & Distilleries	-0.49
Grain & oilseed milling	-0.69
Animal slaughtering & processing	-0.81
Aerospace product & parts	-1.16
Nonferrous metal (except aluminum) smelt. & ref.	-1.31
Dairy product (except frozen)	-1.39
Fruit & vegetable preserving & specialty food	-1.44

Panel c. Absolute increment in adj. R^2

Sector	$\Delta\bar{R}^2$
Ten sectors with largest change in $\bar{R}^2$	
Mining & oil & gas field machinery	44.03
Veneer, plywood, & engineered wood product	37.63
Millwork	30.46
All other wood product	26.31
Sawmills & wood preservation	24.87
Major appliance	24.78
Resin & synthetic rubber	23.86
Support activities for mining	23.85
Paperboard container	23.46
Fiber, yarn, & thread mills Tobacco	22.94
Ten sectors with smallest change in $\bar{R}^2$	
Audio & video equipment	-0.63
Soap, cleaning compound, & toilet preparation	-0.63
Other transportation equipment	-0.65
Aerospace product & parts	-0.66
Ship & boat building	-0.69
Petroleum refineries	-0.70
Animal food	-0.71
Fruit & vegetable preserving & specialty food	-0.73
Nonferrous metal (except aluminum) smelt. & ref.	-0.73
Dairy product (except frozen)	-0.73

In the table we display the adjusted R^2, denoted $\bar{R}^2$, for the time series regressions of the growth rates of the of 117 Industrial Production indices on the estimated factors. The factors are estimated from the panel of 42 GDP sectors and 117 Industrial Production indices using a mixed frequency factor model with $k^C = k^H = k^L = 1$. The regressions in *Panel a* involve a HF explained variable and the estimated common factor. The regressions in *Panel b* involve a HF explained variable and both the common and HF-specific estimated factors. In *Panel c* we display the absolute increment $\bar{R}^2$, when the HF-specific factor is added to the common factor. The sample period for the estimation of both the factor model and the regressions is 1977-2011.

Table D.23: Regression of quarterly sectoral IP growth on the common and HF-specific factors extracted from original IP series: adjusted R^2 (PCA first)

Panel a. Regressor: common factor		**Panel b. Regressors: common and HF spec. factors**		**Panel c. Absolute increment in adj. 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 $\bar{R}^2$*	
Forging and stamping	80.28	Forging and stamping	82.86	Mining and oil and gas field machinery	43.10
Other fabricated metal product”	74.90	Com. & serv. ind. machin. & other gen. purp. machin.	76.95	Veneer, plywood, & engineered wood product	38.65
Coating, engraving, heat treating, & allied activities	74.50	Plastics product	76.71	Millwork	31.62
Com. & serv. ind. machin. & other gen. purp. machin.	72.63	Metalworking machinery	75.76	All other wood product	27.34
Machine shops, turned product, and screw, nut, & bolt	71.82	Other fabricated metal product”	75.66	Sawmills and wood preservation	25.70
Foundries	69.75	Coating, engraving, heat treating, & allied activities	74.75	Major appliance	25.67
Other electrical equipment	66.99	Household and institutional furniture & kitchen cabinet	73.34	Resin and synthetic rubber	24.73
Metalworking machinery	66.71	Machine shops, turned product, & screw, nut, and bolt	73.11	Paperboard container	24.14
Other Miscellaneous Manufacturing	60.62	Millwork	71.27	Fiber, yarn, & thread mills Tobacco	23.53
Plastics product	60.53	Foundries	69.54	Support activities for mining	23.23
Ten sectors with smallest $\bar{R}^2$		*Ten sectors with smallest $\bar{R}^2$*		*Ten sectors with smallest change in $\bar{R}^2$*	
Other Food Except Coffee and Tea	-0.27	Sugar and confectionery product	0.77	Audio and video equipment	-0.62
Grain and oilseed milling	-0.46	Other Food Except Coffee and Tea	0.68	Other transportation equipment	-0.62
Pharmaceutical and medicine	-0.48	Seafood product preparation and packaging	0.21	Soap, cleaning compound, and toilet preparat.	-0.64
Aerospace product and parts	-0.51	Wineries and Distilleries	-0.52	Aerospace product and parts	-0.67
Oil and gas extraction	-0.54	Grain and oilseed milling	-0.71	Ship and boat building	-0.68
Wineries and Distilleries	-0.56	Animal slaughtering and processing	-0.81	Petroleum refineries	-0.69
Seafood product preparation and packaging	-0.57	Aerospace product and parts	-1.18	Animal food	-0.72
Nonferrous metal (except aluminum) smelting & refining	-0.58	Nonferrous metal (except alum.) smelt. & refin.	-1.31	Fruit and vegetable preserving and specialty food	-0.72
Dairy product (except frozen)	-0.66	Dairy product (except frozen)	-1.39	Nonferrous metal (except aluminum) smelting & refining	-0.73
Fruit and vegetable preserving and specialty food	-0.71	Fruit and vegetable preserving and specialty food	-1.44	Dairy product (except frozen)	-0.73

In the table we display the adjusted R^2, denoted $\bar{R}^2$, for the time series regressions of the growth rates of the of 117 Industrial Production indices on the estimated factors. The factors are estimated from the panel of 42 GDP sectors and 117 Industrial Production indices using a mixed frequency factor model with $k^C = k^H = k^L = 1$. Principal components are extracted on the original quarterly IP dataset and then aggregated yearly. The regressions in *Panel a* involve a HF explained variable and the estimated common factor. The regressions in *Panel b* involve a HF explained variable and both the common and HF-specific estimated factors. In *Panel c* we display the absolute increment $\bar{R}^2$, when the HF-specific factor is added to the common factor. The sample period for the estimation of both the factor model and the regressions is 1977-2011.

Table D.24: Adjusted R^2 of the regression of yearly sectoral GDP growth on the common factor. (Aggregation first)

Sector	$\bar{R}^2$
Truck transportation	63.10
Accommodation	62.43
Construction	44.05
Other transportation and support activities	43.31
Administrative and support services	42.69
Other services, except government	42.53
Warehousing and storage	40.95
Air transportation	31.58
Retail trade	30.70
Amusements, gambling, and recreation industries	29.17
Government enterprises (federal)	28.91
Rail transportation	24.84
Performing arts, spectator sports, museums, and related activities	22.63
Publishing industries (includes software)	22.02
Computer systems design and related services	21.24
Food services and drinking places	20.59
Wholesale trade	20.35
Miscellaneous professional, scientific, and technical services	16.98
Waste management and remediation services	14.79
Social assistance	12.91
General government (federal)	11.97
Government enterprises (state & local)	11.10
Real estate	10.39
Legal services	10.19
Federal Reserve banks, credit intermediation, and related activities	9.74
Educational services	3.97
Rental and leasing services and lessors of intangible assets	2.81
Broadcasting and telecommunications	1.24
Ambulatory health care services	1.01
Farms	0.93
Hospitals and nursing and residential care facilities	0.64
Management of companies and enterprises	-0.45
Funds, trusts, and other financial vehicles	-1.23
Motion picture and sound recording industries	-1.68
Pipeline transportation	-1.74
Information and data processing services	-1.84
Transit and ground passenger transportation	-2.05
General government (state & local)	-2.12
Forestry, fishing, and related activities	-2.33
Water transportation	-2.94
Securities, commodity contracts, and investments	-2.99
Insurance carriers and related activities	-3.03

In the table we display the adjusted R^2, denoted $\bar{R}^2$, for the time series regressions of each of the 42 GDP sectors on the estimated common factor. The factors are estimated from the panel of 42 GDP sectors and 117 Industrial Production 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.Q1-2011.Q4. The regressions in this table are restricted MIDAS regressions.

Table D.25: Adjusted R^2 of the regression of yearly sectoral GDP growth on the common and LF-specific factors. (Aggregation first)

Sector	$\bar{R}^2$
Miscellaneous professional, scientific, and technical services	66.67
Administrative and support services	62.63
Truck transportation	62.51
Accommodation	61.48
Construction	59.75
Warehousing and storage	52.53
”Government enterprises (STATES AND LOCAL)”	45.78
Other services, except government	41.75
Other transportation and support activities	41.71
”Government enterprises (FEDERAL)”	37.78
Legal services	34.51
Social assistance	32.82
Rental and leasing services and lessors of intangible assets	32.32
Wholesale trade	30.83
Performing arts, spectator sports, museums, and related activities	30.49
Federal Reserve banks, credit intermediation, and related activities	30.05
Air transportation	29.81
Retail trade	28.56
Real estate	28.53
Computer systems design and related services	27.07
Amusements, gambling, and recreation industries	27.02
Publishing industries (includes software)	23.85
Rail transportation	23.68
”General government (STATES AND LOCAL)”	22.78
Food services and drinking places	21.67
Motion picture and sound recording industries	21.10
Hospitals and nursing and residential care facilities	17.47
Broadcasting and telecommunications	14.46
Waste management and remediation services	14.24
Pipeline transportation	14.13
”General government (FEDERAL)”	11.11
Transit and ground passenger transportation	9.18
Ambulatory health care services	7.76
Management of companies and enterprises	7.52
Funds, trusts, and other financial vehicles	6.15
Information and data processing services	1.96
Educational services	1.35
Insurance carriers and related activities	0.36
Water transportation	-0.64
Farms	-1.87
Forestry, fishing, and related activities	-5.31
Securities, commodity contracts, and investments	-5.99

In the table we display the adjusted R^2, denoted $\bar{R}^2$, for the time series regressions of each of the 42 GDP sectors on the estimated common and LF-specific factors. The factors are estimated from the panel of 42 GDP sectors and 117 Industrial Production 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.Q1-2011.Q4. The regressions in this table are restricted MIDAS regressions.

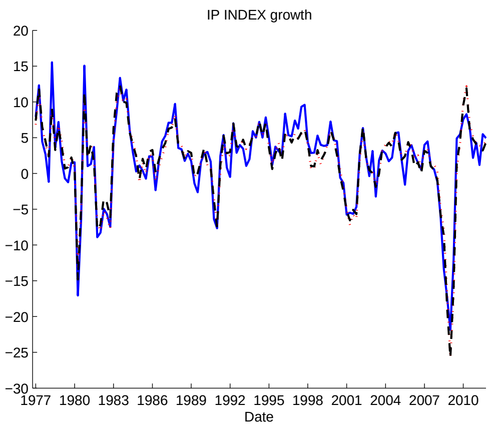
Table D.26: Change in adjusted R^2 of the regression of yearly sectoral GDP growth on the common and LF-specific factors vs. the regression on the common factor only. (Aggregation first)

Sector	change in $\bar{R}^2$
Miscellaneous professional, scientific, and technical services	49.69
Government enterprises (STATES AND LOCAL)	34.69
Rental and leasing services and lessors of intangible assets	29.52
General government (STATES AND LOCAL)	24.90
Legal services	24.32
Motion picture and sound recording industries	22.77
Federal Reserve banks, credit intermediation, and related activities	20.31
Administrative and support services	19.95
Social assistance	19.91
Real estate	18.14
Hospitals and nursing and residential care facilities	16.84
Pipeline transportation	15.87
Construction	15.70
Broadcasting and telecommunications	13.23
Warehousing and storage	11.58
Transit and ground passenger transportation	11.23
Wholesale trade	10.48
Government enterprises (federal)	8.87
Management of companies and enterprises	7.98
Performing arts, spectator sports, museums, and related activities	7.87
Funds, trusts, and other financial vehicles	7.39
Ambulatory health care services	6.76
Computer systems design and related services	5.83
Information and data processing services	3.80
Insurance carriers and related activities	3.39
Water transportation	2.30
Publishing industries (includes software)	1.83
Food services and drinking places	1.07
Waste management and remediation services	-0.54
Truck transportation	-0.60
Other services, except government	-0.78
General government (federal)	-0.86
Accommodation	-0.96
Rail transportation	-1.16
Other transportation and support activities	-1.59
Air transportation	-1.77
Retail trade	-2.15
Amusements, gambling, and recreation industries	-2.15
Educational services	-2.62
Farms	-2.80
Forestry, fishing, and related activities	-2.98
Securities, commodity contracts, and investments	-3.00

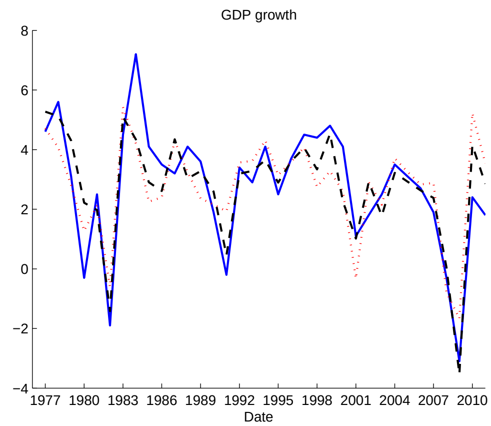
In the table we display the difference in the adjusted R^2 ($\bar{R}^2$) from the regressions of each industrial production index growth on the common and LF-specific estimated factors, and on the HF factor only. The factors are estimated from the panel of 42 GDP sectors and 117 Industrial Production 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.Q1-2011.Q4. The regressions in this table are restricted MIDAS regressions.

Figure D.12 displays the trajectories of the fitted values of some of the regressions in Table III, i.e. the regressions of aggregated indexes on the estimated common factors only, and on the common and frequency-specific factors together. As already remarked for Table III, adding the frequency-specific factor in the regressions improves the fit for some non-IP service series such as the Professional and Business Services Index (panel (d) in Figure D.12).

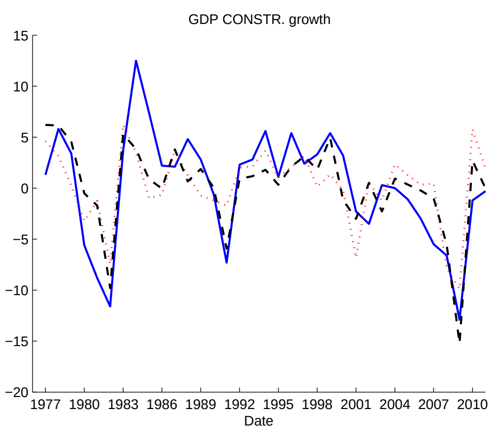
Figure D.12: Regression of LF and HF indices on estimated factors. (Aggregation first)



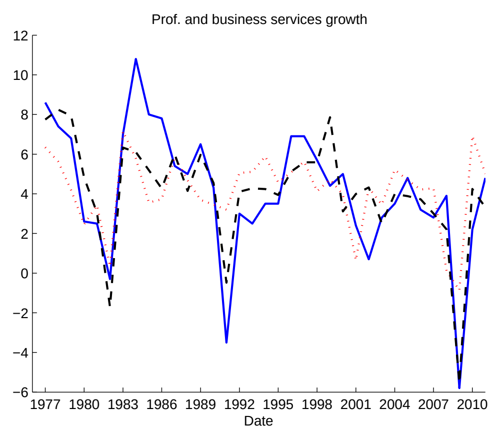
(a) HF Index: Industrial Production Index growth.



(b) LF Index: Aggregate GDP Index growth.



(c) LF Index: GDP-Construction Index growth.



(d) LF Index: Prof. and Busn. Serv. Index growth.

Each panel displays the time series of the growth rate of an observed index (solid blue line) and its fitted value obtained from a regression on the common factor (dotted red line). Fitted values from a regression on multiple factors (dashed black line) are also displayed. In the first panel we regress the IP index on both the common and HF-specific factors, in the second panel we regress the aggregate GDP Index (LF) on $\hat{\hat{g}}_t^C$, $\hat{\hat{g}}_t^H$ and $\hat{\hat{g}}_t^L$. In the third and fourth panels we regress the growth rates of the LF Construction Index and of Professional and Business Services Index, respectively, on both the common and LF-specific factors. The indices considered in the first, second and fourth panels are aggregates of the indices used to estimate the factors. The factors are estimated from the panel of 42 GDP sectors and 117 Industrial Production 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.Q1-2011.Q4.

In Table D.27 we report the results of regressions of aggregated indices on the estimated factors. In particular, we regress the output of each aggregate index either on the estimated common factor only, the LF-specific factor only, the HF-specific factor only, and all the three estimated factors together. We report the adjusted R^2s of these four types of regressions, and the increment in the adjusted R^2 when the two frequency-specific factors are added as regressors to the common factor. This table completes the information in Table III of the paper.

Table D.27: Adj. R^2 of selected GDP and IP indices growth rates on the estimated factors

Yearly observations, 1977-2011

Sector	(1) $\bar{R}^2(C)$	(2) $\bar{R}^2(L)$	(3) $\bar{R}^2(H)$	(4) $\bar{R}^2(C+L+H)$	(4) - (1)
LF observations					
GDP	60.54	8.59	-2.04	73.39	12.85
GDP - Manufacturing	81.88	-3.03	0.22	82.90	1.02
GDP - Agriculture, forestry, fishing, and hunting	1.43	-2.52	-0.85	-2.16	-3.59
GDP - Construction	44.05	11.22	-2.84	58.70	14.64
GDP - Wholesale trade	20.35	7.90	-2.94	29.77	9.41
GDP - Retail trade	30.70	-2.86	2.67	33.71	3.00
GDP - Transportation and warehousing	62.14	-2.95	0.19	61.95	-0.19
GDP - Information	12.14	22.28	-3.03	36.53	24.39
GDP - Finance, insurance, real estate, rental, and leasing	-1.42	21.22	-2.19	18.58	20.00
GDP - Professional and business services	30.02	30.21	-1.98	64.52	34.50
GDP - Educational services, health care, and social assistance	-1.38	18.38	-0.60	16.25	17.63
GDP - Arts, entert., recreat., accommodation, and food serv.	53.51	-2.23	-0.50	57.00	3.49
GDP - Government	-2.12	22.37	-2.95	18.96	21.08

In the table we display the adjusted R^2, denoted $\bar{R}^2$, of the regression of growth rates of selected HF and LF indices on the common factor (column $\bar{R}^2(C)$), the specific HF and LF factors (columns $\bar{R}^2(L)$ and $\bar{R}^2(H)$) and on these three factors together (column (4)). The last column displays the difference between the values in the fourth and the first columns, i.e. the increment in the adjusted R^2 when both specific factors are added as regressors to the common factor.

D.11.4 Supplementary tables for 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. In this section we consider different 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 D.28. In Table D.29 we report the canonical correlation analysis common factor tests. 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 D.28 and D.29 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 D.9 (performing PCA on the flow-sampled HF data).[10]

[10]For both subsamples 1984.Q1-2007.Q4 and 1984.Q1-2011.Q4, when we selecting a model with $k_1 = k_2 = 1$ pervasive factor in each subpanel, we do not reject the null hypotheses of 1 common factor. On the other hand, when we select $k_1 = k_2 = 2$ we reject the null of both 1 and 2 common factors for the subsample 1984.Q1-2011.Q4, and we cannot reject

Table D.28: Estimated number of factors for different subsamples

	X_{HF}	X^H	X^L
IP data: 1984.Q1-2007.Q4. Non-IP data: Gross Domestic Product, 1984-2007			
IC_{p2}	1	2	1
IP data: 1984.Q1-2011.Q4. Non-IP data: Gross Domestic Product, 1984-2011			
IC_{p2}	1	2	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. Panel X^L corresponds to non-IP data. We use $k_{max} = 15$ as maximum number of factors when computing IC_{p2}.

Table D.29: Canonical correlations and test statistics for common factors (Aggregation first)

$\hat{\rho}_1$	$\hat{\rho}_2$	$\tilde{\xi}(2)$	$\tilde{\xi}(1)$
IP data: 1984.Q1-2007.Q4. Non-IP data: Gross Domestic Product, 1984-2007			
0.58	-	-	-1.48
0.70	0.33	-1.50	-2.55
IP data: 1984.Q1-2011.Q4. Non-IP data: Gross Domestic Product, 1984-2011			
0.76	-	-	-0.92
0.81	0.13	-4.01	-2.81

For each subsample, the first line reports the canonical correlation of the first PCs computed in each subpanel of IP and non-IP data (i.e. when we select $k_1 = k_2 = 1$), and the values of the test statistic $\tilde{\xi}(r)$, for the null hypothesis of $r = 1$ common factors. The second line reports the canonical correlation of the first two PCs computed in each subpanel of IP and non-IP data (i.e. when we select $k_1 = k_2 = 2$), and the values of the test statistic $\tilde{\xi}(r)$, for the null hypotheses of $r = 1, 2$ common factors.

In Table D.30 we report the results of regressions of aggregated indices on the estimated factors for the two subsamples. 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 either on the estimated common factor only, or on the frequency-specific factor, or both, and concentrate our attention on the adjusted R^2s of these regressions. The results in Table D.30 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 Table III of the paper. This also transpires from Panels B and C, when examining the total GDP variations projected on the

the null of 2 common factors for the subsample 1984.Q1-2007.Q4. We impute these instabilities to the small time-series dimensions of the subsamples.

common factor. During the Great Moderation the common factor only explains around 30 %, which goes to 56 % when we add the Great Depression. The other findings, i.e. the exposure of the various subindices, appear to be similar to those in the full sample.

Table D.30: Adj. R^2 of aggregate IP and selected GDP indices growth rates on estimated factors (Aggregation first)

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

	(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 regressions 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 only (columns $\bar{R}^2(H)$ and $\bar{R}^2(L)$), 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.

E Monte Carlo experiments

Appendix E describes the Monte Carlo simulation study used to assess the small sample properties of the test statistics proposed in Theorems 1 and 2, and those of the sequential testing procedure for the selection of k^C introduced in Proposition 2. Our selection procedure is also compared with the one based on penalized information criteria for group-factor models proposed in Chen (2012), and the three-steps procedure of Wang (2012). This appendix is composed of seven parts.

Section E.1 defines the model design used for the simulations. Section E.2 gives the values of simulation design parameters and the sample sizes. Section E.3 describes the content of the tables summarizing the results of MC simulations, and how they have been obtained. Section E.4 includes a discussion of all simulation results. Section E.5 displays the tables of results relative to size and power properties of the test for the number of common factors k^C based on the test statistics of Theorems 1 and 2. Section E.6 displays the tables of results comparing the performance of the sequential testing procedure for the selection of k^C introduced in Proposition 2 with alternative procedures adapted from earlier literature. These results are summarized also in Section 6 of the main body of the paper. Finally, in Section E.7 we display and discuss the quantiles of the cross-sectional distribution of R^2 and adjusted R^2 for regressions of simulated observables on factors when the number of common factors is either correctly specified, or overestimated, for a DGP in which specific factors at high and low frequency are highly correlated.

E.1 Simulation design model

We consider simulation designs characterized by different numbers of common (k^C) and group-specific (k^H, k^L) factors in different data generating processes (DGPs). We assume that $k^H = k^L$. The number of Monte Carlo (MC) simulations for each design is 4000. In each MC experiment, mixed frequency panels of observations are generated from the high frequency DGP defined in equation (5.1):

$$x^H_{m,t} = \Lambda_{HC} g^C_{m,t} + \Lambda_H g^H_{m,t} + e^H_{m,t},$$

$$x^{L*}_{m,t} = \Lambda_{LC} g^C_{m,t} + \Lambda_L g^L_{m,t} + e^L_{m,t},$$

where $m = 1, ..., M$ and $t = 1, ..., T$. Loading matrices are defined as $\Lambda_{HC} = [\lambda_{HC,1}, ..., \lambda_{HC,N_H}]'$, $\Lambda_H = [\lambda_{H,1}, ..., \lambda_{H,N_H}]'$, $\Lambda_{LC} = [\lambda_{LC,1}, ..., \lambda_{LC,N_L}]'$, $\Lambda_L = [\lambda_{L,1}, ..., \lambda_{L,N_L}]'$. Vectors $\lambda_{HC,i}$ and $\lambda_{H,i}$, have dimensions k^C and k^H, respectively, for $i = 1, ..., N_H$. Vectors $\lambda_{LC,j}$ and $\lambda_{L,j}$, have dimensions k^C and k^L, respectively, for $j = 1, ..., N_L$. We consider the case of flow-sampled low frequency observable variables:

$$x^L_t = \sum_{m=1}^M x^{L*}_{m,t}.$$

Therefore $x^H_{m,t}$ and x^L_t constitute the panels of mixed-frequency observables. Subsections E.1.1, E.1.2, and E.1.3 describe the DGPs for the latent factors, idiosyncratic innovations, and loadings, respectively. These designs extend the ones in Bates, Plagborg-Moller, Stock, and Watson (2013). Table E.1 in Section E.2 displays the different values of k^C, k^H, k^L, while Table E.2 displays the values of N_H, N_L, and T.

E.1.1 Common and group-specific factors

The vectors of latent factors $g^C_{m,t}$, $g^H_{m,t}$, and $g^L_{m,t}$ have dimensions k^C, k^H, and k^L, respectively. We define the new $(k^C + k^H + k^L)$-dimensional vector $g_{m,t} = [g^{C\prime}_{m,t}, g^{H\prime}_{m,t}, g^{L\prime}_{m,t}]'$, and assume the following high-frequency autoregressive dynamics as DGP:

$$g_{m,t} = a_F g_{m-1,t} + \sqrt{\varsigma}\, \eta_{m,t}, \qquad \text{(E.1)}$$

where the scalar a_F is an AR(1) coefficient common to all factors and $\varsigma = (1-a_F^2)/(M^2\kappa)$, with $\kappa = 1 - \frac{2}{M^2}\sum_{m=1}^{M-1} m(1-a_F^{M-m})$. The innovations $\eta_{m,t} = \left[\eta_{m,t}^{C\prime}, \eta_{m,t}^{H\prime}, \eta_{m,t}^{L\prime}\right]'$ are such that:

$$\eta_{m,t} \sim i.i.N(0, \Sigma_\eta), \qquad \Sigma_\eta = \begin{bmatrix} I_{k^C} & 0 & 0 \\ 0 & I_{k^H} & \Phi \\ 0 & \Phi' & I_{k^L} \end{bmatrix}, \tag{E.2}$$

where $\Phi = \phi I_{k^H}$. The scalar parameter $\phi \in (-1,\ 1)$ generates correlation between the first HF-specific factor and the first LF-specific factor, the second HF-specific factor and the second LF-specific, and so on. The term $\sqrt{\varsigma}$ in equation (E.1) implies that the flow-sampled factor vector $\bar{g}_t := \sum_{m=1}^{M} g_{m,t}$ satisfies the normalization in equation (2.2). This can be shown by noting that we have:

$$M^2\kappa = M^2\left(1 - \frac{2}{M^2}\sum_{m=1}^{M-1} m(1-a_F^{M-m})\right) = M + 2\sum_{m=1}^{M-1}(M-m)a_F^m \ ,$$

and from (E.1)-(E.2):

$$V(g_{m,t}) \ = \ \frac{1}{M^2\kappa}\begin{bmatrix} I_{k^C} & 0 & 0 \\ 0 & I_{k^H} & \Phi \\ 0 & \Phi' & I_{k^L} \end{bmatrix}.$$

Then, we get:

$$\begin{aligned} V(\bar{g}_t) \ &= \ \sum_{m=1}^{M} V(g_{m,t}) + \sum_{m=1}^{M}\sum_{n=1,n\neq m}^{M} Cov(g_{m,t}, g_{n,t}) = MV(g_{m,t}) + \sum_{m=1}^{M}\sum_{n=1,n\neq m}^{M} a_F^{|m-n|}V(g_{m,t}) \\ &= \ \left(M + 2\sum_{m=1}^{M-1}(M-m)a_F^m\right)V(g_{m,t}) \ = \ \begin{bmatrix} I_{k^C} & 0 & 0 \\ 0 & I_{k^H} & \Phi \\ 0 & \Phi' & I_{k^L} \end{bmatrix}, \end{aligned}$$

which yields equation (2.2) written for the flow-sampled factor values.
The initial values of the factors are drawn from their stationary distributions. Table E.1 in Section E.2 displays the different values of parameters a_F and ϕ used in each design.

E.1.2 Idiosyncratic innovations

The idiosyncratic innovations vectors $e_{m,t}^H$ and $e_{m,t}^L$ have dimensions N_H, and N_L, respectively. We define the new $(N_H + N_L)$-dimensional vector $e_{m,t} = [e_{m,t}^{H\prime}, e_{m,t}^{L\prime}]'$, and assume the following DGP for the innovations:

$$e_{m,t} \ = \ a_e e_{m-1,t} + v_{m,t}, \tag{E.3}$$

where the scalar a_e is a common AR(1) coefficient for the innovations. The innovations $v_{m,t}$ are such that:

$$v_{m,t} \ \sim \ i.i.N(0, (1-a_e^2)\Sigma_e), \qquad \Sigma_e = \{\beta^{|i-j|}\}_{ij}, \qquad i,j = 1,...,N_H + N_L \ .$$

The scalar β induces cross-sectional dependence among the idiosyncratic innovations, as in Bates, Plagborg-Moller, Stock, and Watson (2013). The initial values of the idiosyncratic innovations are drawn from their stationary distributions. We consider different values of parameter β as reported in Table E.2, and keep $a_e = 0$ for all simulation designs.

E.1.3 Factor loadings

The simulation designs for the factor loadings adapt the designs of Bates, Plagborg-Moller, Stock, and Watson (2013) to our set-up with common and frequency-specific factors. The rows of the loading matrices Λ_{HC}, Λ_{LC}, Λ_H, and Λ_L are simulated from Gaussian distributions with parameters chosen to match the R^2s for the regressions of $x_{m,t}^{Hi}$ and $x_{m,t}^{Li}$ on the factors observed in the empirical analysis. More specifically, for each $i = 1, ..., N_H$, let $R_{all}^{2,Hi}$ denote a measure which is akin to the population R^2 of the regression of $x_{m,t}^{Hi}$ on all k^C common and k^H specific factors and is defined below. Let also $R_C^{2,Hi}$ denote the same measure for the regression of $x_{m,t}^{Hi}$ on the k^C common factors only. For each i we draw a value of $R_{all}^{2,Hi}$ uniformly from the interval $[0.1,\ R_{max}^2]$, where R_{max}^2 is a parameter in $(0.1, 1)$. We also draw a value of $R_C^{2,Hi}$ uniformly from the interval $[0.2 \cdot R_{all}^{2,Hi},\ 0.8 \cdot R_{all}^{2,Hi}]$. Analogously, for each $j = 1, ..., N_L$, we denote with $R_{all}^{2,Lj}$ a measure which is akin to the population R^2 of the regression of $x_{m,t}^{Lj}$ on all k^C common and k^L specific factors as defined below. Let also $R_C^{2,Lj}$ denote the same measure for the population R^2 of the regression of $x_{m,t}^{Lj}$ on the k^C common factors only. For each j, we draw a value of $R_{all}^{2,Lj}$ uniformly from the interval $[0.1,\ R_{max}^2]$, and a value of $R_C^{2,Lj}$ uniformly from the interval $[0.2 \cdot R_{all}^{2,Lj},\ 0.8 \cdot R_{all}^{2,Lj}]$. In every MC simulation, for each $i = 1, ..., N_H$, and $j = 1, ..., N_L$, the loadings are drawn from the following independent Gaussian distributions:

$$
\begin{aligned}
\lambda_{HC,i} &\sim i.i.N\left(0,\ \lambda_{1,i}^{*\,2} \cdot I_{k^C}\right), \\
\lambda_{H,i} &\sim i.i.N\left(0,\ \lambda_{2,i}^{*\,2} \cdot I_{k^H}\right), \\
\lambda_{LC,j} &\sim i.i.N\left(0,\ \lambda_{3,j}^{*\,2} \cdot I_{k^C}\right), \\
\lambda_{L,j} &\sim i.i.N\left(0,\ \lambda_{4,j}^{*\,2} \cdot I_{k^L}\right),
\end{aligned}
\tag{E.4}
$$

where the scalars $\lambda_{1,i}^* \equiv \lambda_{1,i}^*(R_{all}^{2,Hi}, R_C^{2,Hi})$, and $\lambda_{2,i}^* \equiv \lambda_{2,i}^*(R_{all}^{2,Hi}, R_C^{2,Hi})$ are chosen such that:

$$
\begin{aligned}
\frac{E\left[(\lambda_{HC,i}' g_{m,t}^C + \lambda_{H,i}' g_{m,t}^H)^2 \mid \lambda_{1,i}^*, \lambda_{2,i}^*\right]}{E\left[(x_{m,t}^{Hi})^2 \mid \lambda_{1,i}^*, \lambda_{2,i}^*\right]} &= R_{all}^{2,Hi} \\
\frac{E\left[(\lambda_{HC,i}' g_{m,t}^C)^2 \mid \lambda_{1,i}^*\right]}{E\left[(x_{m,t}^{Hi})^2 \mid \lambda_{1,i}^*, \lambda_{2,i}^*\right]} &= R_C^{2,Hi},
\end{aligned}
$$

for $i = 1, ..., N_H$, and the scalars $\lambda_{3,j}^* \equiv \lambda_{3,j}^*(R_{all}^{2,Lj}, R_C^{2,Lj})$, and $\lambda_{4,j}^* \equiv \lambda_{4,j}^*(R_{all}^{2,Lj}, R_C^{2,Lj})$ are chosen such that:

$$
\begin{aligned}
\frac{E\left[\left(\lambda_{LC,j}' \sum_{m=1}^M g_{m,t}^C + \lambda_{L,j}' \sum_{m=1}^M g_{m,t}^L\right)^2 \mid \lambda_{3,j}^*, \lambda_{4,j}^*\right]}{E\left[(x_t^{Lj})^2 \mid \lambda_{3,j}^*, \lambda_{4,j}^*\right]} &= R_{all}^{2,Lj}, \\
\frac{E\left[\left(\lambda_{LC,j}' \sum_{m=1}^M g_{m,t}^C\right)^2 \mid \lambda_{3,j}^*\right]}{E\left[(x_t^{Lj})^2 \mid \lambda_{3,j}^*\right]} &= R_C^{2,Lj},
\end{aligned}
$$

for $j = 1, ..., N_L$. Hence, $R_{all}^{2,Hi}$ is the ratio of factor-explained variance to total variance of the HF data, accounting for randomness in factors, loadings and errors, and similarly for the other R^2 measures. From (E.1)-

(E.4) we have:

$$\begin{cases} \lambda_{1,i}^{*\,2} &= \dfrac{1}{k^C} \cdot \dfrac{R_C^{2,Hi}}{1-R_{all}^{2,Hi}} \cdot \left(M + 2\displaystyle\sum_{m=1}^{M-1} (M-m) a_F^m \right) \\ \lambda_{3,j}^{*\,2} &= \dfrac{1}{k^C} \dfrac{R_C^{2,Lj}}{1-R_{all}^{2,Lj}} \cdot \left(M + 2\displaystyle\sum_{m=1}^{M-1} (M-m) a_e^m \right), \end{cases} \qquad \text{if } k^C \neq 0$$

$$\lambda_{2,i}^{*\,2} = \frac{1}{k^H} \cdot \frac{R_{all}^{2,Hi} - R_C^{2,Hi}}{1-R_{all}^{2,Hi}} \cdot \left(M + 2\sum_{m=1}^{M-1} (M-m) a_F^m \right), \qquad \text{if } k^H \neq 0$$

$$\lambda_{4,j}^{*\,2} = \frac{1}{k^L} \cdot \frac{R_{all}^{2,Lj} - R_C^{2,Lj}}{1-R_{all}^{2,Lj}} \cdot \left(M + 2\sum_{m=1}^{M-1} (M-m) a_e^m \right), \qquad \text{if } k^L \neq 0$$

and:

$$\begin{aligned} \lambda_{1,i}^{*\,2} &= \lambda_{3,j}^{*\,2} = 0, \text{ if } k^C = 0, \\ \lambda_{2,i}^{*\,2} &= 0, \text{ if } k^H = 0, \\ \lambda_{4,j}^{*\,2} &= 0, \text{ if } k^L = 0, \end{aligned}$$

for $i = 1, ..., N_H$, $j = 1, ..., N_L$. The draws of innovations in factors and errors, the loadings and the R^2 measures are all mutually independent.

E.2 Simulation design parameters and sample sizes

Table E.1 displays the values of the parameters k^C, k^H, k^L, β, a_F, and ϕ associated to each simulation design:

Table E.1: Parameters k^C, k^H, k^L, β, a_F, and ϕ, for each simulation design

Design # / Parameter	k^C	$k^H = k^L$	β	a_F	ϕ
Design 1	1	1	0	0	0
Design 2	**2**	**0**	0	0	0
Design 3	**2**	**1**	0	0	0
Design 4	1	1	**0.2**	0	0
Design 5	1	1	0	**0.6**	0
Design 6	1	1	0	0	**0.7**
Design 7	1	1	0	0	**0.95**
Design 8	1	**5**	0	0	0
Design 9	1	**5**	0	0	**0.5**
Design 10	1	**5**	0	0	**0.7**
Design 11	1	**5**	0	0	**0.95**

In this appendix we report results fixing $a_e = 0$, $R^2_{max} = 0.8$, and the number of HF sub-periods $M = 4$ for all designs. Results for $R^2 = 0.6$ and 0.95, $M = 1, 12, 21$ are available upon request, and produce results qualitatively similar to the one presented here. For each design we consider the following sample sizes:

Table E.2: Sample sizes considered in each simulation design

N_H	N_L	T
50	50	35
100	50	35
100	100	35
100	100	50
200	100	50
200	200	50
200	200	100
500	500	100
500	500	200
500	500	300
800	800	500
1000	1000	600

Design 5, corresponding to the case $k^C = k^H = k^L = 1$, with $[N_H, N_L, T] = [\ 100,\ 50,\ 35\ ]$, $\phi = 0.0$, $a_F = 0.6$, and $M = 4$, is the most similar to the setting of the empirical application in terms of sample sizes and parameter values. Tables E.3 and E.4 show that the values of the parameter $R^2_{max} = 0.8, 0.6$ produce cross-sectional distributions of adjusted-R^2s for the regressions of observables on the factors in line with those found in the empirical application of the paper (see Table I).

Table E.3: Sample averages over 2000 MC simulations of the quantiles of adjusted R^2 of regressions on true and estimated factors, with $R^2_{max} = 0.8$, $N_H = 100$, $N_L = 50$, $T = 35$, $M = 4$, $a_F = 0.6$, $\beta = 0$, $a_e = 0.0$, $\phi = 0.0$.

Panel A: $R^2_{max} = 0.8$**, true factors**

	$\bar{R}^2$: Quantile				
Factors	10%	25%	50%	75%	90%
Observables: LF variables					
common	-2.4	0.1	9.4	29.1	52.0
common, LF-spec.	0.9	11.4	31.8	56.7	74.5
LF-spec.	-2.4	0.1	9.4	28.6	51.3
Observables: HF variables					
common	-0.3	1.7	9.8	28.3	50.8
common, HF-spec.	4.0	13.1	32.0	56.2	74.1
HF-spec.	-0.3	1.7	9.7	28.0	50.6

Panel B: $R^2_{max} = 0.8$**, estimated factors**

	$\bar{R}^2$: Quantile				
Factors	10%	25%	50%	75%	90%
Observables: LF variables					
common	-2.4	0.0	9.3	28.8	51.5
common, LF-spec.	1.5	12.6	33.5	57.7	74.3
LF-spec.	-2.4	0.2	9.8	29.4	51.3
Observables: HF variables					
common	-0.3	1.9	10.4	29.2	51.7
common, HF-spec.	4.5	14.1	33.6	57.4	74.2
HF-spec.	-0.3	1.7	9.7	28.1	50.0

In each line we report the sample averages, computed over 2000 MC simulations, of the quantiles of adj. R^2 of regressions on true and estimated factors. In both Panels A and B, the regressions in the first three lines involve the growth rates of the 50 LF observables as dependent variables, while those in the last three lines involve the growth rates of the 100 HF observables as dependent variables. In Panel A the explanatory variables are the true simulated factors. In Panel B the explanatory variables are the estimated factors, assuming that the true numbers of factors in the DGP ($k^C = k^H = k^L = 1$) are known.

Table E.4: Sample averages over 2000 MC simulations of the quantiles of adj. R^2 of regressions on true and estimated factors, with $R^2_{max} = 0.6$, $N_H = 100$, $N_L = 50$, $T = 35$, $M = 4$, $a_F = 0.6$, $\beta = 0$, $a_e = 0.0$, $\phi = 0.0$.

Panel A: $R^2_{max} = 0.6$, true factors

Factors	$\bar{R}^2$: Quantile 10%	25%	50%	75%	90%
Observables: LF variables					
common	-2.6	-0.6	6.8	22.6	41.8
common, LF-spec.	-0.8	7.4	23.4	41.1	61.8
LF-spec.	-2.6	-0.6	6.7	22.3	41.5
Observables: HF variables					
common	-0.4	1.1	7.2	21.3	40.0
common, HF-spec.	2.7	9.5	23.5	43.2	60.4
HF-spec.	-0.4	1.1	7.2	21.3	39.9

Panel B: $R^2_{max} = 0.6$, estimated factors

Factors	$\bar{R}^2$: Quantile 10%	25%	50%	75%	90%
Observables: LF variables					
common	-2.6	-0.6	6.7	22.3	41.2
common, LF-spec.	-0.2	8.6	25.3	45.7	62.3
LF-spec.	-2.5	-0.4	7.4	23.5	42.4
Observables: HF variables					
common	-0.4	1.3	7.8	22.5	41.2
common, HF-spec.	3.1	10.5	25.2	44.8	61.2
HF-spec.	-0.4	1.2	7.5	21.7	40.0

In each line we report the sample averages, computed over 2000 MC simulations, of the quantiles of adj. R^2 of regressions on true and estimated factors. In both Panels A and B, the regressions in the first three lines involve the growth rates of the 50 LF observables as dependent variables, while those in the last three lines involve the growth rates of the 100 HF observables as dependent variables. In Panel A the explanatory variables are the true simulated factors. In Panel B the explanatory variables are the estimated factors, assuming that the true numbers of factors in the DGP ($k^C = k^H = k^L = 1$) are known.

E.3 Description of content of tables of results

Size and power properties, Section E.5 (p. 120 - 130)

The simulation designs described above allow to study the small sample size and power properties of the *feasible* test statistic $\tilde{\xi}(k^C)$ in equation (4.6) in Theorem 2, and the size of the *infeasible* test statistics $\hat{\xi}(k^C)$ of Theorem 1. The upper panel in each table corresponds to cases in which the *feasible* and *infeasible* statistics are computed from factors estimated by PCA from HF data directly, and then flow-sampled in order to compute the test statistics. The lower panel in each table corresponds to cases in which HF observables are first flow-sampled, and then factors are estimated by PCA on this flow-sampled panel of HF data before computing the test statistics. We refer to Section D.9 for details on the practical implementation. Data are simulated under the DGP with parameters listed in the title of the corresponding page. Each of the tables on p. 120 - 130 displays in the first three columns the values of N_H, N_L and T. Moreover:

- columns 4 - 6 display the empirical size of the *infeasible* test statistic $\hat{\xi}(k^C)$ defined in Theorem 1, and computed under the null hypothesis of a number of common factors k^C equal to the one specified in the title of the corresponding page;
- columns 7 - 9 display the empirical size of the *feasible* test statistic $\tilde{\xi}(k^C)$ defined in Theorem 2, and computed under the null hypothesis of a number of common factors k^C equal to the one specified in the title of the corresponding page;
- columns 10 - 12 display the empirical power of the *feasible* test statistic $\tilde{\xi}(k^C)$ defined in Theorem 2. The number of common factors in the DGP is k^C and is specified in the title of the corresponding page. The empirical power is computed as the empirical frequency of rejection of the test of the null hypothesis of $k^C + 1$ common factors, against the alternative of a number strictly smaller than $k^C + 1$. [11]

[11] For Design 2 the power has not been computed because $k_1 = k_2 = k^C = 2$ and thus the null hypothesis of $k^C + 1 = 3$

The feasible statistic is computed as in (4.6), i.e. assuming conditionally heteroschedastic and serially as well as cross-sectionally independent errors. Design 4 with $\beta = 0.2$ allows us to investigate the robustness of the statistic computed as in (4.6) to small levels of cross-sectional correlation. In all tests we consider nominal sizes of 1%, 5%, 10%. All empirical size and power are computed as the empirical rejection frequencies of the tests obtained over 4000 MC simulations from the same DGP. The null of each test is rejected when the value of the test statistic computed on simulated data is strictly smaller than the 1%, 5%, 10% quantiles of the asymptotic distribution of the test statistic, which is a standard Gaussian.
In each simulation we draw new random samples of the factors, the loadings and the idiosyncratic innovations. In unreported experiments we fix the same values for the factors and the loadings in all the 4000 MC simulations, and draw new random samples only for the idiosyncratic innovations. Also in this case, we obtain results analogous to the ones presented in this Online Appendix.

Selection of number of factors, Section E.6 (p. 131 - 141)

MC simulations are used to evaluate the accuracy of the estimators for the number of common factors k^C provided by *i)* our consistent sequential testing procedure defined in Proposition 2, *ii)* the selection procedure based on the penalized information criterion of Theorem 3.7 in Chen (2012), and *iii)* the three-steps selection procedure proposed by Wang (2012).
The estimators are evaluated by comparing the average estimated number of common (k^C), high-frequency-specific (k^H), and low-frequency-specific (k^L) factors, computed across the 4000 simulations for each DGP described in the title of the corresponding page. For all the competing estimators we consider both the case in which the true numbers of pervasive factors $k_1 = k^C + k^H$ and $k_2 = k^C + k^L$ in the two panels are known, and only k^C needs to be estimated (*lower* panel in each table), and also the case in which k_1 and k_2 are estimated (*upper* panel in each table). Each of the tables on p. 131 - 141 displays in the first three columns the values of N_H, N_L and T. Moreover:

- columns 4-6 display the average number of estimated factors for our sequential testing procedure of Proposition 2. The *feasible* statistics in the sequential testing procedure are computed from factors estimated by PCA from HF data directly, and then flow-sampled in order to compute the test statistics. These columns are labelled "AGGR (2016), HF data: PCA first";

- columns 7-9 display the average number of estimated factors for our sequential testing procedure of Proposition 2 for the number of common factors. The *feasible* statistics in the sequential testing procedure are computed from factors estimated by PCA on flow-sampled HF data. These columns are labelled "AGGR (2016), HF data: flow samp. first";

- columns 10-12 display the average number of estimated factors by the selection procedure based on the information criterion of Theorem 3.7 in Chen (2012). These columns are labelled "CHEN (2012)";

- columns 13-15 display the average number of estimated factors by the following three-steps procedure to determine k^C: (1) estimate the number of pervasive factors in each of the two panels separately, and denote them as $\hat{k}_1$ and $\hat{k}_2$, (2) estimate the number R of pervasive factors in the stacked panel of HF (flow sampled) and LF data[12], and denote it as $\hat{R}$, (3) determine the number of common factors k^C as $\hat{k}_1 + \hat{k}_2 - \hat{R}$.[13] This procedure is a special case of the one suggested by Wang (2012). These last three columns are labeled "WANG (2012)".

common factors cannot be considered in that setting.

[12]Note that $R = k^C + k^H + k^L$.

[13]Note that $k_1 + k_2 - R = (k^C + k^H) + (k^C + k^L) - (k^C + k^H + k^L) = k^C$.

The critical values $z_{\alpha_{N,T}}$ for our sequential testing procedure - AGGR (2016) - are determined by assigning the values $\gamma = 0.1$, and $c = 0.95$ to the functional form $z_{\alpha_{N,T}} = -c(N\sqrt{T})^{\gamma}$ given in equation (4.7). This choice of the functional form satisfies the conditions *(i)* and *(ii)* in Proposition 2. Moreover, the values of c and γ imply that that $z_{\alpha_{N,T}} = -1.64 \sim z_{0.05}$ when $N = \min(N_1, N_2) = 40$, and $T = 35$, which are analogous to the smallest cross-sectional and time series dimensions in our empirical application. This choice of the functional form for $z_{\alpha_{N,T}}$, and the parameters γ and c proved to work well in all our MC simulation experiments.
The estimation of k_1 and k_2, which is a necessary first step in all the three procedures described above, is based on the information criteria IC_{p2} or IC_{p3} of Bai and Ng (2002), and thoroughly discussed in Section E.4.2.

E.4 Discussion of results

E.4.1 Size and power properties

We are interested in verifying whether the Gaussian asymptotic distribution provides a good small sample approximation for the left tail of the re-centered and re-scaled infeasible statistic $\hat{\xi}(k^C)$, and the feasible $\tilde{\xi}(k^C)$. We compute the empirical size of the test for the null hypothesis of k^C common factors corresponding to nominal sizes of 1%, 5% and 10%. We also report the empirical power of the feasible statistic for the null hypothesis of $k^C + 1$ common factors, when the true number of common factors is k^C, for the same nominal sizes. Following the discussion in Section 5 and Subsection D.9, we consider both (1) factors estimated via PCA applied to HF data, and then flow sampled in order to compute the test statistic, and (2) factors estimated on flow sampled HF data. For case (1) the variance of the flow sampled HF innovations residuals, denoted by $\gamma_{1,ii} = V(\bar{e}_t^{Hi}) = MV(e_{m,t}^{Hi})$, is estimated from the HF residuals $\hat{\varepsilon}_{1,i,m,t}$ obtained from the regressions of the HF data on the estimated common and HF factors. The estimator is: $\tilde{\gamma}_{1,ii} = \dfrac{1}{T}\sum_{m=1}^{M}\sum_{t=1}^{T}\hat{\varepsilon}_{1,i,m,t}^2$. For sample sizes ($T \leq 200$) this estimator improved by an amount of 0.01 - 0.08 all the empirical sizes with respect to an estimator using the residuals from regressions of flow sampled HF data on flow sampled factors.

Infeasible statistic: size

The tables in Section E.5 show that the asymptotic Gaussian distribution provides a very good approximation for the left tail of the infeasible test statistic $\hat{\xi}(k^C)$ under the null, even for sample sizes as small as $N_H = N_L = 50$, and $T = 35$. For the vast majority of sample sizes, and simulation designs, the size distortions for aforementioned case (1) are in the order of 1% to maximum 3%. Analogous results hold for case (2) PCA is performed on flow sampled HF data, with the exception of Designs 2 and 3, where the number of common factors is $k^C = 2$ (see discussion below). For instance, in the baseline Design 1 in which $k^C = k^H = k^L = 1$ and all factors and idiosyncratic innovations are i.i.d. in cross-section and over time, the maximum size distortion is 0.02 and is observed only when $T \leq 50$. The same results hold for DGPs with the same number of factors as in Design 1, but featuring a moderate level of cross-sectional correlation among the idiosyncratic innovations as measured by the coefficient $\beta = 0.2$ in Design 4, or a moderately high level of correlation ($\phi = 0.7$) among the specific factors in the two panels for Design 6. Analogous results hold also when the factors feature an autocorrelation coefficient similar to the one in the empirical analysis, that is $a_F = 0.6$. This can be seen in the tables for Design 5, where the only notable difference compared to the baseline case $a_F = 0$ is an increase in the empirical size of a maximum of 0.05 for the smaller sample sizes.
For Design 2 (resp. Design 3) in which $k^C = 2$ and $k^H = k^L = 0$ (resp. $k^H = k^L = 1$), for sample sizes as small as $T \leq 50$ and $\max(N_H, N_L) \leq 200$, the size distortions increase to a maximum of 6% (resp. 10%), which occurs when PCA is performed post aggregation. This result is due to the fact that, by construction, the signal-to-noise ratio for each of the two common factors in these designs is halved compared to those with

$k^C = 1$. In unreported simulation results, available upon request, we increased the signal-to-noise ratio of the common factors and - as expected - we noticed a reduction in all size distortions, for all designs and sample sizes. This reduction is more pronounced for smaller sample sizes, and is discussed below also for the feasible statistics. We finally note that in the two designs with $k^C = 2$ performing PCA on HF data has the effect of approximately halving the size distortions, compared to the case in which PCA is performed on flow-sampled HF data.

Moreover, we note that the infeasible test seems to be undersized for sample sizes as small as $T \leq 200$ only when the number of specific factors is high in both panels. As shown by the values close to 0 for the empirical sizes of the infeasible statistic for Designs 8 - 11, where $k^H = k^L = 5$, this effect is independent of the level of the correlation among the specific factors. Importantly, this fact does not affect significantly the performance of the sequential procedure for the selection of the number of common factors (see Section E.4.2).

The size distortions disappear in all simulation designs for large values of N_H, N_L, and T, which corroborates our asymptotic theory of Theorem 1. In Particular, for all the Designs 1 - 7, that is when $k^C = 1$ or 2, and $k^H = k^L = 1$ or 0, the size distortions of all feasible statistics are not larger than 0.01 when $T \geq 200$.

Feasible statistics: size and power

Turning to the feasible statistic $\tilde{\xi}(k^C)$, we note that the size distortions are larger than those of the feasible statistic, when $\max(N_H, N_L) \leq 200$, and $T \leq 50$. As the feasible and infeasible statistics use the same estimates of the canonical correlation $\hat{\rho}_\ell$, the increase in the size distortion is due to the fact that the matrix Σ_U appearing in both the bias and the variance of the test statistics is replaced by its estimator $\hat{\Sigma}_U$ defined in Theorem 2, and matrix $\tilde{\Sigma}_{cc}$ is replaced by I_{k^c}. Nevertheless, as the sample sizes increase all size distortions vanish, consistently with the asymptotic theory developed in Theorem 2. For instance, in the baseline Design 1 we note that when PCA is performed on HF data first, the size distortions increase by a maximum of 0.08 when $T \leq 50$, and by a maximum of 0.02 when $T \leq 300$. The same holds for Designs 4 and 6 where cross-sectional correlation among residuals and autocorrelation in the factors are introduced. As it was the case for the infeasible statistic, performing PCA prior to aggregation yields smaller size distortions than the approach performing PCA post aggregation, by amounts in the range of 0.01 - 0.10, when $T \leq 50$.

Designs 2 and 3, where $k^C = 2$, and $k^H = k^L = 0$ or 1, feature the largest size distortions among all the designs when $T \leq 200$, and $\max(N_H, N_L) \leq 200$ for the same reason discussed above for the infeasible statistic. In these two designs, performing PCA first instead of PCA on the flow sampled HF data, drastically reduces the size distortion: for sample sizes with $T \leq 100$, for instance, the size distortions are halved. As expected, when the signal-to-noise ratio of the common factors is increased, the size distortions monotonically improve for all sample sizes, and especially for the very small ones. More specifically, in unreported simulation results, available upon request, we increased the signal-to-noise of all factors by simulating $R^{2,Hi}_{all}$ and $R^{2,Lj}_{all}$ uniformly in the interval $[0.1, 0.95]$, instead of simulating from our baseline interval $[0.1, 0.8]$. Moreover we also increased the signal-to-noise of common factors only, simulating $R^{2,Hi}_C$ and $R^{2,Lj}_C$ uniformly in the intervals $[0.5R^{2,Hi}_{all},\ 0.90R^{2,Hi}_{all}]$ and $[0.5R^{2,Lj}_{all},\ 0.90R^{2,Lj}_{all}]$, respectively. This last case generated the most evident improvements in all empirical sizes, and especially for the designs in which $k^C = 2$.

The power of the feasible test statistic is always equal, or very close to 1, for all designs with the exception of the cases in which $\min(N_H, N_L) \leq 50$, and $T = 35$. This is a remarkable result as our simulation designs include cases in which the specific factors in the different panels are highly correlated. In Design 7, for instance, the correlation coefficient among the specific factors in the two panels is $\phi = 0.95$. This value is the same for the correlations among the 5 specific factors in each of the two panels in Design 12, and implies that in both these designs the specific factors could be confounded with at least 1 additional common factor. This explains the lower power e.g. in Design 7 for the smallest sample sizes. It is also important to note that in the case of many specific factors as in Designs 8-11, where $k^H = k^L = 5$ and $k^C = 1$, the test is less undersized when performed using the feasible statistic than it is with the infeasible statistics.

Finally, in unreported simulation results, we note that the above results are almost exactly the same when M increases from 4 to 12, and to 21. Conversely, we find a slight increase in size distortions when $M = 1$.

E.4.2 Estimation of the number of factors

We are interested in comparing the following three procedures to determine the number of common factors k^C : (i) our consistent sequential testing procedure defined in Proposition 2, which is based on the feasible test statistics of Theorem 2, and uses the critical values as described in Section E.3, (ii) the selection procedure based on the penalized information criterion of Theorem 3.7 in Chen (2012), and (iii) the three-steps selection procedure proposed by Wang (2012).

The estimators are evaluated by comparing the average, computed across the 4000 MC simulations, estimated number of common, high-frequency-specific, and low-frequency-specific factors. For all the estimators of k^C, k^H, k^L we consider the case in which the true numbers of pervasive factors $k_1 = k^C + k^H$ and $k_2 = k^C + k^L$ in the two panels are known, and only k^C needs to be estimated, and also the case in which k_1 and k_2 are estimated using the IC_p information criteria proposed by Bai and Ng (2002). More specifically, we present estimation results for k_1 and k_2 where we used the IC_{p2} criterion for all designs in which $k^H = k^L = 0$ or 1, and the IC_{p3} criterion when $k^H = k^L = 5$. The same criteria are used to estimate the number of pervasive factors in the stacked panel of HF (flow-sampled) and LF data in the second step of the procedure suggested by Wang (2012). In line with the results of Bai and Ng (2002), we noted that for small sample sizes ($T \leq 50$, and especially for $T = 35$) and in the case of many pervasive factors in the LF panel (that is $k_2 \geq 5$) the IC_{p2} criterion tends to severely underestimate the values of k_2, while the IC_{p3} produces better estimates. Underestimating k_2 affects considerably the estimates of k^C for all the three procedures considered. In unreported results available upon request we have estimated k_1, k_2, and also R (that is the number of pervasive factors in the panel formed by stacking together both the flow-sampled HF data, and the LF data), using the ER and GR ratios of Ahn and Horenstein (2013), and noted that they perform similarly or worse than the IC_{p2} criterion. [14] The first thing to note is that for all the simulation designs considered, the results for the estimation of the number of factors are very similar both in the cases in which k_1 and k_2 are known (lower panels in all the tables in Section E.6), and when they are estimated as we have just described (upper panels in same tables).

We also remark that across all our designs, performing PCA first on HF data instead of performing PCA on the flow-sampled HF data, produces consistently better estimation of the number of factors for all the sample sizes in which $T \geq 50$. The same holds true also for sample sizes as small as $T = 35$, with the exception of Designs 7 - 11 when $N = \min(N_H, N_L) = 50$. For the same small value of T, as soon as $N > 50$, PCA on HF data is always preferred to PCA on flow-sampled HF data, with the exception of Design 11, where this happens only for $T > 100$.

In all the designs with $k^C = 1$ common factor, and $k^H = k^L = 1$, zero or moderate values for the correlation of specific factors - that is when $\phi = 0$ (Designs 1, 4, and 5) or 0.7 (Design 7) - the average estimated number of common factors obtained with our procedure ranges between 0.90 and 1, that is below but close to the true value $k^C = 1$. Analogous results hold for the designs in which $k^C = 2$ (Designs 2 and 3), where the estimated number of common factors is always below 2. These results confirm both the ones on the very good empirical power of the test statistics, and are also compatible with the fact that the statistic is slightly oversized in (very) small samples, for these designs.

As predicted by the consistency result for our sequential testing procedure in Proposition 2, and by the empirical size and power properties of the feasible statistics, the average estimated number of common factors for our selection procedure approaches quickly the true value k^C as the sample sizes increase in all our designs. This is true also for the most challenging Designs 7 and 12, in which the specific factors in the different panels feature an extremely high value of correlation, that is $\phi = 0.95$.

[14] Alternative estimators, such as the one proposed by Onatski (2010), could also be considered.

In cases with a small number, say 1 or 3, of uncorrelated specific factors (that is when $\phi = 0$), the penalized information criterion proposed in Chen (2012) yields the correct number of factors in almost all Monte Carlo simulations for any sample size, confirming the results in Chen (2012). To save space only the tables of results for the case $k^C = 1$ (or 2) and $k^H = k^L = 1$ have been reported, and correspond to Designs 1 - 5. On the other hand, the tables for the case $k^H = k^L = 3$ are available upon request, and are analogous to the ones for the case $k^H = k^L = 1$. For the same DGPs our selection procedure has comparable performance as that in Chen (2012), and it is less accurate than Chen (2012)'s one only for sample sizes as small as $\max(N_H, N_L) \leq 200$, and $T \leq 50$, where the average estimated number of common factors ranges between 0.85 and 1 if $k^C = 1$. In particular, for the baseline Design 1 in which $k^C = k^H = k^L = 1$, and all factors and idiosyncratic errors are uncorrelated both in cross-section and over time, the average estimated value of k^C is between 0.89 and 0.94 when $T \leq 50$. The same holds for a moderate value of cross-sectional correlation among residuals as in Design 4. When $a_F = 0.6$, $k^C = k^H = k^L = 1$, and for small sample sizes, we note a moderate deterioration of the performance of our sequential testing procedure compared to the case $a_F = 0.0$, that is when we compare the results in Designs 1 and 5, respectively. In this case our procedure tends to slightly underestimate the number of common factors as expected from the increase of the empirical size discussed in Section E.4.1, with the minimum average estimated value for k^C equal to 0.84 only when $T \leq 35$.

The procedure of Chen (2012) tends to overestimate the number of common factors when the correlation ϕ among the specific factors increases from 0 (see Designs 1 and 8 for the cases $k^C = 1$ and $k^H = k^L = 5$, respectively) to 0.5 (Design 9 for the cases $k^C = 1$ and $k^H = k^L = 5$ only[15]), 0.7 (Designs 6 and 10, for the cases $k^C = 1$ and $k^H = k^L = 5$, respectively) and 0.95 (Designs 7 and 11, for the cases $k^C = 1$ and $k^H = k^L = 5$, respectively). This deterioration in the performance is much less dramatic for our sequential testing procedure. As expected from our results on the empirical power in the previous section, in Designs 8 - 11 we also observe a monotonic decrease in the precision across all the estimators when the number of specific factors becomes relatively large, namely $k^H = k^L = 5$. In this case all the three procedures considered tend to overestimate the true number of common factors, namely $k^C = 1$. Nevertheless, in all Designs 8 -11 when $T \leq 50$ our procedure consistently outperforms Chen (2012). Importantly, this result holds true also when the specific factors in the two panels are not correlated, as in Design 8 where $\phi = 0$. For larger values of the correlation coefficient ϕ, the better performance of our procedure is even more evident also in larger sample sizes (see Designs 9 - 11). It is noteworthy that as ϕ increases the deterioration for our sequential procedure is much less dramatic than in Chen (2012), suggesting that it is preferable in these more general cases. Furthermore, our sequential testing procedure also exhibits a faster improvement in performance as the sample size increases.

Finally, the consistent three-steps selection procedure of Wang (2012) performs similarly, or worse than the one of Chen (2012) in DGPs with a small number of uncorrelated specific factors. More specifically, when $k^C = 1$ as in Designs 1, and 4 - 11, Wang (2012) procedure tends to overestimate the true value of k^C for sample sizes as small as $T \leq 50$. Moreover, as either the number of specific factors increases (Design 8), or ϕ increases from 0 to 0.7 (see Designs 6 and 10) or 0.95 (see Designs 7 and 11) the procedure overestimates k^C also for sample sizes as large as $N = 500$ and $T = 100$ (Design 6), or $N = 800$ and $T = 500$ (Designs 10 and 11), and clearly becomes the worse among the three considered.

[15] Due to space limitations the results for the case ϕ = 0.5, when $k^C = k^L = k^H = 1$ have not been reported, but are available upon request.

E.5 Tables: Size and Power

DESIGN 1 : $k^C = 1$, $k^H = k^L = 1$, $\beta = 0.0$, $a_F = 0.0$, $\phi = 0.0$, $a_e = 0.0$, $R^2_{max} = 0.8$

			SIZE						**POWER**		
			Infeasible			Feasible			Feasible		
			Nominal size						*Nominal size*		
N_H	N_L	T	0.01	0.05	0.10	0.01	0.05	0.10	0.01	0.05	0.10
HF data: PCA first, then flow sample the factors											
50	50	35	0.02	0.05	0.08	0.03	0.08	0.14	0.93	0.98	0.99
100	50	35	0.03	0.06	0.10	0.04	0.12	0.18	0.95	0.98	0.99
100	100	35	0.02	0.05	0.09	0.04	0.10	0.15	1.00	1.00	1.00
100	100	50	0.02	0.05	0.09	0.03	0.08	0.14	1.00	1.00	1.00
200	100	50	0.02	0.06	0.10	0.04	0.11	0.18	1.00	1.00	1.00
200	200	50	0.02	0.05	0.09	0.03	0.09	0.15	1.00	1.00	1.00
200	200	100	0.01	0.05	0.09	0.02	0.07	0.12	1.00	1.00	1.00
500	500	100	0.01	0.05	0.08	0.02	0.07	0.13	1.00	1.00	1.00
500	500	200	0.01	0.05	0.09	0.02	0.06	0.12	1.00	1.00	1.00
500	500	300	0.01	0.05	0.09	0.02	0.06	0.12	1.00	1.00	1.00
800	800	500	0.01	0.05	0.09	0.01	0.05	0.10	1.00	1.00	1.00
1000	1000	600	0.01	0.05	0.10	0.01	0.06	0.11	1.00	1.00	1.00
HF data: flow sample first, then PCA											
50	50	35	0.03	0.06	0.10	0.06	0.15	0.23	0.94	0.98	0.99
100	50	35	0.03	0.07	0.11	0.07	0.17	0.25	0.95	0.98	0.99
100	100	35	0.03	0.07	0.11	0.08	0.18	0.26	1.00	1.00	1.00
100	100	50	0.02	0.06	0.10	0.05	0.13	0.22	1.00	1.00	1.00
200	100	50	0.03	0.07	0.12	0.06	0.15	0.24	1.00	1.00	1.00
200	200	50	0.03	0.07	0.11	0.06	0.15	0.23	1.00	1.00	1.00
200	200	100	0.02	0.06	0.10	0.03	0.10	0.17	1.00	1.00	1.00
500	500	100	0.02	0.06	0.10	0.04	0.11	0.18	1.00	1.00	1.00
500	500	200	0.01	0.06	0.10	0.02	0.08	0.15	1.00	1.00	1.00
500	500	300	0.02	0.05	0.10	0.02	0.08	0.14	1.00	1.00	1.00
800	800	500	0.01	0.05	0.09	0.02	0.06	0.12	1.00	1.00	1.00
1000	1000	600	0.01	0.06	0.10	0.01	0.07	0.12	1.00	1.00	1.00

DESIGN 2 : $k^C = 2$, $k^H = k^L = 0$, $\beta = 0.0$, $a_F = 0.0$, $\phi = 0.0$, $a_e = 0.0$, $R^2_{max} = 0.8$

			SIZE						POWER		
			Infeasible			Feasible			Feasible		
			Nominal size						*Nominal size*		
N_H	N_L	T	0.01	0.05	0.10	0.01	0.05	0.10	0.01	0.05	0.10
HF data: PCA first, then flow sample the factors											
50	50	35	0.03	0.07	0.11	0.03	0.10	0.17	-	-	-
100	50	35	0.03	0.07	0.12	0.05	0.15	0.23	-	-	-
100	100	35	0.03	0.08	0.12	0.06	0.15	0.23	-	-	-
100	100	50	0.02	0.06	0.10	0.03	0.10	0.17	-	-	-
200	100	50	0.03	0.07	0.11	0.05	0.14	0.22	-	-	-
200	200	50	0.02	0.07	0.10	0.05	0.13	0.20	-	-	-
200	200	100	0.01	0.05	0.10	0.02	0.08	0.14	-	-	-
500	500	100	0.01	0.05	0.10	0.03	0.09	0.16	-	-	-
500	500	200	0.01	0.05	0.09	0.02	0.07	0.12	-	-	-
500	500	300	0.01	0.05	0.09	0.01	0.06	0.11	-	-	-
800	800	500	0.01	0.05	0.09	0.02	0.06	0.12	-	-	-
1000	1000	600	0.01	0.05	0.10	0.02	0.06	0.12	-	-	-
HF data: flow sample first, then PCA											
50	50	35	0.04	0.09	0.15	0.10	0.23	0.34	-	-	-
100	50	35	0.04	0.09	0.14	0.10	0.24	0.34	-	-	-
100	100	35	0.05	0.10	0.16	0.15	0.30	0.41	-	-	-
100	100	50	0.03	0.08	0.12	0.08	0.20	0.30	-	-	-
200	100	50	0.03	0.08	0.13	0.09	0.21	0.32	-	-	-
200	200	50	0.03	0.08	0.13	0.11	0.23	0.33	-	-	-
200	200	100	0.02	0.06	0.11	0.04	0.13	0.22	-	-	-
500	500	100	0.02	0.07	0.12	0.05	0.15	0.25	-	-	-
500	500	200	0.01	0.05	0.10	0.03	0.10	0.17	-	-	-
500	500	300	0.01	0.06	0.10	0.02	0.09	0.14	-	-	-
800	800	500	0.01	0.05	0.10	0.02	0.08	0.15	-	-	-
1000	1000	600	0.01	0.05	0.10	0.02	0.08	0.14	-	-	-

DESIGN 3 : $k^C = 2$, $k^H = k^L = 1$, $\beta = 0.0$, $a_F = 0.0$, $\phi = 0.0$, $a_e = 0.0$, $R^2_{max} = 0.8$

			SIZE						POWER		
			Infeasible			Feasible			Feasible		
			Nominal size						*Nominal size*		
N_H	N_L	T	0.01	0.05	0.10	0.01	0.05	0.10	0.01	0.05	0.10
HF data: PCA first, then flow sample the factors											
50	50	35	0.03	0.07	0.10	0.09	0.22	0.32	0.91	0.96	0.98
100	50	35	0.06	0.11	0.15	0.18	0.32	0.42	0.94	0.98	0.99
100	100	35	0.03	0.07	0.12	0.13	0.27	0.38	1.00	1.00	1.00
100	100	50	0.02	0.06	0.10	0.07	0.18	0.28	1.00	1.00	1.00
200	100	50	0.03	0.08	0.13	0.11	0.26	0.37	1.00	1.00	1.00
200	200	50	0.03	0.07	0.11	0.09	0.22	0.31	1.00	1.00	1.00
200	200	100	0.01	0.05	0.09	0.04	0.11	0.19	1.00	1.00	1.00
500	500	100	0.02	0.06	0.10	0.05	0.14	0.23	1.00	1.00	1.00
500	500	200	0.01	0.04	0.09	0.02	0.09	0.17	1.00	1.00	1.00
500	500	300	0.01	0.05	0.09	0.02	0.08	0.14	1.00	1.00	1.00
800	800	500	0.01	0.04	0.09	0.02	0.07	0.14	1.00	1.00	1.00
1000	1000	600	0.01	0.04	0.08	0.02	0.06	0.12	1.00	1.00	1.00
HF data: flow sample first, then PCA											
50	50	35	0.08	0.14	0.19	0.29	0.48	0.59	0.93	0.97	0.99
100	50	35	0.08	0.14	0.20	0.30	0.49	0.60	0.94	0.98	0.99
100	100	35	0.07	0.14	0.20	0.35	0.55	0.65	1.00	1.00	1.00
100	100	50	0.04	0.10	0.16	0.21	0.41	0.53	1.00	1.00	1.00
200	100	50	0.05	0.11	0.17	0.22	0.41	0.52	1.00	1.00	1.00
200	200	50	0.05	0.13	0.18	0.26	0.44	0.55	1.00	1.00	1.00
200	200	100	0.03	0.07	0.12	0.09	0.23	0.34	1.00	1.00	1.00
500	500	100	0.03	0.09	0.14	0.12	0.27	0.38	1.00	1.00	1.00
500	500	200	0.02	0.07	0.12	0.05	0.17	0.26	1.00	1.00	1.00
500	500	300	0.01	0.06	0.11	0.04	0.13	0.20	1.00	1.00	1.00
800	800	500	0.01	0.05	0.11	0.03	0.10	0.18	1.00	1.00	1.00
1000	1000	600	0.01	0.05	0.10	0.03	0.09	0.16	1.00	1.00	1.00

DESIGN 4 : $k^C = 1$, $k^H = k^L = 1$, $\beta = 0.2$, $a_F = 0.0$, $\phi = 0.0$, $a_e = 0.0$, $R^2_{max} = 0.8$

			SIZE						POWER		
			Infeasible			Feasible			Feasible		
			Nominal size						*Nominal size*		
N_H	N_L	T	0.01	0.05	0.10	0.01	0.05	0.10	0.01	0.05	0.10
HF data: PCA first, then flow sample the factors											
50	50	35	0.02	0.05	0.08	0.03	0.09	0.14	0.93	0.98	0.99
100	50	35	0.03	0.06	0.10	0.04	0.11	0.17	0.94	0.98	0.99
100	100	35	0.02	0.05	0.09	0.04	0.10	0.17	1.00	1.00	1.00
100	100	50	0.02	0.05	0.08	0.03	0.08	0.14	1.00	1.00	1.00
200	100	50	0.02	0.06	0.10	0.04	0.11	0.17	1.00	1.00	1.00
200	200	50	0.02	0.06	0.09	0.03	0.09	0.15	1.00	1.00	1.00
200	200	100	0.02	0.05	0.09	0.02	0.07	0.13	1.00	1.00	1.00
500	500	100	0.02	0.05	0.09	0.03	0.07	0.13	1.00	1.00	1.00
500	500	200	0.02	0.05	0.10	0.02	0.07	0.12	1.00	1.00	1.00
500	500	300	0.01	0.04	0.09	0.01	0.06	0.11	1.00	1.00	1.00
800	800	500	0.01	0.05	0.10	0.01	0.06	0.11	1.00	1.00	1.00
1000	1000	600	0.01	0.05	0.09	0.01	0.06	0.11	1.00	1.00	1.00
HF data: flow sample first, then PCA											
50	50	35	0.03	0.07	0.10	0.07	0.16	0.23	0.94	0.98	0.99
100	50	35	0.03	0.07	0.11	0.07	0.16	0.25	0.95	0.98	0.99
100	100	35	0.03	0.07	0.11	0.09	0.19	0.28	1.00	1.00	1.00
100	100	50	0.02	0.06	0.10	0.05	0.14	0.21	1.00	1.00	1.00
200	100	50	0.03	0.07	0.11	0.06	0.15	0.22	1.00	1.00	1.00
200	200	50	0.03	0.06	0.11	0.07	0.16	0.23	1.00	1.00	1.00
200	200	100	0.02	0.06	0.10	0.03	0.10	0.18	1.00	1.00	1.00
500	500	100	0.02	0.06	0.10	0.04	0.11	0.18	1.00	1.00	1.00
500	500	200	0.02	0.06	0.11	0.03	0.09	0.15	1.00	1.00	1.00
500	500	300	0.01	0.05	0.10	0.02	0.07	0.13	1.00	1.00	1.00
800	800	500	0.01	0.05	0.10	0.02	0.07	0.13	1.00	1.00	1.00
1000	1000	600	0.01	0.06	0.10	0.02	0.07	0.13	1.00	1.00	1.00

DESIGN 5 : $k^C = 1$, $k^H = k^L = 1$, $\beta = 0.0$, $a_F = 0.6$, $\phi = 0.0$, $a_e = 0.0$, $R^2_{max} = 0.8$

			SIZE						**POWER**		
			Infeasible			Feasible			Feasible		
			Nominal size						*Nominal size*		
N_H	N_L	T	0.01	0.05	0.10	0.01	0.05	0.10	0.01	0.05	0.10
HF data: PCA first, then flow sample the factors											
50	50	35	0.04	0.09	0.13	0.05	0.13	0.20	0.95	0.98	0.99
100	50	35	0.04	0.09	0.13	0.07	0.15	0.23	0.95	0.98	0.99
100	100	35	0.04	0.08	0.13	0.06	0.14	0.20	1.00	1.00	1.00
100	100	50	0.03	0.07	0.12	0.03	0.11	0.19	1.00	1.00	1.00
200	100	50	0.04	0.08	0.13	0.05	0.14	0.21	1.00	1.00	1.00
200	200	50	0.03	0.07	0.12	0.04	0.11	0.19	1.00	1.00	1.00
200	200	100	0.02	0.06	0.11	0.03	0.08	0.15	1.00	1.00	1.00
500	500	100	0.02	0.06	0.11	0.03	0.09	0.15	1.00	1.00	1.00
500	500	200	0.01	0.06	0.10	0.01	0.07	0.13	1.00	1.00	1.00
500	500	300	0.01	0.05	0.10	0.01	0.06	0.11	1.00	1.00	1.00
800	800	500	0.01	0.05	0.09	0.01	0.06	0.11	1.00	1.00	1.00
1000	1000	600	0.01	0.05	0.10	0.02	0.06	0.12	1.00	1.00	1.00
HF data: flow sample first, then PCA											
50	50	35	0.04	0.09	0.12	0.06	0.15	0.22	0.95	0.98	0.99
100	50	35	0.04	0.09	0.13	0.07	0.16	0.24	0.95	0.98	0.99
100	100	35	0.04	0.08	0.12	0.07	0.16	0.23	1.00	1.00	1.00
100	100	50	0.03	0.07	0.12	0.04	0.13	0.20	1.00	1.00	1.00
200	100	50	0.04	0.08	0.12	0.06	0.15	0.22	1.00	1.00	1.00
200	200	50	0.03	0.07	0.12	0.05	0.13	0.20	1.00	1.00	1.00
200	200	100	0.02	0.06	0.11	0.03	0.09	0.16	1.00	1.00	1.00
500	500	100	0.02	0.06	0.11	0.03	0.10	0.17	1.00	1.00	1.00
500	500	200	0.01	0.05	0.10	0.02	0.08	0.14	1.00	1.00	1.00
500	500	300	0.01	0.05	0.10	0.02	0.07	0.12	1.00	1.00	1.00
800	800	500	0.01	0.05	0.09	0.01	0.06	0.11	1.00	1.00	1.00
1000	1000	600	0.01	0.05	0.10	0.01	0.06	0.12	1.00	1.00	1.00

DESIGN 6 : $k^C = 1$, $k^H = k^L = 1$, $\beta = 0.0$, $a_F = 0.0$, $\phi = 0.70$, $a_e = 0.0$, $R^2_{max} = 0.8$

			SIZE						POWER		
			Infeasible			Feasible			Feasible		
			Nominal size						*Nominal size*		
N_H	N_L	T	0.01	0.05	0.10	0.01	0.05	0.10	0.01	0.05	0.10
HF data: PCA first, then flow sample the factors											
50	50	35	0.02	0.04	0.07	0.03	0.08	0.13	1.00	1.00	1.00
100	50	35	0.03	0.06	0.09	0.04	0.11	0.18	1.00	1.00	1.00
100	100	35	0.02	0.05	0.09	0.04	0.10	0.17	1.00	1.00	1.00
100	100	50	0.02	0.05	0.08	0.02	0.08	0.14	1.00	1.00	1.00
200	100	50	0.02	0.06	0.10	0.04	0.11	0.17	1.00	1.00	1.00
200	200	50	0.02	0.05	0.08	0.03	0.09	0.15	1.00	1.00	1.00
200	200	100	0.01	0.05	0.09	0.02	0.07	0.12	1.00	1.00	1.00
500	500	100	0.02	0.06	0.09	0.02	0.08	0.14	1.00	1.00	1.00
500	500	200	0.02	0.05	0.09	0.02	0.06	0.11	1.00	1.00	1.00
500	500	300	0.01	0.04	0.08	0.01	0.05	0.11	1.00	1.00	1.00
800	800	500	0.01	0.05	0.09	0.01	0.05	0.11	1.00	1.00	1.00
1000	1000	600	0.01	0.04	0.09	0.01	0.05	0.10	1.00	1.00	1.00
HF data: flow sample first, then PCA											
50	50	35	0.03	0.06	0.09	0.07	0.15	0.23	1.00	1.00	1.00
100	50	35	0.03	0.07	0.10	0.07	0.16	0.25	1.00	1.00	1.00
100	100	35	0.03	0.07	0.10	0.08	0.19	0.27	1.00	1.00	1.00
100	100	50	0.02	0.06	0.10	0.05	0.14	0.23	1.00	1.00	1.00
200	100	50	0.03	0.07	0.10	0.06	0.14	0.22	1.00	1.00	1.00
200	200	50	0.03	0.06	0.10	0.06	0.15	0.23	1.00	1.00	1.00
200	200	100	0.02	0.06	0.10	0.04	0.10	0.17	1.00	1.00	1.00
500	500	100	0.02	0.07	0.11	0.04	0.12	0.19	1.00	1.00	1.00
500	500	200	0.02	0.05	0.09	0.03	0.08	0.14	1.00	1.00	1.00
500	500	300	0.01	0.05	0.09	0.02	0.07	0.13	1.00	1.00	1.00
800	800	500	0.01	0.05	0.10	0.02	0.07	0.12	1.00	1.00	1.00
1000	1000	600	0.01	0.05	0.09	0.02	0.06	0.12	1.00	1.00	1.00

DESIGN 7 : $k^C = 1$, $k^H = k^L = 1$, $\beta = 0.0$, $a_F = 0.0$, $\phi = 0.95$, $a_e = 0.0$, $R^2_{max} = 0.8$

			SIZE						POWER		
			Infeasible			Feasible			Feasible		
			Nominal size						*Nominal size*		
N_H	N_L	T	0.01	0.05	0.10	0.01	0.05	0.10	0.01	0.05	0.10
HF data: PCA first, then flow sample the factors											
50	50	35	0.01	0.02	0.03	0.04	0.10	0.15	0.53	0.72	0.81
100	50	35	0.00	0.02	0.04	0.06	0.12	0.18	0.75	0.87	0.92
100	100	35	0.01	0.03	0.05	0.05	0.11	0.18	0.93	0.98	0.98
100	100	50	0.01	0.03	0.06	0.03	0.09	0.14	0.98	0.99	1.00
200	100	50	0.02	0.05	0.07	0.04	0.11	0.16	1.00	1.00	1.00
200	200	50	0.01	0.04	0.08	0.03	0.09	0.15	1.00	1.00	1.00
200	200	100	0.01	0.04	0.08	0.02	0.07	0.12	1.00	1.00	1.00
500	500	100	0.02	0.05	0.09	0.03	0.08	0.13	1.00	1.00	1.00
500	500	200	0.01	0.05	0.09	0.02	0.07	0.12	1.00	1.00	1.00
500	500	300	0.01	0.05	0.08	0.01	0.06	0.10	1.00	1.00	1.00
800	800	500	0.01	0.05	0.10	0.01	0.07	0.12	1.00	1.00	1.00
1000	1000	600	0.01	0.05	0.09	0.01	0.05	0.11	1.00	1.00	1.00
HF data: flow sample first, then PCA											
50	50	35	0.01	0.02	0.04	0.08	0.17	0.24	0.68	0.84	0.90
100	50	35	0.01	0.03	0.05	0.08	0.17	0.24	0.81	0.91	0.95
100	100	35	0.01	0.04	0.07	0.09	0.19	0.27	0.97	0.99	0.99
100	100	50	0.02	0.04	0.07	0.06	0.14	0.22	0.99	1.00	1.00
200	100	50	0.02	0.05	0.09	0.06	0.14	0.21	1.00	1.00	1.00
200	200	50	0.02	0.06	0.10	0.06	0.15	0.23	1.00	1.00	1.00
200	200	100	0.01	0.05	0.09	0.03	0.10	0.17	1.00	1.00	1.00
500	500	100	0.02	0.06	0.10	0.04	0.12	0.19	1.00	1.00	1.00
500	500	200	0.02	0.06	0.10	0.03	0.09	0.15	1.00	1.00	1.00
500	500	300	0.01	0.05	0.09	0.02	0.07	0.13	1.00	1.00	1.00
800	800	500	0.01	0.06	0.11	0.02	0.08	0.14	1.00	1.00	1.00
1000	1000	600	0.01	0.05	0.10	0.02	0.06	0.13	1.00	1.00	1.00

DESIGN 8 : $k^C = 1$, $k^H = k^L = 5$, $\beta = 0.0$, $a_F = 0.0$, $\phi = 0.00$, $a_e = 0.0$, $R^2_{max} = 0.8$

			SIZE						**POWER**		
			Infeasible			Feasible			Feasible		
			Nominal size						*Nominal size*		
N_H	N_L	T	0.01	0.05	0.10	0.01	0.05	0.10	0.01	0.05	0.10
HF data: PCA first, then flow sample the factors											
50	50	35	0.00	0.00	0.01	0.00	0.01	0.03	0.27	0.56	0.72
100	50	35	0.00	0.00	0.01	0.01	0.04	0.08	0.75	0.92	0.97
100	100	35	0.00	0.01	0.01	0.01	0.03	0.05	1.00	1.00	1.00
100	100	50	0.00	0.01	0.02	0.01	0.04	0.06	1.00	1.00	1.00
200	100	50	0.00	0.01	0.02	0.02	0.04	0.08	1.00	1.00	1.00
200	200	50	0.00	0.01	0.02	0.01	0.03	0.06	1.00	1.00	1.00
200	200	100	0.00	0.01	0.03	0.01	0.04	0.07	1.00	1.00	1.00
500	500	100	0.01	0.02	0.03	0.01	0.04	0.07	1.00	1.00	1.00
500	500	200	0.01	0.02	0.05	0.01	0.04	0.08	1.00	1.00	1.00
500	500	300	0.00	0.02	0.05	0.01	0.04	0.08	1.00	1.00	1.00
800	800	500	0.01	0.03	0.06	0.01	0.05	0.09	1.00	1.00	1.00
1000	1000	600	0.01	0.03	0.06	0.01	0.04	0.09	1.00	1.00	1.00
HF data: flow sample first, then PCA											
50	50	35	0.00	0.01	0.01	0.02	0.09	0.14	0.60	0.85	0.93
100	50	35	0.00	0.01	0.02	0.04	0.11	0.17	0.86	0.97	0.99
100	100	35	0.00	0.01	0.02	0.05	0.12	0.16	1.00	1.00	1.00
100	100	50	0.00	0.01	0.03	0.04	0.10	0.16	1.00	1.00	1.00
200	100	50	0.00	0.02	0.03	0.03	0.09	0.14	1.00	1.00	1.00
200	200	50	0.00	0.01	0.03	0.03	0.09	0.15	1.00	1.00	1.00
200	200	100	0.00	0.02	0.03	0.02	0.08	0.13	1.00	1.00	1.00
500	500	100	0.01	0.02	0.04	0.02	0.07	0.12	1.00	1.00	1.00
500	500	200	0.01	0.03	0.05	0.02	0.07	0.13	1.00	1.00	1.00
500	500	300	0.00	0.03	0.06	0.01	0.06	0.12	1.00	1.00	1.00
800	800	500	0.01	0.03	0.06	0.02	0.06	0.11	1.00	1.00	1.00
1000	1000	600	0.01	0.03	0.06	0.01	0.06	0.11	1.00	1.00	1.00

DESIGN 9 : $k^C = 1$, $k^H = k^L = 5$, $\beta = 0.0$, $a_F = 0.0$, $\phi = 0.5$, $a_e = 0.0$, $R^2_{max} = 0.8$

			SIZE						POWER		
			Infeasible			Feasible			Feasible		
			Nominal size						*Nominal size*		
N_H	N_L	T	0.01	0.05	0.10	0.01	0.05	0.10	0.01	0.05	0.10
HF data: PCA first, then flow sample the factors											
50	50	35	0.00	0.00	0.01	0.00	0.01	0.03	0.14	0.36	0.51
100	50	35	0.00	0.01	0.01	0.01	0.04	0.07	0.55	0.80	0.89
100	100	35	0.00	0.00	0.01	0.01	0.02	0.04	0.96	0.98	0.99
100	100	50	0.00	0.01	0.01	0.01	0.03	0.05	0.99	1.00	1.00
200	100	50	0.00	0.01	0.02	0.01	0.04	0.08	1.00	1.00	1.00
200	200	50	0.00	0.01	0.02	0.01	0.04	0.06	1.00	1.00	1.00
200	200	100	0.00	0.01	0.03	0.01	0.03	0.06	1.00	1.00	1.00
500	500	100	0.00	0.02	0.03	0.01	0.04	0.07	1.00	1.00	1.00
500	500	200	0.00	0.02	0.04	0.01	0.03	0.07	1.00	1.00	1.00
500	500	300	0.01	0.02	0.05	0.01	0.04	0.08	1.00	1.00	1.00
800	800	500	0.00	0.03	0.05	0.01	0.04	0.08	1.00	1.00	1.00
1000	1000	600	0.01	0.03	0.06	0.01	0.04	0.09	1.00	1.00	1.00
HF data: flow sample first, then PCA											
50	50	35	0.00	0.01	0.01	0.02	0.07	0.12	0.43	0.73	0.84
100	50	35	0.00	0.01	0.01	0.04	0.10	0.14	0.77	0.92	0.96
100	100	35	0.00	0.01	0.02	0.04	0.10	0.15	0.99	1.00	1.00
100	100	50	0.00	0.01	0.02	0.03	0.09	0.15	1.00	1.00	1.00
200	100	50	0.00	0.01	0.02	0.03	0.09	0.15	1.00	1.00	1.00
200	200	50	0.00	0.01	0.03	0.04	0.09	0.15	1.00	1.00	1.00
200	200	100	0.00	0.02	0.03	0.02	0.07	0.13	1.00	1.00	1.00
500	500	100	0.00	0.02	0.04	0.03	0.08	0.13	1.00	1.00	1.00
500	500	200	0.01	0.02	0.04	0.01	0.06	0.11	1.00	1.00	1.00
500	500	300	0.01	0.03	0.05	0.02	0.06	0.11	1.00	1.00	1.00
800	800	500	0.01	0.03	0.06	0.01	0.06	0.11	1.00	1.00	1.00
1000	1000	600	0.01	0.03	0.06	0.01	0.06	0.11	1.00	1.00	1.00

DESIGN 10 : $k^C = 1$, $k^H = k^L = 5$, $\beta = 0.0$, $a_F = 0.0$, $\phi = 0.7$, $a_e = 0.0$, $R^2_{max} = 0.8$

			SIZE						**POWER**		
			Infeasible			Feasible			Feasible		
			Nominal size						*Nominal size*		
N_H	N_L	T	0.01	0.05	0.10	0.01	0.05	0.10	0.01	0.05	0.10
HF data: PCA first, then flow sample the factors											
50	50	35	0.00	0.00	0.00	0.00	0.01	0.02	0.04	0.16	0.28
100	50	35	0.00	0.00	0.01	0.01	0.02	0.05	0.27	0.54	0.66
100	100	35	0.00	0.00	0.01	0.01	0.02	0.04	0.76	0.88	0.92
100	100	50	0.00	0.01	0.01	0.01	0.02	0.05	0.93	0.98	0.99
200	100	50	0.00	0.01	0.02	0.02	0.04	0.08	1.00	1.00	1.00
200	200	50	0.00	0.01	0.01	0.01	0.03	0.05	1.00	1.00	1.00
200	200	100	0.00	0.01	0.02	0.01	0.03	0.06	1.00	1.00	1.00
500	500	100	0.00	0.02	0.03	0.01	0.04	0.08	1.00	1.00	1.00
500	500	200	0.00	0.02	0.05	0.01	0.04	0.09	1.00	1.00	1.00
500	500	300	0.00	0.02	0.05	0.01	0.04	0.09	1.00	1.00	1.00
800	800	500	0.01	0.03	0.05	0.01	0.04	0.08	1.00	1.00	1.00
1000	1000	600	0.01	0.03	0.06	0.01	0.04	0.09	1.00	1.00	1.00
HF data: flow sample first, then PCA											
50	50	35	0.00	0.00	0.01	0.01	0.05	0.09	0.24	0.52	0.67
100	50	35	0.00	0.00	0.01	0.02	0.06	0.12	0.53	0.77	0.85
100	100	35	0.00	0.01	0.01	0.03	0.09	0.13	0.92	0.97	0.98
100	100	50	0.00	0.01	0.02	0.03	0.08	0.13	0.98	0.99	1.00
200	100	50	0.00	0.01	0.02	0.03	0.08	0.13	1.00	1.00	1.00
200	200	50	0.00	0.01	0.02	0.03	0.08	0.13	1.00	1.00	1.00
200	200	100	0.00	0.02	0.03	0.02	0.07	0.12	1.00	1.00	1.00
500	500	100	0.00	0.02	0.04	0.03	0.08	0.13	1.00	1.00	1.00
500	500	200	0.01	0.03	0.05	0.02	0.07	0.12	1.00	1.00	1.00
500	500	300	0.00	0.03	0.06	0.02	0.07	0.12	1.00	1.00	1.00
800	800	500	0.01	0.03	0.06	0.01	0.05	0.11	1.00	1.00	1.00
1000	1000	600	0.01	0.03	0.07	0.01	0.06	0.11	1.00	1.00	1.00

DESIGN 11 : $k^C = 1$, $k^H = k^L = 5$, $\beta = 0.0$, $a_F = 0.0$, $\phi = 0.95$, $a_e = 0.0$, $R^2_{max} = 0.8$

			SIZE						POWER		
			Infeasible			Feasible			Feasible		
			Nominal size						*Nominal size*		
N_H	N_L	T	0.01	0.05	0.10	0.01	0.05	0.10	0.01	0.05	0.10
HF data: PCA first, then flow sample the factors											
50	50	35	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.01
100	50	35	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.02	0.05
100	100	35	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.04	0.08
100	100	50	0.00	0.00	0.01	0.00	0.01	0.02	0.02	0.08	0.14
200	100	50	0.00	0.00	0.00	0.00	0.01	0.03	0.10	0.24	0.34
200	200	50	0.00	0.01	0.01	0.00	0.02	0.03	0.28	0.46	0.56
200	200	100	0.00	0.01	0.02	0.00	0.02	0.05	0.73	0.87	0.93
500	500	100	0.00	0.01	0.03	0.01	0.03	0.06	1.00	1.00	1.00
500	500	200	0.00	0.02	0.04	0.01	0.04	0.07	1.00	1.00	1.00
500	500	300	0.00	0.02	0.04	0.01	0.03	0.07	1.00	1.00	1.00
800	800	500	0.01	0.03	0.05	0.01	0.04	0.09	1.00	1.00	1.00
1000	1000	600	0.00	0.02	0.05	0.01	0.04	0.08	1.00	1.00	1.00
HF data: flow sample first, then PCA											
50	50	35	0.00	0.00	0.00	0.01	0.02	0.04	0.02	0.08	0.14
100	50	35	0.00	0.00	0.00	0.00	0.02	0.05	0.04	0.12	0.19
100	100	35	0.00	0.00	0.00	0.01	0.03	0.06	0.13	0.25	0.35
100	100	50	0.00	0.01	0.01	0.01	0.04	0.08	0.14	0.29	0.41
200	100	50	0.00	0.00	0.01	0.01	0.04	0.07	0.23	0.42	0.54
200	200	50	0.00	0.01	0.01	0.02	0.05	0.09	0.54	0.71	0.80
200	200	100	0.00	0.01	0.02	0.01	0.05	0.09	0.86	0.95	0.97
500	500	100	0.00	0.02	0.03	0.02	0.07	0.11	1.00	1.00	1.00
500	500	200	0.00	0.02	0.05	0.02	0.06	0.11	1.00	1.00	1.00
500	500	300	0.00	0.02	0.04	0.01	0.06	0.10	1.00	1.00	1.00
800	800	500	0.01	0.03	0.06	0.02	0.06	0.12	1.00	1.00	1.00
1000	1000	600	0.00	0.03	0.05	0.01	0.05	0.10	1.00	1.00	1.00

E.6 Tables: Selection of number of factors

DESIGN 1 : $k^C = 1$, $k^H = k^L = 1$, $\beta = 0.0$, $a_F = 0.0$, $\phi = 0.0$, $a_e = 0.0$, $R^2_{max} = 0.8$

			AGGR(2016) HF data: PCA first			AGGR(2016) HF data: flow samp. first			CHEN (2012)			WANG (2012)		
			Average number of estimated factors over 4000 MC sim.											
N_H	N_L	T	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$
Estimated k_1 and k_2														
50	50	35	0.92	1.08	1.04	0.85	1.15	1.11	0.98	1.02	0.98	1.37	0.63	0.59
100	50	35	0.89	1.11	1.08	0.83	1.17	1.13	0.98	1.02	0.99	1.52	0.48	0.45
100	100	35	0.92	1.08	1.08	0.85	1.15	1.15	1.00	1.00	1.00	1.12	0.88	0.88
100	100	50	0.94	1.06	1.06	0.89	1.11	1.11	1.00	1.00	1.00	1.01	0.99	0.99
200	100	50	0.91	1.09	1.09	0.87	1.13	1.13	1.00	1.00	1.00	1.09	0.91	0.91
200	200	50	0.94	1.06	1.06	0.90	1.10	1.10	1.00	1.00	1.00	1.00	1.00	1.00
200	200	100	0.97	1.03	1.03	0.95	1.05	1.05	1.00	1.00	1.00	1.00	1.00	1.00
500	500	100	0.97	1.03	1.03	0.96	1.04	1.04	1.00	1.00	1.00	1.00	1.00	1.00
500	500	200	0.98	1.02	1.02	0.98	1.02	1.02	1.00	1.00	1.00	1.00	1.00	1.00
500	500	300	0.98	1.02	1.02	0.98	1.02	1.02	1.00	1.00	1.00	1.00	1.00	1.00
800	800	500	0.99	1.01	1.01	0.99	1.01	1.01	1.00	1.00	1.00	1.00	1.00	1.00
1000	1000	600	0.99	1.01	1.01	0.99	1.01	1.01	1.00	1.00	1.00	1.00	1.00	1.00
True k_1 and k_2														
50	50	35	0.95	1.05	1.05	0.88	1.12	1.12	1.00	1.00	1.00	1.41	0.59	0.59
100	50	35	0.91	1.09	1.09	0.85	1.15	1.15	1.00	1.00	1.00	1.55	0.45	0.45
100	100	35	0.92	1.08	1.08	0.85	1.15	1.15	1.00	1.00	1.00	1.12	0.88	0.88
100	100	50	0.94	1.06	1.06	0.89	1.11	1.11	1.00	1.00	1.00	1.01	0.99	0.99
200	100	50	0.91	1.09	1.09	0.87	1.13	1.13	1.00	1.00	1.00	1.09	0.91	0.91
200	200	50	0.94	1.06	1.06	0.90	1.10	1.10	1.00	1.00	1.00	1.00	1.00	1.00
200	200	100	0.97	1.03	1.03	0.95	1.05	1.05	1.00	1.00	1.00	1.00	1.00	1.00
500	500	100	0.97	1.03	1.03	0.96	1.04	1.04	1.00	1.00	1.00	1.00	1.00	1.00
500	500	200	0.98	1.02	1.02	0.98	1.02	1.02	1.00	1.00	1.00	1.00	1.00	1.00
500	500	300	0.98	1.02	1.02	0.98	1.02	1.02	1.00	1.00	1.00	1.00	1.00	1.00
800	800	500	0.99	1.01	1.01	0.99	1.01	1.01	1.00	1.00	1.00	1.00	1.00	1.00
1000	1000	600	0.99	1.01	1.01	0.99	1.01	1.01	1.00	1.00	1.00	1.00	1.00	1.00

The numbers of pervasive factors k_1 and k_2 in the first step of all the four estimation procedures considered are estimated using the IC_{p2} information criteria of Bai and Ng (2002), with $kmax = 10$. The same information criterion is used also in the second step of the Wang (2012) procedure for the selection of the number of pervasive factors R in the stacked panel of HF (flow sampled) and LF data; in this case $kmax = 20$.

DESIGN 2 : $k^C = 2$, $k^H = k^L = 0$, $\beta = 0.0$, $a_F = 0.0$, $\phi = 0.0$, $a_e = 0.0$, $R^2_{max} = 0.8$

			AGGR(2016) HF data: PCA first			AGGR(2016) HF data: flow samp. first			CHEN (2012)			WANG (2012)		
			Average number of estimated factors over 4000 MC sim.											
N_H	N_L	T	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$
Estimated k_1 and k_2														
50	50	35	1.87	0.13	0.10	1.73	0.27	0.24	1.97	0.03	0.00	1.97	0.03	-0.00
100	50	35	1.83	0.17	0.15	1.73	0.27	0.25	1.98	0.02	0.00	1.98	0.02	0.00
100	100	35	1.86	0.14	0.14	1.71	0.29	0.29	2.00	0.00	0.00	2.00	0.00	-0.00
100	100	50	1.91	0.09	0.09	1.82	0.18	0.18	2.00	0.00	0.00	2.00	0.00	0.00
200	100	50	1.88	0.12	0.12	1.81	0.19	0.19	2.00	0.00	0.00	2.00	0.00	0.00
200	200	50	1.91	0.09	0.09	1.82	0.18	0.18	2.00	0.00	0.00	2.00	0.00	0.00
200	200	100	1.96	0.04	0.04	1.93	0.07	0.07	2.00	0.00	0.00	2.00	0.00	0.00
500	500	100	1.97	0.03	0.03	1.93	0.07	0.07	2.00	0.00	0.00	2.00	0.00	0.00
500	500	200	1.98	0.02	0.02	1.97	0.03	0.03	2.00	0.00	0.00	2.00	0.00	0.00
500	500	300	1.99	0.01	0.01	1.98	0.02	0.02	2.00	0.00	0.00	2.00	0.00	0.00
800	800	500	1.99	0.01	0.01	1.99	0.01	0.01	2.00	0.00	0.00	2.00	0.00	0.00
1000	1000	600	1.99	0.01	0.01	1.99	0.01	0.01	2.00	0.00	0.00	2.00	0.00	0.00
True k_1 and k_2														
50	50	35	1.89	0.11	0.11	1.75	0.25	0.25	2.00	0.00	0.00	2.00	-0.00	-0.00
100	50	35	1.85	0.15	0.15	1.74	0.26	0.26	2.00	0.00	0.00	2.00	0.00	0.00
100	100	35	1.86	0.14	0.14	1.71	0.29	0.29	2.00	0.00	0.00	2.00	-0.00	-0.00
100	100	50	1.91	0.09	0.09	1.82	0.18	0.18	2.00	0.00	0.00	2.00	0.00	0.00
200	100	50	1.88	0.12	0.12	1.81	0.19	0.19	2.00	0.00	0.00	2.00	0.00	0.00
200	200	50	1.91	0.09	0.09	1.82	0.18	0.18	2.00	0.00	0.00	2.00	0.00	0.00
200	200	100	1.96	0.04	0.04	1.93	0.07	0.07	2.00	0.00	0.00	2.00	0.00	0.00
500	500	100	1.97	0.03	0.03	1.93	0.07	0.07	2.00	0.00	0.00	2.00	0.00	0.00
500	500	200	1.98	0.02	0.02	1.97	0.03	0.03	2.00	0.00	0.00	2.00	0.00	0.00
500	500	300	1.99	0.01	0.01	1.98	0.02	0.02	2.00	0.00	0.00	2.00	0.00	0.00
800	800	500	1.99	0.01	0.01	1.99	0.01	0.01	2.00	0.00	0.00	2.00	0.00	0.00
1000	1000	600	1.99	0.01	0.01	1.99	0.01	0.01	2.00	0.00	0.00	2.00	0.00	0.00

The numbers of pervasive factors k_1 and k_2 in the first step of all the four estimation procedures considered are estimated using the IC_{p2} information criteria of Bai and Ng (2002), with $kmax = 10$. The same information criterion is used also in the second step of the Wang (2012) procedure for the selection of the number of pervasive factors R in the stacked panel of HF (flow sampled) and LF data; in this case $kmax = 20$.

DESIGN 3 : $k^C = 2$, $k^H = k^L = 1$, $\beta = 0.0$, $a_F = 0.0$, $\phi = 0.0$, $a_e = 0.0$, $R^2_{max} = 0.8$

			AGGR(2016) HF data: PCA first			AGGR(2016) HF data: flow samp. first			CHEN (2012)			WANG (2012)		
			Average number of estimated factors over 4000 MC sim.											
N_H	N_L	T	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$
Estimated k_1 and k_2														
50	50	35	1.16	1.81	1.07	1.00	1.98	1.24	1.28	1.70	0.96	2.03	0.95	0.21
100	50	35	1.17	1.83	1.09	1.05	1.95	1.21	1.29	1.71	0.98	1.99	1.01	0.27
100	100	35	1.61	1.39	1.21	1.32	1.68	1.50	1.82	1.18	0.99	2.10	0.90	0.72
100	100	50	1.81	1.19	1.14	1.58	1.42	1.37	1.95	1.05	1.00	2.00	1.00	0.96
200	100	50	1.74	1.26	1.22	1.58	1.42	1.37	1.95	1.05	1.00	2.07	0.93	0.88
200	200	50	1.83	1.17	1.17	1.60	1.40	1.39	2.00	1.00	1.00	2.00	1.00	1.00
200	200	100	1.93	1.07	1.07	1.85	1.15	1.15	2.00	1.00	1.00	2.00	1.00	1.00
500	500	100	1.94	1.06	1.06	1.86	1.14	1.14	2.00	1.00	1.00	2.00	1.00	1.00
500	500	200	1.97	1.03	1.03	1.94	1.06	1.06	2.00	1.00	1.00	2.00	1.00	1.00
500	500	300	1.98	1.02	1.02	1.96	1.04	1.04	2.00	1.00	1.00	2.00	1.00	1.00
800	800	500	1.99	1.01	1.01	1.98	1.02	1.02	2.00	1.00	1.00	2.00	1.00	1.00
1000	1000	600	1.99	1.01	1.01	1.99	1.01	1.01	2.00	1.00	1.00	2.00	1.00	1.00
True k_1 and k_2														
50	50	35	1.81	1.19	1.19	1.49	1.51	1.51	1.99	1.01	1.01	2.83	0.17	0.17
100	50	35	1.68	1.32	1.32	1.46	1.54	1.54	1.99	1.01	1.01	2.73	0.27	0.27
100	100	35	1.74	1.26	1.26	1.41	1.59	1.59	2.00	1.00	1.00	2.28	0.72	0.72
100	100	50	1.85	1.15	1.15	1.61	1.39	1.39	2.00	1.00	1.00	2.04	0.96	0.96
200	100	50	1.76	1.24	1.24	1.60	1.40	1.40	2.00	1.00	1.00	2.12	0.88	0.88
200	200	50	1.83	1.17	1.17	1.61	1.39	1.39	2.00	1.00	1.00	2.00	1.00	1.00
200	200	100	1.93	1.07	1.07	1.85	1.15	1.15	2.00	1.00	1.00	2.00	1.00	1.00
500	500	100	1.94	1.06	1.06	1.86	1.14	1.14	2.00	1.00	1.00	2.00	1.00	1.00
500	500	200	1.97	1.03	1.03	1.94	1.06	1.06	2.00	1.00	1.00	2.00	1.00	1.00
500	500	300	1.98	1.02	1.02	1.96	1.04	1.04	2.00	1.00	1.00	2.00	1.00	1.00
800	800	500	1.99	1.01	1.01	1.98	1.02	1.02	2.00	1.00	1.00	2.00	1.00	1.00
1000	1000	600	1.99	1.01	1.01	1.99	1.01	1.01	2.00	1.00	1.00	2.00	1.00	1.00

The numbers of pervasive factors k_1 and k_2 in the first step of all the four estimation procedures considered are estimated using the IC_{p2} information criteria of Bai and Ng (2002), with $kmax = 10$. The same information criterion is used also in the second step of the Wang (2012) procedure for the selection of the number of pervasive factors R in the stacked panel of HF (flow sampled) and LF data; in this case $kmax = 20$.

DESIGN 4 : $k^C = 1$, $k^H = k^L = 1$, $\beta = 0.2$, $a_F = 0.0$, $\phi = 0.0$, $a_e = 0.0$, $R^2_{max} = 0.8$

			AGGR(2016) HF data: PCA first			AGGR(2016) HF data: flow samp. first			CHEN (2012)			WANG (2012)		
			Average number of estimated factors over 4000 MC sim.											
N_H	N_L	T	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$
Estimated k_1 and k_2														
50	50	35	0.92	1.08	1.04	0.85	1.15	1.12	0.98	1.02	0.98	1.37	0.63	0.60
100	50	35	0.90	1.10	1.07	0.85	1.15	1.12	0.98	1.02	0.99	1.52	0.48	0.45
100	100	35	0.92	1.08	1.08	0.84	1.16	1.16	1.00	1.00	1.00	1.12	0.88	0.88
100	100	50	0.94	1.06	1.06	0.89	1.11	1.11	1.00	1.00	1.00	1.01	0.99	0.99
200	100	50	0.92	1.08	1.08	0.88	1.12	1.12	1.00	1.00	1.00	1.09	0.91	0.91
200	200	50	0.94	1.06	1.06	0.90	1.10	1.10	1.00	1.00	1.00	1.00	1.00	1.00
200	200	100	0.96	1.04	1.04	0.94	1.06	1.06	1.00	1.00	1.00	1.00	1.00	1.00
500	500	100	0.97	1.03	1.03	0.95	1.05	1.05	1.00	1.00	1.00	1.00	1.00	1.00
500	500	200	0.98	1.02	1.02	0.97	1.03	1.03	1.00	1.00	1.00	1.00	1.00	1.00
500	500	300	0.99	1.01	1.01	0.98	1.02	1.02	1.00	1.00	1.00	1.00	1.00	1.00
800	800	500	0.99	1.01	1.01	0.99	1.01	1.01	1.00	1.00	1.00	1.00	1.00	1.00
1000	1000	600	0.99	1.01	1.01	0.99	1.01	1.01	1.00	1.00	1.00	1.00	1.00	1.00
True k_1 and k_2														
50	50	35	0.95	1.05	1.05	0.87	1.13	1.13	1.00	1.00	1.00	1.40	0.60	0.60
100	50	35	0.92	1.08	1.08	0.86	1.14	1.14	1.00	1.00	1.00	1.55	0.45	0.45
100	100	35	0.92	1.08	1.08	0.84	1.16	1.16	1.00	1.00	1.00	1.12	0.88	0.88
100	100	50	0.94	1.06	1.06	0.89	1.11	1.11	1.00	1.00	1.00	1.01	0.99	0.99
200	100	50	0.92	1.08	1.08	0.88	1.12	1.12	1.00	1.00	1.00	1.09	0.91	0.91
200	200	50	0.94	1.06	1.06	0.90	1.10	1.10	1.00	1.00	1.00	1.00	1.00	1.00
200	200	100	0.96	1.04	1.04	0.94	1.06	1.06	1.00	1.00	1.00	1.00	1.00	1.00
500	500	100	0.97	1.03	1.03	0.95	1.05	1.05	1.00	1.00	1.00	1.00	1.00	1.00
500	500	200	0.98	1.02	1.02	0.97	1.03	1.03	1.00	1.00	1.00	1.00	1.00	1.00
500	500	300	0.99	1.01	1.01	0.98	1.02	1.02	1.00	1.00	1.00	1.00	1.00	1.00
800	800	500	0.99	1.01	1.01	0.99	1.01	1.01	1.00	1.00	1.00	1.00	1.00	1.00
1000	1000	600	0.99	1.01	1.01	0.99	1.01	1.01	1.00	1.00	1.00	1.00	1.00	1.00

The numbers of pervasive factors k_1 and k_2 in the first step of all the four estimation procedures considered are estimated using the IC_{p2} information criteria of Bai and Ng (2002), with $kmax = 10$. The same information criterion is used also in the second step of the Wang (2012) procedure for the selection of the number of pervasive factors R in the stacked panel of HF (flow sampled) and LF data; in this case $kmax = 20$.

DESIGN 5 : $k^C = 1$, $k^H = k^L = 1$, $\beta = 0.0$, $a_F = 0.6$, $\phi = 0.0$, $a_e = 0.0$, $R^2_{max} = 0.8$

			AGGR(2016) HF data: PCA first			AGGR(2016) HF data: flow samp. first			CHEN (2012)			WANG (2012)		
			Average number of estimated factors over 4000 MC sim.											
N_H	N_L	T	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$
Estimated k_1 and k_2														
50	50	35	0.87	1.13	1.08	0.85	1.15	1.10	0.98	1.02	0.98	1.08	0.92	0.87
100	50	35	0.84	1.16	1.11	0.84	1.16	1.12	0.96	1.04	0.99	1.26	0.74	0.70
100	100	35	0.88	1.12	1.11	0.87	1.13	1.13	1.00	1.00	1.00	1.03	0.97	0.96
100	100	50	0.92	1.08	1.08	0.90	1.10	1.10	1.00	1.00	1.00	1.00	1.00	1.00
200	100	50	0.89	1.11	1.11	0.89	1.11	1.11	1.00	1.00	1.00	1.02	0.98	0.98
200	200	50	0.93	1.07	1.07	0.92	1.08	1.08	1.00	1.00	1.00	1.00	1.00	1.00
200	200	100	0.95	1.05	1.05	0.95	1.05	1.05	1.00	1.00	1.00	1.00	1.00	1.00
500	500	100	0.97	1.03	1.03	0.96	1.04	1.04	1.00	1.00	1.00	1.00	1.00	1.00
500	500	200	0.98	1.02	1.02	0.98	1.02	1.02	1.00	1.00	1.00	1.00	1.00	1.00
500	500	300	0.99	1.01	1.01	0.98	1.02	1.02	1.00	1.00	1.00	1.00	1.00	1.00
800	800	500	0.99	1.01	1.01	0.99	1.01	1.01	1.00	1.00	1.00	1.00	1.00	1.00
1000	1000	600	0.99	1.01	1.01	0.99	1.01	1.01	1.00	1.00	1.00	1.00	1.00	1.00
True k_1 and k_2														
50	50	35	0.90	1.10	1.10	0.88	1.12	1.12	1.00	1.00	1.00	1.13	0.87	0.87
100	50	35	0.87	1.13	1.13	0.86	1.14	1.14	1.00	1.00	1.00	1.30	0.70	0.70
100	100	35	0.89	1.11	1.11	0.87	1.13	1.13	1.00	1.00	1.00	1.04	0.96	0.96
100	100	50	0.92	1.08	1.08	0.90	1.10	1.10	1.00	1.00	1.00	1.00	1.00	1.00
200	100	50	0.89	1.11	1.11	0.89	1.11	1.11	1.00	1.00	1.00	1.02	0.98	0.98
200	200	50	0.93	1.07	1.07	0.92	1.08	1.08	1.00	1.00	1.00	1.00	1.00	1.00
200	200	100	0.95	1.05	1.05	0.95	1.05	1.05	1.00	1.00	1.00	1.00	1.00	1.00
500	500	100	0.97	1.03	1.03	0.96	1.04	1.04	1.00	1.00	1.00	1.00	1.00	1.00
500	500	200	0.98	1.02	1.02	0.98	1.02	1.02	1.00	1.00	1.00	1.00	1.00	1.00
500	500	300	0.99	1.01	1.01	0.98	1.02	1.02	1.00	1.00	1.00	1.00	1.00	1.00
800	800	500	0.99	1.01	1.01	0.99	1.01	1.01	1.00	1.00	1.00	1.00	1.00	1.00
1000	1000	600	0.99	1.01	1.01	0.99	1.01	1.01	1.00	1.00	1.00	1.00	1.00	1.00

The numbers of pervasive factors k_1 and k_2 in the first step of all the four estimation procedures considered are estimated using the IC_{p2} information criteria of Bai and Ng (2002), with $kmax = 10$. The same information criterion is used also in the second step of the Wang (2012) procedure for the selection of the number of pervasive factors R in the stacked panel of HF (flow sampled) and LF data; in this case $kmax = 20$.

DESIGN 6 : $k^C = 1$, $k^H = k^L = 1$, $\beta = 0.0$, $a_F = 0.0$, $\phi = 0.7$, $a_e = 0.0$, $R^2_{max} = 0.8$

			AGGR(2016) HF data: PCA first			AGGR(2016) HF data: flow samp. first			CHEN (2012)			WANG (2012)		
			Average number of estimated factors over 4000 MC sim.											
N_H	N_L	T	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$
Estimated k_1 and k_2														
50	50	35	0.90	1.10	1.07	0.83	1.17	1.13	1.02	0.98	0.95	1.96	0.04	-0.00
100	50	35	0.87	1.13	1.10	0.82	1.18	1.14	0.98	1.02	0.98	1.96	0.04	-0.00
100	100	35	0.92	1.08	1.08	0.84	1.16	1.16	1.04	0.96	0.96	2.00	0.00	0.00
100	100	50	0.94	1.06	1.06	0.89	1.11	1.11	1.00	1.00	1.00	1.99	0.01	0.01
200	100	50	0.92	1.08	1.08	0.89	1.11	1.11	1.00	1.00	1.00	2.00	0.00	0.00
200	200	50	0.95	1.05	1.05	0.90	1.10	1.10	1.00	1.00	1.00	1.99	0.01	0.01
200	200	100	0.97	1.03	1.03	0.94	1.06	1.06	1.00	1.00	1.00	1.67	0.33	0.33
500	500	100	0.97	1.03	1.03	0.95	1.05	1.05	1.00	1.00	1.00	1.34	0.66	0.66
500	500	200	0.98	1.02	1.02	0.97	1.03	1.03	1.00	1.00	1.00	1.00	1.00	1.00
500	500	300	0.99	1.01	1.01	0.98	1.02	1.02	1.00	1.00	1.00	1.00	1.00	1.00
800	800	500	0.99	1.01	1.01	0.99	1.01	1.01	1.00	1.00	1.00	1.00	1.00	1.00
1000	1000	600	0.99	1.01	1.01	0.99	1.01	1.01	1.00	1.00	1.00	1.00	1.00	1.00
True k_1 and k_2														
50	50	35	0.92	1.08	1.08	0.85	1.15	1.15	1.03	0.97	0.97	2.00	-0.00	-0.00
100	50	35	0.89	1.11	1.11	0.84	1.16	1.16	1.00	1.00	1.00	2.00	-0.00	-0.00
100	100	35	0.92	1.08	1.08	0.84	1.16	1.16	1.04	0.96	0.96	2.00	0.00	0.00
100	100	50	0.94	1.06	1.06	0.89	1.11	1.11	1.00	1.00	1.00	1.99	0.01	0.01
200	100	50	0.92	1.08	1.08	0.89	1.11	1.11	1.00	1.00	1.00	2.00	0.00	0.00
200	200	50	0.95	1.05	1.05	0.90	1.10	1.10	1.00	1.00	1.00	1.99	0.01	0.01
200	200	100	0.97	1.03	1.03	0.94	1.06	1.06	1.00	1.00	1.00	1.67	0.33	0.33
500	500	100	0.97	1.03	1.03	0.95	1.05	1.05	1.00	1.00	1.00	1.34	0.66	0.66
500	500	200	0.98	1.02	1.02	0.97	1.03	1.03	1.00	1.00	1.00	1.00	1.00	1.00
500	500	300	0.99	1.01	1.01	0.98	1.02	1.02	1.00	1.00	1.00	1.00	1.00	1.00
800	800	500	0.99	1.01	1.01	0.99	1.01	1.01	1.00	1.00	1.00	1.00	1.00	1.00
1000	1000	600	0.99	1.01	1.01	0.99	1.01	1.01	1.00	1.00	1.00	1.00	1.00	1.00

The numbers of pervasive factors k_1 and k_2 in the first step of all the four estimation procedures considered are estimated using the IC_{p2} information criteria of Bai and Ng (2002), with $kmax = 10$. The same information criterion is used also in the second step of the Wang (2012) procedure for the selection of the number of pervasive factors R in the stacked panel of HF (flow sampled) and LF data; in this case $kmax = 20$.

DESIGN 7 : $k^C = 1$, $k^H = k^L = 1$, $\beta = 0.0$, $a_F = 0.0$, $\phi = 0.95$, $a_e = 0.0$, $R^2_{max} = 0.8$

			AGGR(2016) HF data: PCA first			AGGR(2016) HF data: flow samp. first			CHEN (2012)			WANG (2012)		
			Average number of estimated factors over 4000 MC sim.											
N_H	N_L	T	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$
Estimated k_1 and k_2														
50	50	35	1.17	0.83	0.79	0.99	1.01	0.98	1.96	0.04	0.00	1.96	0.04	-0.00
100	50	35	1.00	1.00	0.96	0.92	1.08	1.04	1.89	0.11	0.07	1.96	0.04	0.00
100	100	35	0.94	1.06	1.06	0.85	1.15	1.14	2.00	0.00	0.00	2.00	0.00	0.00
100	100	50	0.95	1.05	1.05	0.89	1.11	1.11	2.00	0.00	0.00	2.00	0.00	0.00
200	100	50	0.92	1.08	1.08	0.88	1.12	1.12	1.97	0.03	0.03	2.00	0.00	0.00
200	200	50	0.94	1.06	1.06	0.90	1.10	1.10	2.00	0.00	0.00	2.00	0.00	0.00
200	200	100	0.97	1.03	1.03	0.95	1.05	1.05	2.00	0.00	0.00	2.00	0.00	0.00
500	500	100	0.97	1.03	1.03	0.95	1.05	1.05	2.00	0.00	0.00	2.00	0.00	0.00
500	500	200	0.98	1.02	1.02	0.97	1.03	1.03	1.38	0.62	0.62	2.00	0.00	0.00
500	500	300	0.99	1.01	1.01	0.98	1.02	1.02	1.00	1.00	1.00	2.00	0.00	0.00
800	800	500	0.99	1.01	1.01	0.99	1.01	1.01	1.00	1.00	1.00	2.00	0.00	0.00
1000	1000	600	0.99	1.01	1.01	0.99	1.01	1.01	1.00	1.00	1.00	2.00	0.00	0.00
True k_1 and k_2														
50	50	35	1.19	0.81	0.81	1.00	1.00	1.00	2.00	0.00	0.00	2.00	-0.00	-0.00
100	50	35	1.02	0.98	0.98	0.93	1.07	1.07	1.93	0.07	0.07	2.00	0.00	0.00
100	100	35	0.94	1.06	1.06	0.86	1.14	1.14	2.00	0.00	0.00	2.00	0.00	0.00
100	100	50	0.95	1.05	1.05	0.89	1.11	1.11	2.00	0.00	0.00	2.00	0.00	0.00
200	100	50	0.92	1.08	1.08	0.88	1.12	1.12	1.97	0.03	0.03	2.00	0.00	0.00
200	200	50	0.94	1.06	1.06	0.90	1.10	1.10	2.00	0.00	0.00	2.00	0.00	0.00
200	200	100	0.97	1.03	1.03	0.95	1.05	1.05	2.00	0.00	0.00	2.00	0.00	0.00
500	500	100	0.97	1.03	1.03	0.95	1.05	1.05	2.00	0.00	0.00	2.00	0.00	0.00
500	500	200	0.98	1.02	1.02	0.97	1.03	1.03	1.38	0.62	0.62	2.00	0.00	0.00
500	500	300	0.99	1.01	1.01	0.98	1.02	1.02	1.00	1.00	1.00	2.00	0.00	0.00
800	800	500	0.99	1.01	1.01	0.99	1.01	1.01	1.00	1.00	1.00	2.00	0.00	0.00
1000	1000	600	0.99	1.01	1.01	0.99	1.01	1.01	1.00	1.00	1.00	2.00	0.00	0.00

The numbers of pervasive factors k_1 and k_2 in the first step of all the four estimation procedures considered are estimated using the IC_{p2} information criteria of Bai and Ng (2002), with $kmax = 10$. The same information criterion is used also in the second step of the Wang (2012) procedure for the selection of the number of pervasive factors R in the stacked panel of HF (flow sampled) and LF data; in this case $kmax = 20$.

DESIGN 8 : $k^C = 1$, $k^H = k^L = 5$, $\beta = 0.0$, $a_F = 0.0$, $\phi = 0.0$, $a_e = 0.0$, $R^2_{max} = 0.8$

			AGGR(2016) HF data: PCA first			AGGR(2016) HF data: flow samp. first			CHEN (2012)			WANG (2012)		
			Average number of estimated factors over 4000 MC sim.											
N_H	N_L	T	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$
Estimated k_1 and k_2														
50	50	35	1.79	3.07	8.11	1.35	3.51	8.55	2.57	2.29	7.33	12.97	-8.12	-3.07
100	50	35	1.33	4.67	8.55	1.13	4.87	8.75	2.23	3.77	7.65	14.36	-8.36	-4.48
100	100	35	0.93	5.06	3.22	0.82	5.17	3.33	1.96	4.04	2.20	9.06	-3.07	-4.91
100	100	50	0.96	5.03	4.39	0.90	5.09	4.45	1.51	4.48	3.85	9.78	-3.79	-4.42
200	100	50	0.95	5.05	4.41	0.91	5.09	4.45	1.04	4.96	4.32	9.55	-3.55	-4.20
200	200	50	0.95	5.05	4.00	0.90	5.10	4.05	1.45	4.55	3.50	9.79	-3.79	-4.85
200	200	100	0.98	5.02	5.02	0.96	5.04	5.04	1.00	5.00	5.00	6.28	-0.28	-0.28
500	500	100	0.99	5.01	5.01	0.97	5.03	5.03	1.00	5.00	5.00	7.32	-1.32	-1.32
500	500	200	0.99	5.01	5.01	0.98	5.02	5.02	1.00	5.00	5.00	1.01	4.99	4.99
500	500	300	0.99	5.01	5.01	0.99	5.01	5.01	1.00	5.00	5.00	1.00	5.00	5.00
800	800	500	0.99	5.01	5.01	0.99	5.01	5.01	1.00	5.00	5.00	1.00	5.00	5.00
1000	1000	600	1.00	5.00	5.00	0.99	5.01	5.01	1.00	5.00	5.00	1.00	5.00	5.00
True k_1 and k_2														
50	50	35	1.66	4.34	4.34	1.17	4.83	4.83	2.14	3.86	3.86	10.22	-4.22	-4.22
100	50	35	1.07	4.93	4.93	0.96	5.04	5.04	1.47	4.53	4.53	10.49	-4.49	-4.49
100	100	35	0.98	5.02	5.02	0.90	5.10	5.10	2.66	3.34	3.34	10.91	-4.91	-4.91
100	100	50	0.98	5.02	5.02	0.92	5.08	5.08	1.72	4.28	4.28	10.43	-4.43	-4.43
200	100	50	0.96	5.04	5.04	0.93	5.07	5.07	1.08	4.92	4.92	10.20	-4.20	-4.20
200	200	50	0.98	5.02	5.02	0.94	5.06	5.06	1.77	4.23	4.23	10.85	-4.85	-4.85
200	200	100	0.98	5.02	5.02	0.96	5.04	5.04	1.00	5.00	5.00	6.28	-0.28	-0.28
500	500	100	0.99	5.01	5.01	0.97	5.03	5.03	1.00	5.00	5.00	7.32	-1.32	-1.32
500	500	200	0.99	5.01	5.01	0.98	5.02	5.02	1.00	5.00	5.00	1.01	4.99	4.99
500	500	300	0.99	5.01	5.01	0.99	5.01	5.01	1.00	5.00	5.00	1.00	5.00	5.00
800	800	500	0.99	5.01	5.01	0.99	5.01	5.01	1.00	5.00	5.00	1.00	5.00	5.00
1000	1000	600	1.00	5.00	5.00	0.99	5.01	5.01	1.00	5.00	5.00	1.00	5.00	5.00

The numbers of pervasive factors k_1 and k_2 in the first step of all the four estimation procedures considered are estimated using the IC_{p2} information criteria of Bai and Ng (2002), with $kmax = 10$. The same information criterion is used also in the second step of the Wang (2012) procedure for the selection of the number of pervasive factors R in the stacked panel of HF (flow sampled) and LF data; in this case $kmax = 20$.

DESIGN 9 : $k^C = 1$, $k^H = k^L = 5$, $\beta = 0.0$, $a_F = 0.0$, $\phi = 0.5$, $a_e = 0.0$, $R^2_{max} = 0.8$

			AGGR(2016) HF data: PCA first			AGGR(2016) HF data: flow samp. first			CHEN (2012)			WANG (2012)		
			Average number of estimated factors over 4000 MC sim.											
N_H	N_L	T	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$
Estimated k_1 and k_2														
50	50	35	2.32	2.55	7.58	1.68	3.19	8.23	3.06	1.81	6.85	12.26	-7.40	-2.36
100	50	35	1.65	4.34	8.22	1.30	4.70	8.58	2.74	3.26	7.14	13.82	-7.82	-3.94
100	100	35	0.95	5.04	3.21	0.85	5.14	3.31	2.90	3.10	1.26	8.55	-2.55	-4.38
100	100	50	0.97	5.02	4.34	0.91	5.08	4.40	2.68	3.31	2.63	8.46	-2.46	-3.14
200	100	50	0.95	5.05	4.41	0.91	5.09	4.45	1.56	4.44	3.80	8.51	-2.51	-3.15
200	200	50	0.95	5.05	3.97	0.90	5.10	4.03	2.95	3.05	1.97	8.68	-2.68	-3.76
200	200	100	0.99	5.01	5.01	0.96	5.04	5.04	1.21	4.79	4.79	6.20	-0.20	-0.20
500	500	100	0.99	5.01	5.01	0.97	5.03	5.03	1.17	4.83	4.83	6.25	-0.25	-0.25
500	500	200	0.99	5.01	5.01	0.99	5.01	5.01	1.00	5.00	5.00	4.91	1.09	1.09
500	500	300	0.99	5.01	5.01	0.98	5.02	5.02	1.00	5.00	5.00	1.38	4.62	4.62
800	800	500	0.99	5.01	5.01	0.99	5.01	5.01	1.00	5.00	5.00	1.00	5.00	5.00
1000	1000	600	1.00	5.00	5.00	0.99	5.01	5.01	1.00	5.00	5.00	1.00	5.00	5.00
True k_1 and k_2														
50	50	35	1.99	4.01	4.01	1.31	4.69	4.69	2.70	3.30	3.30	9.49	-3.49	-3.49
100	50	35	1.20	4.80	4.80	1.00	5.00	5.00	2.01	3.99	3.99	9.94	-3.94	-3.94
100	100	35	1.00	5.00	5.00	0.92	5.08	5.08	3.65	2.35	2.35	10.38	-4.38	-4.38
100	100	50	0.98	5.02	5.02	0.93	5.07	5.07	2.92	3.08	3.08	9.14	-3.14	-3.14
200	100	50	0.97	5.03	5.03	0.93	5.07	5.07	1.68	4.32	4.32	9.15	-3.15	-3.15
200	200	50	0.98	5.02	5.02	0.94	5.06	5.06	3.39	2.61	2.61	9.76	-3.76	-3.76
200	200	100	0.99	5.01	5.01	0.96	5.04	5.04	1.21	4.79	4.79	6.20	-0.20	-0.20
500	500	100	0.99	5.01	5.01	0.97	5.03	5.03	1.17	4.83	4.83	6.25	-0.25	-0.25
500	500	200	0.99	5.01	5.01	0.99	5.01	5.01	1.00	5.00	5.00	4.91	1.09	1.09
500	500	300	0.99	5.01	5.01	0.98	5.02	5.02	1.00	5.00	5.00	1.38	4.62	4.62
800	800	500	0.99	5.01	5.01	0.99	5.01	5.01	1.00	5.00	5.00	1.00	5.00	5.00
1000	1000	600	1.00	5.00	5.00	0.99	5.01	5.01	1.00	5.00	5.00	1.00	5.00	5.00

The numbers of pervasive factors k_1 and k_2 in the first step of all the four estimation procedures considered are estimated using the IC_{p3} information criteria of Bai and Ng (2002), with $kmax = 10$. The same information criterion is used also in the second step of the Wang (2012) procedure for the selection of the number of pervasive factors R in the stacked panel of HF (flow sampled) and LF data; in this case $kmax = 20$.

DESIGN 10 : $k^C = 1$, $k^H = k^L = 5$, $\beta = 0.0$, $a_F = 0.0$, $\phi = 0.7$, $a_e = 0.0$, $R^2_{max} = 0.8$

			AGGR(2016) HF data: PCA first			AGGR(2016) HF data: flow samp. first			CHEN (2012)			WANG (2012)		
			Average number of estimated factors over 4000 MC sim.											
N_H	N_L	T	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$
Estimated k_1 and k_2														
50	50	35	3.00	1.88	6.89	2.09	2.79	7.80	3.59	1.29	6.31	11.62	-6.73	-1.72
100	50	35	2.22	3.78	7.66	1.66	4.34	8.22	3.41	2.58	6.46	13.21	-7.21	-3.33
100	100	35	1.02	4.98	3.15	0.89	5.10	3.27	3.70	2.29	0.46	7.93	-1.94	-3.77
100	100	50	1.00	5.00	4.37	0.94	5.06	4.43	4.12	1.87	1.24	7.63	-1.63	-2.26
200	100	50	0.96	5.04	4.39	0.93	5.07	4.43	2.82	3.18	2.53	7.72	-1.72	-2.37
200	200	50	0.97	5.03	3.95	0.92	5.08	4.00	4.50	1.50	0.42	7.64	-1.64	-2.72
200	200	100	0.98	5.02	5.01	0.97	5.03	5.03	3.22	2.78	2.78	6.04	-0.04	-0.04
500	500	100	0.99	5.01	5.01	0.97	5.03	5.03	3.50	2.50	2.50	6.03	-0.03	-0.03
500	500	200	0.99	5.01	5.01	0.98	5.02	5.02	1.01	4.99	4.99	6.00	0.00	0.00
500	500	300	0.99	5.01	5.01	0.98	5.02	5.02	1.00	5.00	5.00	5.91	0.09	0.09
800	800	500	0.99	5.01	5.01	0.99	5.01	5.01	1.00	5.00	5.00	1.80	4.20	4.20
1000	1000	600	1.00	5.00	5.00	0.99	5.01	5.01	1.00	5.00	5.00	1.01	4.99	4.99
True k_1 and k_2														
50	50	35	2.52	3.48	3.48	1.57	4.43	4.43	3.34	2.66	2.66	8.84	-2.84	-2.84
100	50	35	1.55	4.45	4.45	1.21	4.79	4.79	2.70	3.30	3.30	9.33	-3.33	-3.33
100	100	35	1.13	4.87	4.87	0.97	5.03	5.03	4.75	1.25	1.25	9.77	-3.77	-3.77
100	100	50	1.01	4.99	4.99	0.95	5.05	5.05	4.42	1.58	1.58	8.27	-2.27	-2.27
200	100	50	0.97	5.03	5.03	0.94	5.06	5.06	2.98	3.02	3.02	8.37	-2.37	-2.37
200	200	50	0.98	5.02	5.02	0.95	5.05	5.05	5.26	0.74	0.74	8.72	-2.72	-2.72
200	200	100	0.98	5.02	5.02	0.97	5.03	5.03	3.22	2.78	2.78	6.04	-0.04	-0.04
500	500	100	0.99	5.01	5.01	0.97	5.03	5.03	3.50	2.50	2.50	6.03	-0.03	-0.03
500	500	200	0.99	5.01	5.01	0.98	5.02	5.02	1.01	4.99	4.99	6.00	0.00	0.00
500	500	300	0.99	5.01	5.01	0.98	5.02	5.02	1.00	5.00	5.00	5.91	0.09	0.09
800	800	500	0.99	5.01	5.01	0.99	5.01	5.01	1.00	5.00	5.00	1.80	4.20	4.20
1000	1000	600	1.00	5.00	5.00	0.99	5.01	5.01	1.00	5.00	5.00	1.01	4.99	4.99

The numbers of pervasive factors k_1 and k_2 in the first step of all the four estimation procedures considered are estimated using the IC_{p3} information criteria of Bai and Ng (2002), with $kmax = 10$. The same information criterion is used also in the second step of the Wang (2012) procedure for the selection of the number of pervasive factors R in the stacked panel of HF (flow sampled) and LF data; in this case $kmax = 20$.

DESIGN 11 : $k^C = 1$, $k^H = k^L = 5$, $\beta = 0.0$, $a_F = 0.0$, $\phi = 0.95$, $a_e = 0.0$, $R^2_{max} = 0.8$

			AGGR(2016) HF data: PCA first			AGGR(2016) HF data: flow samp. first			CHEN (2012)			WANG (2012)		
			Average number of estimated factors over 4000 MC sim.											
N_H	N_L	T	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$	$\hat{k}^C$	$\hat{k}^H$	$\hat{k}^L$
Estimated k_1 and k_2														
50	50	35	4.34	0.53	5.54	3.39	1.48	6.50	4.48	0.40	5.41	10.73	-5.86	-0.85
100	50	35	4.19	1.81	5.68	3.15	2.85	6.72	4.94	1.06	4.93	12.34	-6.35	-2.47
100	100	35	2.49	3.51	1.70	1.74	4.26	2.45	4.18	1.81	0.01	6.94	-0.94	-2.75
100	100	50	2.73	3.27	2.63	1.93	4.07	3.43	5.34	0.66	0.01	6.57	-0.57	-1.22
200	100	50	2.10	3.90	3.23	1.72	4.28	3.61	5.31	0.69	0.02	6.70	-0.70	-1.36
200	200	50	1.64	4.36	3.29	1.29	4.71	3.64	4.93	1.07	0.00	6.42	-0.42	-1.49
200	200	100	1.19	4.81	4.81	1.07	4.93	4.93	6.00	0.00	0.00	6.00	-0.00	-0.00
500	500	100	0.99	5.01	5.01	0.98	5.02	5.02	6.00	0.00	0.00	6.00	-0.00	-0.00
500	500	200	0.99	5.01	5.01	0.98	5.02	5.02	6.00	0.00	0.00	6.00	0.00	0.00
500	500	300	0.99	5.01	5.01	0.99	5.01	5.01	6.00	0.00	0.00	6.00	0.00	0.00
800	800	500	0.99	5.01	5.01	0.99	5.01	5.01	6.00	0.00	0.00	6.00	0.00	0.00
1000	1000	600	1.00	5.00	5.00	0.99	5.01	5.01	6.00	0.00	0.00	6.00	0.00	0.00
True k_1 and k_2														
50	50	35	4.09	1.91	1.91	2.73	3.27	3.27	4.53	1.47	1.47	7.98	-1.98	-1.98
100	50	35	3.33	2.67	2.67	2.49	3.51	3.51	4.30	1.70	1.70	8.47	-2.47	-2.47
100	100	35	2.96	3.04	3.04	2.05	3.95	3.95	5.90	0.10	0.10	8.75	-2.75	-2.75
100	100	50	2.86	3.14	3.14	2.03	3.97	3.97	5.96	0.04	0.04	7.22	-1.22	-1.22
200	100	50	2.17	3.83	3.83	1.77	4.23	4.23	5.87	0.13	0.13	7.36	-1.36	-1.36
200	200	50	1.76	4.24	4.24	1.36	4.64	4.64	6.00	0.00	0.00	7.49	-1.49	-1.49
200	200	100	1.19	4.81	4.81	1.07	4.93	4.93	6.00	0.00	0.00	6.00	-0.00	-0.00
500	500	100	0.99	5.01	5.01	0.98	5.02	5.02	6.00	0.00	0.00	6.00	-0.00	-0.00
500	500	200	0.99	5.01	5.01	0.98	5.02	5.02	6.00	0.00	0.00	6.00	0.00	0.00
500	500	300	0.99	5.01	5.01	0.99	5.01	5.01	6.00	0.00	0.00	6.00	0.00	0.00
800	800	500	0.99	5.01	5.01	0.99	5.01	5.01	6.00	0.00	0.00	6.00	0.00	0.00
1000	1000	600	1.00	5.00	5.00	0.99	5.01	5.01	6.00	0.00	0.00	6.00	0.00	0.00

The numbers of pervasive factors k_1 and k_2 in the first step of all the four estimation procedures considered are estimated using the IC_{p3} information criteria of Bai and Ng (2002), with $kmax = 10$. The same information criterion is used also in the second step of the Wang (2012) procedure for the selection of the number of pervasive factors R in the stacked panel of HF (flow sampled) and LF data; in this case $kmax = 20$.

E.7 Cross-sectional distribution of R^2 and adjusted R^2 for correctly specified and misspecified number of common factors

In the empirical analysis of Section 7 we use adjusted R^2 as a measure to compare the fraction of variability of the original data explained by the different estimated factors. One referee raised the important issue of whether two highly correlated specific factors (as the ones considered in some of the previous simulation designs) maybe better interpreted as a common factor when the sample sizes are as small as in our empirical application, especially when adjusted R^2 is used. In fact, it is possible that interpreting two highly correlated specific factors as one common factor will increase the adjusted adjusted R^2 even more due to the efficiency gain from a larger combined sample.

We address this concern in a MC experiment where data are generated from a DGP with $k^C = 1$ common factor, and 1 HF-specific and 1 LF-specific factors ($k^H = k^L = 1$) which are highly correlated. We consider two values of the correlation coefficient among the specific factors, namely $\phi = 0.7$, and 0.9. Then, we estimate on the simulated data both a correctly specified model with $k^C = k^H = k^L = 1$, and a misspecified model with $k^C = 2$ and $k^H = k^L = 0$. In both cases we compute the quantiles of the CS distributions for both the R^2, and the adjusted R^2, of the regressions of the observed data on the different estimated factors, as in Table I. That is, for the correctly specified model we regress the simulated LF and HF data on (i) the common factor only, (ii) the specific factor only, and (iii) both common and specific factors. For the misspecified model the regressors include the two common factors only. Table E.5 ($\phi = 0.7$) and Table E.6 ($\phi = 0.9$) report the sample average computed over 2000 MC simulations for each of the 10%, 25%, 50%, 75%, and 90% quantiles of both the R^2s, and the adjusted R^2s. The results clearly show that the factors estimated from a correctly specified model, when both the common and the specific factors are included in the regressions, produce both R^2 and adjusted R^2 for the LF data which are consistently higher than those of a misspecified model. As expected, the cross-sectional distribution of the regressions of the HF data on one common and one HF-specific factors produce exactly the same R^2 and adjusted R^2 as the regressions on 2 common factors. This happens because the common and the specific factors are estimated from a rotation of the same $k_1 = 2$ pervasive factors estimated by PCA in the first step of our procedure, which are also used to estimate the two common factors in the misspecified model. The results are qualitatively the same for both values of the correlation coefficients among the specific factors $\phi = 0.7$, and 0.9. They suggest that even for prediction purposes distinguishing two highly correlated specific factors from a single common factor is valuable.

Table E.5: Sample averages over 2000 MC simulations of the quantiles of R^2 and adjusted R^2 of regressions on true and estimated factors, with $R^2_{all} = 0.8$, $N_H = 100$, $N_L = 50$, $T = 35$, $M = 4$, $a_F = 0.0$, $\beta = 0$, $a_e = 0.0$, $\phi = 0.7$.

Panel A: Estimated with $k^C = k^H = k^L = 1$, as in DGP

	R^2: Quantile				
Factors	10%	25%	50%	75%	90%
Observables: LF variables					
common	0.6	3.0	12.2	31.2	52.7
common, LF-spec.	7.0	17.2	36.9	60.5	77. 0
LF-spec.	0.6	2.9	11.9	30.9	53. 0
Observables: HF variables					
common	0.4	2.5	10.7	29.5	52.4
common, HF-spec.	5.8	15.0	34.3	58.5	76.2
HF-spec.	0.4	2.5	10.6	29.0	51.7

Panel B: Estimated with $k^C = 2, k^H = k^L = 0$

	R^2: Quantile				
Factors	10%	25%	50%	75%	90%
Observables: LF variables					
common	5.1	12.5	26.7	44.7	61.1
common, LF-spec.	-	-	-	-	-
LF-spec.	-	-	-	-	-
Observables: HF variables					
common	5.8	15.0	34.3	58.5	76.2
common, HF-spec.	-	-	-	-	-
HF-spec.	-	-	-	-	-

Panel C: Estimated with $k^C = k^H = k^L = 1$, as in DGP

	$\bar{R}^2$: Quantile				
Factors	10%	25%	50%	75%	90%
Observables: LF variables					
common	-2.4	0.1	9.6	29.2	51.2
common, LF-spec.	1.2	12.0	33.0	58.0	75.6
LF-spec.	-2.4	0	9.2	28.8	51.6
Observables: HF variables					
common	-0.3	1.8	10.1	29.0	52. 0
common, HF-spec.	4.4	13.8	33.3	57.9	75.9
HF-spec.	-0.3	1.8	10.0	28.5	51.4

Panel D: Estimated with $k^C = 2, k^H = k^L = 0$

	$\bar{R}^2$: Quantile				
Factors	10%	25%	50%	75%	90%
Observables: LF variables					
common	-0.9	7.1	22.2	41.3	58.7
common, LF-spec.	-	-	-	-	-
LF-spec.	-	-	-	-	-
Observables: HF variables					
common	4.4	13.8	33.3	57.9	75.9
common, HF-spec.	-	-	-	-	-
HF-spec.	-	-	-	-	-

In each line we report the sample averages, computed over 2000 MC simulations, of the quantiles of R^2 (Panels A and B) and adj. R^2 (Panels C and D) of regressions on estimated factors. In all panels, the regressions in the first three lines involve the growth rates of the 50 LF observables as dependent variables, while those in the last three lines involve the growth rates of the 100 HF observables as dependent variables. In Panels A and C the explanatory variables are the factors estimated assuming that $k^C = k^H = k^L = 1$, as in the DGP. In Panels B and D the explanatory variables are the factors estimated assuming that $k^C = 2$ and $k^H = k^L = 0$, differently from the true number of factors in the DGP. The low-frequency sample size T is set equal to 35, and the number of high-frequency subperiods is $M = 4$. All low frequency observables are flow-sampled. For all the DGPs we set $a_F = 0.0$, $\beta = 0$, $a_e = 0.0$, $\phi = 0.7$, and $R^2_{all} = 0.8$.

Table E.6: Sample averages over 2000 MC simulations of the quantiles of R^2 and adjusted R^2 of regressions on true and estimated factors, with $R^2_{all} = 0.8$, $N_H = 100$, $N_L = 50$, $T = 35$, $M = 4$, $a_F = 0.0$, $\beta = 0$, $a_e = 0.0$, $\phi = 0.9$.

Panel A: Estimated with $k^C = k^H = k^L = 1$, as in DGP

	R^2: Quantile				
Factors	10%	25%	50%	75%	90%
Observables: LF variables					
common	0.6	3.1	12.6	32.0	53.1
common, LF-spec.	6.8	17.2	36.8	60.2	76.9
LF-spec.	0.5	2.8	11.5	30.2	52.2
Observables: HF variables					
common	0.4	2.5	10.8	29.6	52.4
common, HF-spec.	5.6	14.9	34.1	58.6	76.2
HF-spec.	0.4	2.4	10.6	28.9	51.5

Panel B: Estimated with $k^C = 2, k^H = k^L = 0$

	R^2: Quantile				
Factors	10%	25%	50%	75%	90%
Observables: LF variables					
common	6.1	15.2	32.7	53.6	69.0
common, LF-spec.	-	-	-	-	-
LF-spec.	-	-	-	-	-
Observables: HF variables					
common	5.6	14.9	34.1	58.6	76.2
common, HF-spec.	-	-	-	-	-
HF-spec.	-	-	-	-	-

Panel C: Estimated with $k^C = k^H = k^L = 1$, as in DGP

	$\bar{R}^2$: Quantile				
Factors	10%	25%	50%	75%	90%
Observables: LF variables					
common	-2.4	0.2	10.0	29.9	51.7
common, LF-spec.	1.0	12.0	32.9	57.8	75.4
LF-spec.	-2.5	-0.2	8.9	28.1	50.8
Observables: HF variables					
common	-0.3	1.8	10.1	29.1	52.1
common, HF-spec.	4.3	13.6	33.1	58.0	75.9
HF-spec.	-0.3	1.7	10.0	28.4	51.1

Panel D: Estimated with $k^C = 2, k^H = k^L = 0$

	$\bar{R}^2$: Quantile				
Factors	10%	25%	50%	75%	90%
Observables: LF variables					
common	0.2	9.9	28.5	50.7	67.1
common, LF-spec.	-	-	-	-	-
LF-spec.	-	-	-	-	-
Observables: HF variables					
common	4.3	13.6	33.1	58.0	75.9
common, HF-spec.	-	-	-	-	-
HF-spec.	-	-	-	-	-

In each line we report the sample averages, computed over 2000 MC simulations, of the quantiles of R^2 (Panels A and B) and adj. R^2 (Panels C and D) of regressions on estimated factors. In all panels, the regressions in the first three lines involve the growth rates of the 50 LF observables as dependent variables, while those in the last three lines involve the growth rates of the 100 HF observables as dependent variables. In Panels A and C the explanatory variables are the factors estimated assuming that $k^C = k^H = k^L = 1$, as in the DGP. In Panels B and D the explanatory variables are the factors estimated assuming that $k^C = 2$ and $k^H = k^L = 0$, differently from the true number of factors in the DGP. The low-frequency sample size T is set equal to 35, and the number of high-frequency subperiods is $M = 4$. All low frequency observables are flow-sampled. For all the DGPs we set $a_F = 0.0$, $\beta = 0$, $a_e = 0.0$, $\phi = 0.9$, and $R^2_{all} = 0.8$.

References

AHN, S. C., AND A. R. HORENSTEIN (2013): "Eigenvalue Ratio Test for the Number of Factors," *Econometrica*, 81, 1203–1227.

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

ANDREOU, E., P. GAGLIARDINI, E. GHYSELS, AND M. RUBIN (2019): "Inference in Group Factor Models with an Application to Mixed Frequency Data," *Econometrica, forthcoming*.

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

——— (2009): "Panel Data Models with Interactive Fixed Effects," *Econometrica*, 77(4), 1229–1279.

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

BATES, B., M. PLAGBORG-MOLLER, J. STOCK, AND M. WATSON (2013): "Consistent Factor Estimation in Dynamic Factor Models with Structural Instability," *Journal of Econometrics*, 177, 289–304.

BILLINGSLEY, P. (1995): *Probability and Measure*. Wiley.

CHEN, P. (2012): "Common Factors and Specific Factors," *Unpublished Manuscript, MPRA Paper No. 36114*.

DAVIDSON, J. (1994): *Stochastic Limit Theory*. Oxford University Press.

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, 1–38.

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.

HARVEY, A. (1989): *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press, Cambridge.

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

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

ONATSKI, A. (2010): "Determining the Number of Factors from Empirical Distribution of Eigenvalues," *Review of Economics and Statistics*, 92, 1004–1016.

STOCK, J. H., AND M. W. WATSON (2002): "Forecasting Using Principal Components from a Large Number of Predictors," *Journal of the American Statistical Association*, 97, 1167 – 1179.

WANG, P. (2012): "Large Dimensional Factor Models with a Multi-Level Factor Structure: Identification, Estimation, and Inference," *Unpublished Manuscript, Hong Kong University of Science and Technology*.