Inferential Theory for Generalized Dynamic Factor Models

Matteo Barigozzi
Department of Economics, Università di Bologna, Italy

Marc Hallin
ECARES and Département de Mathématique
Université libre de Bruxelles, Belgium

Matteo Luciani
FFederal Reserve Board of Governors, Washington, DC, USA

Paolo ZaffaroniImperial College Business School, London, UK

August 2021

ECARES working paper 2021-20

ECARES
ULB - CP 114/04
50, F.D. Roosevelt Ave., B-1050 Brussels BELGIUM
www.ecares.org

Inferential Theory for Generalized Dynamic Factor Models[•]

Matteo Barigozzi[†] Marc Hallin[‡] Matteo Luciani[*] Paolo Zaffaroni[**]

August 29, 2021

Abstract

We provide the asymptotic distributional theory for the so-called *General* or *Generalized Dynamic Factor Model* (GDFM), laying the foundations for an inferential approach in the GDFM analysis of high-dimensional time series. Our results are exploiting the duality between common shocks and dynamic loadings under a random cross-section approach to derive the asymptotic distribution of a class of estimators for common shocks, dynamic loadings, common components, and impulse response functions. An empirical application aimed at the construction of a "core" inflation indicator for the U.S. economy is presented, empirically demonstrating the superiority of the GDFM-based indicator over the most commonly adopted approaches, outperforming, in particular, the one based on Principal Components.

Keywords: High-dimensional time series, Generalized Dynamic Factor Models, One-sided representations of dynamic factor models, Asymptotic distribution, Confidence intervals.

JEL subject classification : C0, C01, E0.

1 Introduction

This paper provides the asymptotic distribution theory, and hence the inferential method, for the estimator recently proposed in Forni et al. (2015) (hereafter, FHLZ) for the so-called *General* or *Generalized Dynamic Factor Model* (GDFM) introduced by Forni et al. (2000). Our approach combines the flexibility of the GDFM in terms of dynamics with the possibility, bestowed by the *Dynamic Factor Models* (DFMs) of Stock and Watson (2002a,b) and Bai and Ng (2002), of estimating the common shocks and their impulse response functions (IRFs).

Under the GDFM, the statistical analysis of a countable family $\{x_{it}|\, t \in \mathbb{Z}\}$, $i \in \mathbb{N}$ of observable stochastic processes is based on the decomposition of x_{it} into

$$x_{it} = \chi_{it} + \xi_{it} = b_{i1}(L)u_{1t} + b_{i2}(L)u_{2t} + \ldots + b_{iq}(L)u_{qt} + \xi_{it}, \quad i \in \mathbb{N},\ t \in \mathbb{Z}, \tag{1}$$

where $\mathbf{u}_t = (u_{1t}\ u_{2t}\ \ldots\ u_{qt})'$ is an unobservable q-dimensional vector of mutually orthogonal *common shocks* driving $\{\chi_{it}|\, t \in \mathbb{Z},\, i \in \mathbb{N}\}$ and $b_{if}(L), i \in \mathbb{N},\ f = 1, \ldots, q$, are square-summable filters (L, as

† Department of Economics, Università di Bologna, Italy matteo.barigozzi@unibo.it
‡ ECARES, Université libre de Bruxelles, Belgium mhallin@ulb.ac.be
[*] Federal Reserve Board of Governors, Washington, DC, USA matteo.luciani@frb.gov
[**] Imperial College Business School, London, UK p.zaffaroni@imperial.ac.uk

[•]Disclaimer: the views expressed in this paper are those of the authors and do not necessarily reflect the views and policies of the Board of Governors or the Federal Reserve System.

usual, stands for the lag operator). The unobservable χ_{it} and ξ_{it}, for which identifying assumptions are provided in Section 2, are called x_{it}'s *common* and *idiosyncratic components*, respectively; at minimum, it is assumed that the idiosyncratic components ξ_{it} are "weakly" cross-correlated (in a sense to be made precise) and orthogonal at any lead and lag to the common shocks $u_{1t}, \dots, u_{qt}$ driving the common components χ_{it}.

The literature on DFMs is based on the assumption that the space spanned by the common components $\{\chi_{it} | i \in \mathbb{N}\}$ is, for any given t, *finite-dimensional*—with dimension r, say, independent of t. Under that assumption, the decomposition (1) can be rewritten as

$$\begin{aligned} x_{it} &= \lambda_{i1}F_{1t} + \lambda_{i2}F_{2t} + \ldots + \lambda_{ir}F_{rt} + \xi_{it}, \\ \mathbf{F}_t &= (F_{1t} \ \ldots \ F_{rt})' = \mathbf{N}(L)\mathbf{u}_t, \end{aligned} \tag{2}$$

since the vector $\mathbf{F}_t$ of r *static factors* F_{jt} is loaded contemporaneously via scalar *loadings* λ_{ij}, call this a *static representation.* The factors $\mathbf{F}_t$ and the loadings λ_{ij} can be estimated consistently (after imposing adequate identification constraints) using the first r *standard* principal components (see Stock and Watson, 2002a,b and Bai, 2003; see also Fan et al., 2013, 2015, 2016, 2017, 2021 where, in a finance context, several refinements of the PCA approach are proposed). Inference based on these estimators can be carried out thanks to Bai (2003), who establishes the asymptotic normality of the PCA estimators of the DFM (2), together with consistent estimation of the corresponding asymptotic covariance matrices. Bai and Ng (2002), among many others, also propose criteria to determine r consistently. Moreover, the second equation in (2) is usually specified as a Singular Vector Autogression (SVAR), so that (2) takes the form

$$\begin{aligned} x_{it} &= \lambda_{i1}F_{1t} + \lambda_{i2}F_{2t} + \ldots + \lambda_{ir}F_{rt} \ + \xi_{it} \\ \mathbf{D}(L)\mathbf{F}_t &= (\mathbf{I} - \mathbf{D}_1 L - \mathbf{D}_2 L^2 - \ldots - \mathbf{D}_p L^p)\mathbf{F}_t \ = \mathbf{K}\mathbf{u}_t, \end{aligned} \tag{3}$$

where the matrices $\mathbf{D}_j$ are $r \times r$ while $\mathbf{K}$ is $r \times q$. Under (3), Amengual and Watson (2007) and Bai and Ng (2007) provide consistent criteria to jointly determine q and r.

As already mentioned, an appealing consequence of the DFM decomposition (2) is that it readily permits to derive the IRFs of the common shocks from the estimation of $\boldsymbol{\lambda}_i = (\lambda_{i1}, \dots, \lambda_{ir})'$, $\mathbf{D}(L)$, and $\mathbf{K}$. Identification of the matrix $\mathbf{K}$ by suitable restrictions allows for an interpretation of the shocks $\mathbf{u_t}$ as the *structural* common shocks while the simplicity of the finite-dimensional nature of (3) enhances its use for out-of-sample forecasting (Forni et al., 2018).

The same finite-dimensional nature of (3), however, rules out a number of quite plausible dynamic structure such as simple AR(1) models for the observables x_{it}.[1] Recognizing that the space spanned by the common components of (1), in many applications, is likely to be infinite-dimensional ($r = \infty$), hence cannot be recovered from a finite number of standard principal components, Forni et al. (2000) use q principal components in the frequency domain (the *dynamic principal components* introduced by Brillinger (2001)) to estimate the common components χ_{it}, where q can be obtained, for instance, from the identification methods proposed by Hallin and Liška (2007) or Onatski (2009). However, being based on dynamic principal components, their estimators involve two-sided filters acting on the observations x_{it}, hence do not allow to estimate the common shocks at the end of the observation period, nor their IRFs: their methods thus are unsuitable for out-of-sample prediction.

[1] Simple cases like $x_{it} = a_i(1 - \alpha_i L)^{-1}u_t + \xi_{it}$, that is $q = 1$, with white noise u_t and AR coefficients α_i admitting an absolutely continuous distribution, are ruled out.

FHLZ remedy this problem and bring together the virtues of the (infinite-dimensional) GDFM (1) and the simplicity of the (finite-dimensional) DFM (3). Under the mild assumption of *rationality* of the spectral density of the common components χ_{it}—that is, assuming that each filter $b_{if}(L)$ in (1) is a ratio of *finite-degree* polynomials in L[2]—and elaborating upon results by Anderson and Deistler (2008a,b), FHLZ prove that for generic values of the parameters in $b_{if}(L)$ (i.e. apart from a lower-dimensional subset in the parameter space) $\boldsymbol{\chi}_{nt} := (\chi_{1t}, \chi_{2t}, \dots, \chi_{nt})'$ in (1) admits a *unique autoregressive representation with block structure* of the form

$$\mathbf{A}_n(L)\boldsymbol{\chi}_{nt} = \mathbf{R}_n\mathbf{u}_{\mathrm{t}}, \tag{4}$$

where $\mathbf{A}_n(L)$ is a $n \times n$ block-diagonal matrix polynomial in L of bounded degree (as $n \to \infty$), consisting of $(q+1) \times (q+1)$-dimensional blocks, and $\mathbf{R}_n$ a $n \times q$ matrix of rank q.

Since $\boldsymbol{\chi}_{nt} = \mathbf{x}_{nt} - \boldsymbol{\xi}_{nt}$ with $\mathbf{x}_{nt} := (x_{1t}, x_{2t}, \dots, x_{nt})'$ and $\boldsymbol{\xi}_{nt} := (\xi_{1t}, \xi_{2t}, \dots, \xi_{nt})'$, the filtered process $\mathbf{z}_{nt} := \mathbf{A}_n(L)\mathbf{x}_{nt}$ satisfies

$$\mathbf{z}_{nt} = \mathbf{R}_n\mathbf{u}_t + \boldsymbol{\phi}_{nt} \tag{5}$$

with $\boldsymbol{\phi}_{nt} := \mathbf{A}_n(L)\boldsymbol{\xi}_{nt}$. Expression (5) is key because it shows how to represent the GDFM (1) with infinite-dimensional factor space and observations $\mathbf{x}_{nt}$ as a DFM (3) with finite-dimensional factor space and observations $\mathbf{Z}_{nt}$. Indeed, under our assumptions, it can be shown that $\boldsymbol{\phi}_{nt}$ is idiosyncratic, so that (5) is a representation of the form (3) with $\mathbf{D}_n(L) = \mathbf{I}_n$, $\mathbf{F}_t = \mathbf{u}_{\mathrm{t}}$, $\boldsymbol{\lambda}_i = \mathbf{R}_i$, the ith row of $\mathbf{R}_n$, and thus $r = q$. As explained below, a crucial advantage of (4) is that the high-dimensional VAR operator $\mathbf{A}_n(L)$, thanks to its block-diagonality, is piecing together a set of k (with $1 \leq k \leq \lfloor n/(q+1) \rfloor$) $(q+1)$-dimensional VARs $\mathbf{A}^{(\mathrm{k})}(\mathrm{L})$. Our procedure does not run into curse of dimensionality problems because q is finite, and typically small.

The FHLZ estimation of the GDFM decomposition (1) mainly consists of three steps which can be summarized as follows.[3] First, by means of Hallin and Liška (2007) and dynamic PCA, estimate q and the spectral density matrix of $\boldsymbol{\chi}_{nt}$. Second, by Fourier inversion, derive the corresponding autocovariance matrices and the Yule-Walker estimators of the $(q+1) \times (q+1)$ blocks of $\mathbf{A}_n(L)$ in (4). This yields an estimated $\mathbf{z}_{nt}$, hence, up to estimation errors, allowing us to switch from the dynamic to the static representation (5) of the GDFM. Third, exploiting the finite-dimensional nature of (5), apply static PCA to the estimated $\mathbf{z}_{nt}$.

Challenges arise, however, when one needs to carry out inference. Indeed, one cannot rely on the asymptotic results of Bai (2003) because, unlike $\mathbf{x}_{nt}$ in (3), $\mathbf{z}_{nt}$ in (5) is unobserved and one can only consider its sample counterpart, which depends on estimators of the filter $\mathbf{A}_n(L)$. This makes the existing limiting theory for estimators of DFM (3) invalid.

The objective of this paper is to fill that theoretical gap and provide, for the FHLZ estimation of the GDFM (1) with infinite-dimensional factor space, the analogous type of results as Bai (2003) does for the PCA estimator of the loadings and factors in (3).

In Sections 2 and 3, we formalize the general assumptions needed for the GDFM setup and its static representation (5), and the reinforcements of these assumptions required for our distributional results. The FHLZ estimation procedure is described in Section 4 and its limiting statistical properties

[2]Namely, $b_{if}(L) = c_{if}(L)/d_{if}(L)$ where $c_{if}(L)$ and $d_{if}(L)$ are finite-degree polynomials in L.

[3]Strictly speaking, a consistent reconstruction of χ_{it}, ξ_{it} or the shocks $\mathbf{u}_t$ is not an *estimation* since χ_{it}, ξ_{it}, and $\mathbf{u}_t$ are not parameters, but we consistenly will indulge in this convenient abuse of terminology.

are established in Section 5. Section 6 presents the Monte Carlo experiments, and Section 7 illustrates the empirical application. Section 8 concludes. Technical proofs are relegated to the final Appendix.

2 GDFM: General Representation

We assume throughout that all stochastic variables in this paper belong to the Hilbert space $L_2(\Omega, \mathcal{F}, P)$ where $(\Omega, \mathcal{F}, P)$ is some common probability space. We study double-indexed zero-mean stochastic processes of the form $\mathbf{x} := \{x_{it} |\, i \in \mathbb{N},\, t \in \mathbb{Z}\}$, of which we only observe a finite $n \times T$ realization $\{x_{it} |\, 1 \le i \le n,\, 1 \le t \le T\}$. Denoting by $\mathbf{x}_n$ the n-dimensional subprocess $\{x_{it} |\, i = 1, \dots, n,\, t \in \mathbb{Z}\}$, the lag-$k$ autocovariance matrix of $\mathbf{x}_{nt}$ is defined as $\mathbf{\Gamma}_{n,k} := \mathbb{C}\mathrm{ov}(\mathbf{x}_{nt}, \mathbf{x}_{n,t-k})$, with $\mathbf{\Gamma}_n := \mathbf{\Gamma}_{n,0} = \mathbb{V}\mathrm{ar}(\mathbf{x}_{nt})$ for simplicity. On $\mathbf{x}$, the basic GDFM assumptions are as follows.

Assumption (A0) (GDFM)**.** *(a) The process $\mathbf{x}$ is second-order stationary with respect to time;*

(b) for all $n \in \mathbb{N}$, $\mathbf{x}_n$ admits the spectral density matrix

$$\mathbf{\Sigma}_n(\theta) := \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \mathbf{\Gamma}_{n,k}\, e^{-\iota k\theta}, \quad \theta \in [-\pi, \pi]$$

where $\iota = \sqrt{-1}$, with jth largest eigenvalue $\lambda_{nj}(\theta)$, $j = 1, \dots, n$;

(c) the number $q := \min\{j : \limsup_{n\to\infty} \sup_{\theta\in[-\pi,\pi]} \lambda_{nj}(\theta) < \infty\} - 1$ of diverging eigenvalues of $\mathbf{\Sigma}_n(\theta)$ is finite.

Under Assumption (A0), it can be shown (see, e.g. Forni and Lippi, 2001; Hallin and Lippi, 2013) that each element of $\mathbf{x}$ decomposes into the sum

$$x_{it} = \chi_{it} + \xi_{it}, \quad i \in \mathbb{N},\ t \in \mathbb{Z}, \tag{6}$$

of an unobserved *common component* χ_{it} and an unobserved *idiosyncratic component* ξ_{it} where, denoting by $\mathbf{\Gamma}^{\chi}_{n,k}$ and $\mathbf{\Gamma}^{\xi}_{n,k}$ the lag-k autocovariance matrices of the n-dimensional subprocesses $\boldsymbol{\chi}_n$ and $\boldsymbol{\xi}_n$ of $\boldsymbol{\chi} := \{\chi_{it} |\, i \in \mathbb{N},\, t \in \mathbb{Z}\}$ and $\boldsymbol{\xi} := \{\xi_{it} |\, i \in \mathbb{N},\, t \in \mathbb{Z}\}$, the spectral density matrices $\mathbf{\Sigma}^{\chi}_n(\theta)$ and $\mathbf{\Sigma}^{\xi}_n(\theta)$ exist, with eigenvalues $\lambda^{\chi}_{nj}(\theta)$ and $\lambda^{\xi}_{nj}(\theta)$ satisfying

$$\lim_{n\to\infty} \lambda^{\chi}_{nq}(\theta) = \infty \quad \text{and} \quad \limsup_{n\to\infty} \lambda^{\xi}_{n1}(\theta) < \infty \quad \theta\text{-a.e in } [-\pi, \pi]. \tag{7}$$

The common component process $\boldsymbol{\chi}$, moreover, is driven by a q-dimensional second-order white noise of *common shocks*, that is,

$$\chi_{it} = \sum_{j=1}^{q} \sum_{k=0}^{\infty} b_{ij,k} u_{j,t-k}, \quad i \in \mathbb{N},\ t \in \mathbb{Z}, \tag{8}$$

for some square-summable loadings $b_{ij,k}$, $i \in \mathbb{N}$, $1 \le j \le q$ and some q-dimensional second-order white noise $\{\mathbf{u}_t = (u_{1t}, \dots, u_{qt}) | t \in \mathbb{Z}\}$, which implies that $\lambda^{\chi}_{n,q+1}(\theta) = 0$ θ-a.e. in $[-\pi, \pi]$. In vector notation, (6) takes the form

$$\mathbf{x}_{nt} = \boldsymbol{\chi}_{nt} + \boldsymbol{\xi}_{nt}, \quad n \in \mathbb{N},\ t \in \mathbb{Z}, \tag{9}$$

with obvious definitions of the common and idiosyncratic subprocesses $\boldsymbol{\chi}_{nt}$ and $\boldsymbol{\xi}_{nt}$.

If distributional results are to be obtained, however, these properties need to be reinforced; on top of Assumption (A0) and its consequences, we assume the following.

Assumption (A1)(GDFM+)**.** *(a) The process* $\mathbf{x}$ *is* stricly *stationary with respect to time;*

(b) The q *diverging eigenvalues* $\lambda^{\chi}_{nj}(\theta)$ *diverge at rate* n *and are well separated, that is, for all* $j = 1, \ldots, q$*, there exist two strictly positive continuous functions* $\theta \mapsto \alpha^{\chi}_{j}(\theta)$ *and* $\theta \mapsto \beta^{\chi}_{j}(\theta)$ *from* $[-\pi, \pi]$ *to* $\mathbb{R}$ *such that, for all* $\theta \in [-\pi, \pi]$*,*

$$\alpha^{\chi}_{j}(\theta) \leq \liminf_{n\to\infty} \frac{\lambda^{\chi}_{nj}(\theta)}{n} \leq \limsup_{n\to\infty} \frac{\lambda^{\chi}_{nj}(\theta)}{n} \leq \beta^{\chi}_{j}(\theta)$$

with $\beta^{\chi}_{j}(\theta) < \alpha^{\chi}_{j-1}(\theta)$ *for all* $j = 2, \ldots, q$*;*

(c) the common shocks $\{\mathbf{u}_t := (u_{1t} \cdots u_{qt})' | \, t \in \mathbb{Z}\}$ *are* q*-dimensional i.i.d. white noise, with positive definite covariance* $\mathbb{E}[\mathbf{u}_t\mathbf{u}_t'] = \mathbf{\Gamma}^u$*; moreover, for all* $j = 1, \ldots, q$ *and all* $t \in \mathbb{Z}$*,* $\mathbb{E}\left[|u_{jt}|^p\right] \leq M_u$ *for some* $p > 5$ *and some finite constant* $M_u > 0$*;*

(d) the idiosyncratic components are such that, for all $i \in \mathbb{N}$ *and* $t \in \mathbb{Z}$*,*

$$\xi_{it} = \sum_{j=1}^{\infty}\sum_{k=0}^{\infty} \beta_{ij,k}\eta_{j,t-k}, \tag{10}$$

where $\{\boldsymbol{\eta}_t := (\eta_{1t} \; \eta_{2t} \cdots)' | \, t \in \mathbb{Z}\}$ *is an infinite-dimensional i.i.d. zero-mean stochastic process; moreover, for all* $i \in \mathbb{N}$*,* $\mathbb{E}[\eta_{it}\eta_{jt}] = 0$ *for* $i \neq j$*,* $\mathbb{E}[\eta_{it}^2] = 1$*, and* $\mathbb{E}[|\eta_{it}|^p] \leq M_\eta$ *for some* $p > 5$ *and some finite constant* $M_\eta > 0$*;*

(e) $\{\mathbf{u}_t\}$ *and* $\{\boldsymbol{\eta}_t\}$ *are mutually independent at all leads and lags;*

(f) for all $i \in \mathbb{N}$*,* $j = 1, \ldots, q$*, and* $z \in \mathbb{C}$*,* $b_{ij}(z) := \sum_{k=0}^{\infty} b_{ij,k}z^k$ *is of the form* $c_{ij}(z)/d_{ij}(z)$ *where*

  (i) $c_{ij}(z) = \sum_{k=0}^{s_1} c_{ij,k}z^k$ *for some positive integer* s_1*, with* $|c_{ij,k}| \leq B^{\chi}$ *for some real* $B^{\chi} > 0$ *independent of* i *and* j*;*

  (ii) $d_{ij}(z) = \sum_{k=0}^{s_2} d_{ij,k}z^k$ *for some positive integer* s_2*, and is such that all the roots of* $d_{ij}(z) = 0$ *satisfy* $|z| \geq \phi > 1$ *for some* $\phi > 0$ *independent of* i *and* j*;*

(g) for all $i, j \in \mathbb{N}$ *and* $k \in \mathbb{Z}^+$*,* $|\beta_{ij,k}| \leq B_{ij}\rho^k$*, with* $\rho \in [0, 1)$*,* $\sum_{i=1}^{\infty} B_{ij} \leq B$*, and* $\sum_{j=1}^{\infty} B_{ij} \leq B$ *for some finite real* $B > 0$ *independent of* i *and* j*.*

Parts *(a)–(e)* are reinforcing the traditional GDFM Assumption (A0) by requiring, among others, stationarity rather than second-order stationarity, and mutually independent strong white noises $\mathbf{u}_t$ and $\boldsymbol{\eta}_t$ rather than mutually orthogonal second-order white noises. Finite fifth moments and i.i.d.-ness of the common and idiosyncratic innovations are needed in order to control the degree of *physical dependence* (Wu, 2005) of the common and idiosyncratic components, hence of each x_{it}. This is what allows us to consistently estimate the spectral density matrix (Wu and Zaffaroni, 2018; Zhang and Wu, 2021; see also Proposition 5 in Forni et al., 2017). Part *(b)* is a classical reinforcement (Forni et al., 2000) of the pervasiveness of common shocks assuming linearly diverging and well-separated common eigenvalues, which avoids the uninteresting difficulties related with asymptotically multiple eigenvalues. Linear divergence rates, moreover, are the only ones compatible with the fact that cross-sectional ordering is completely arbitrary, hence should remain irrelevant—see the stochastic approach in Section 3 for further justification.

The idiosyncratic MA(∞) representation (10) in part *(d)* along with part *(g)* also entail square-summability of the idiosyncratic filters, both along the time and the cross-sectional dimensions. This, in turn, implies limited (lagged) cross-sectional dependence among idiosyncratic components. Indeed, letting $\sigma^{\xi}_{ij}(\theta)$ denote the (i,j)th entry of the idiosyncratic spectral density matrix $\mathbf{\Sigma}^{\xi}_n(\theta)$, it filliws from parts *(d)* and *(g)* of Assumption (A1) that

$$\begin{aligned}\sup_{\theta\in[-\pi,\pi]}\sup_{j\in\mathbb{N}}\sum_{i=1}^{\infty}|\sigma^{\xi}_{ij}(\theta)| &\leq \sup_{\theta\in[-\pi,\pi]}\sup_{j\in\mathbb{N}}\sum_{i=1}^{\infty}\sum_{s=0}^{\infty}|\beta_{is}(e^{-\iota\theta})\beta_{js}(e^{\iota\theta})|\\ &\leq \sup_{\theta\in[-\pi,\pi]}\sup_{j\in\mathbb{N}}\frac{1}{(1-\rho)^2}\sum_{i=1}^{\infty}\sum_{s=0}^{\infty}B_{is}B_{js}\leq \frac{B^2}{(1-\rho)^2}=B^{\xi},\text{ say.}\end{aligned}\tag{11}$$

This immediately implies (see Forni et al., 2017, Proposition 1)

$$\sup_{\theta\in[-\pi,\pi]}\sup_{n\in\mathbb{N}}\lambda^{\xi}_{n1}(\theta)\leq B^{\xi},\tag{12}$$

which is in line with the second part of (7), which *(d)* and *(g)*, thus, are reinforcing. Part *(f)* entails rational filters and, therefore, rational spectral density matrices $\mathbf{\Sigma}^{\chi}_n(\theta)$, as well as square-summability of the common filters along the time dimension. Furthermore, a simple application of Weyl's inequality, allows us to show that part (b) of Assumption (A1), together with (12), imply

$$\alpha^{\chi}_j(\theta)\leq \liminf_{n\to\infty}\frac{\lambda_{nj}(\theta)}{n}\leq \limsup_{n\to\infty}\frac{\lambda_{nj}(\theta)}{n}\leq \beta^{\chi}_j(\theta),\tag{13}$$

and

$$\sup_{n\in\mathbb{N}}\sup_{\theta\in[-\pi,\pi]}\lambda_{n,q+1}(\theta)\leq B^x$$

for some positive real B^x (see Forni et al., 2017, Proposition 1).

3 GDFM: VAR representation

3.1 VAR representation and a duality issue

Let Assumptions (A0) and (A1) hold. For any $s\in\mathbb{N}$ and $t\in\mathbb{Z}$, consider the $(q+1)$-dimensional subvector $\boldsymbol{\chi}^{(s)}_t := (\chi_{(s-1)(q+1)+1,t}\cdots\chi_{s(q+1),t})'$ of $\boldsymbol{\chi}_t$. Forni et al. (2015) prove that the following property is satisfied for generic values of the parameters $c_{ij,k}$ and $d_{ij,k}$ in Assumption 2(f). Turning this property into an assumption, thus, only places an extremely mild restriction on the actual data-generating process.

Assumption (**A2**)(VAR representation)**.** *For all $s\in\mathbb{N}$ and all $t\in\mathbb{Z}$, there exist a unique $(q+1)$-dimensional VAR filter $\mathbf{A}^{(s)}(L)=\mathbf{I}_{q+1}-\sum_{k=1}^{p_s}\mathbf{A}^{(s)}_kL^k$ and a $(q+1)\times q$-dimensional matrix $\boldsymbol{\mathcal{R}}^{(s)}$ such that*

$$\mathbf{A}^{(s)}(L)\boldsymbol{\chi}^{(s)}_t=\boldsymbol{\mathcal{R}}^{(s)}\mathbf{u}_t,\qquad t\in\mathbb{Z}\tag{14}$$

where

(a) $p_s\leq S:=qs_1+q^2s_2<\infty$ and all the roots of the determinant equation $\det\big(\mathbf{A}^{(s)}(z)\big)=0$, $z\in\mathbb{C}$, are such that $|z|>1$;

(b) $\boldsymbol{\mathcal{R}}^{(s)}$ has maximal rank q.

Moreover, denoting by $\mathbf{C}^{\chi}_s$ the $S(q+1)\times S(q+1)$ covariance matrix of $(\boldsymbol{\chi}^{(s)\prime}_t\cdots\boldsymbol{\chi}^{(s)\prime}_{t-S})'$, for all $s\in\mathbb{N}$,

(c) $|\det(\mathbf{C}^{\chi}_s)|\leq d$ for some finite positive real d.

Denote by $\underline{\mathbf{A}}(L)$ the infinite-dimensional block-diagonal matrix with diagonal blocks $\mathbf{A}^{(s)}(L)$, $s \in \mathbb{N}$ and define $\underline{\boldsymbol{\mathcal{R}}} := (\boldsymbol{\mathcal{R}}^{(1)\prime}\, \boldsymbol{\mathcal{R}}^{(2)\prime}\, \cdots)'$ with $(q+1)$ rows and infinitely many columns. Considering, without loss of generality, n such that $n = m(q+1)$ for some integer m, let

$$\mathbf{A}_n(L) := \begin{pmatrix} \mathbf{A}^{(1)}(L) & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}^{(2)}(L) & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{A}^{(m)}(L) \end{pmatrix}, \quad \boldsymbol{\mathcal{R}}_n := \begin{pmatrix} \boldsymbol{\mathcal{R}}^{(1)} \\ \boldsymbol{\mathcal{R}}^{(2)} \\ \vdots \\ \boldsymbol{\mathcal{R}}^{(m)} \end{pmatrix}, \tag{15}$$

which are the upper $n \times n$ and upper $n \times q$ sub-matrices of $\underline{\mathbf{A}}(L)$ and $\underline{\boldsymbol{\mathcal{R}}}'$, respectively. Then, from Assumption (A2), the common component $\boldsymbol{\chi}_{nt}$ admits the finite AR representation

$$\mathbf{A}_n(L)\boldsymbol{\chi}_{nt} = \boldsymbol{\mathcal{R}}_n \mathbf{u}_t, \quad n \in \mathbb{N},\ t \in \mathbb{Z}, \tag{16}$$

so that, with $\mathbf{B}_n(L) = (\mathbf{b}_1(L) \cdots \mathbf{b}_n(L))' := [\mathbf{A}_n(L)]^{-1}$,

$$\boldsymbol{\chi}_{nt} = \mathbf{B}_n(L)\mathbf{u}_t = [\mathbf{A}_n(L)]^{-1}\boldsymbol{\mathcal{R}}_n \mathbf{u}_t, \quad n \in \mathbb{N},\ t \in \mathbb{Z}. \tag{17}$$

Now, letting $\boldsymbol{\mathcal{R}}_n := (\mathbf{R}_1 \cdots \mathbf{R}_i \cdots \mathbf{R}_n)'$, in view of (9), we obtain

$$\mathbf{z}_{nt} := \mathbf{A}_n(L)\mathbf{x}_{nt} = \boldsymbol{\mathcal{R}}_n \mathbf{u}_t + \mathbf{A}_n(L)\boldsymbol{\xi}_{nt} =: \boldsymbol{\psi}_{nt} + \boldsymbol{\phi}_{nt}, \quad n \in \mathbb{N},\ t \in \mathbb{Z}. \tag{18}$$

Let us show that (18) is a static factor model for $\mathbf{z}_{nt}$ in the sense of Bai (2003). For any given n and $T \in \mathbb{N}$, consider

$$\mathbf{X}_{nT} := \begin{pmatrix} x_{11} & \cdots & x_{i1} & \cdots & x_{n1} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ x_{1t} & \cdots & x_{it} & \cdots & x_{nt} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ x_{1T} & \cdots & x_{iT} & \cdots & x_{nT} \end{pmatrix} = \begin{pmatrix} \mathbf{x}'_{n1} \\ \vdots \\ \mathbf{x}'_{nt} \\ \vdots \\ \mathbf{x}'_{nT} \end{pmatrix} = \left(\boldsymbol{x}_T^1 \cdots \boldsymbol{x}_T^i \cdots \boldsymbol{x}_T^n\right), \tag{19}$$

with tth row $\mathbf{x}_{nt} = (x_{1t} \cdots x_{nt})'$ (an n-dimensional vector) and ith column $\boldsymbol{x}_T^i = (x_{i1} \cdots x_{iT})'$ (a T-dimensional column vector), respectively, and the $T \times q$ matrix of common shocks $\boldsymbol{\mathcal{U}}_T := (\mathbf{u}_1 \cdots \mathbf{u}_t \cdots \mathbf{u}_T)'$, with tth row $\mathbf{u}'_t$ (a q-dimensional vector). Recall that the $n \times q$ matrix $\boldsymbol{\mathcal{R}}_n := (\mathbf{R}_1 \cdots \mathbf{R}_i \cdots \mathbf{R}_n)'$ has q-dimensional rows $\mathbf{R}'_i$. Similarly define the idiosyncratic $T \times n$ matrix

$$\boldsymbol{\Phi}_{nT} = (\boldsymbol{\phi}_{n1} \cdots \boldsymbol{\phi}_{nt} \cdots \boldsymbol{\phi}_{nT})' = \left(\boldsymbol{\varphi}_T^1 \cdots \boldsymbol{\varphi}_T^i \cdots \boldsymbol{\varphi}_T^n\right),$$

with tth row $\boldsymbol{\phi}'_{nt} = (\phi_{1t} \cdots \phi_{nt})$ (an n-dimensional vector) and ith column $\boldsymbol{\varphi}_T^i = (\phi_{i1} \cdots \phi_{iT})'$ (a T-dimensionalvector). With this notation, the static representation (18) of the GDFM, henceforth the static model, takes the form of a matrix representation

$$\mathbf{Z}_{nT} := (\mathbf{A}_n(L)\mathbf{X}'_{nT})' = \boldsymbol{\mathcal{U}}_T \boldsymbol{\mathcal{R}}'_n + \boldsymbol{\Phi}_{nT} =: \boldsymbol{\Psi}_{nT} + \boldsymbol{\Phi}_{nT}, \tag{20}$$

where $\mathbf{Z}_{nT}$ is $T \times n$, with rows $\mathbf{z}'_{nt}$ and columns $\boldsymbol{z}_T^i$, and $L\boldsymbol{X}'_{nT} = L(\mathbf{x}_{n1} \cdots \mathbf{x}_{nT}) = (\mathbf{x}_{n0} \cdots \mathbf{x}_{n,T-1})$.

The same matrix representation (20) can be written under (transposed) row-vector form (cross-sectional projection), which, denoting by $\mathbf{e}_{Tt}$ the tth column of the $T \times T$ identity matrix $\mathbf{I}_T$, yields

$$\text{(I)} \qquad \mathbf{z}_{nt} = \left(\mathbf{A}_n(L)\mathbf{X}'_{nT}\right)\mathbf{e}_{Tt} = \boldsymbol{\mathcal{R}}_n\mathbf{u}_t + \boldsymbol{\phi}_{nt}, \qquad n \in \mathbb{N},\ t \in \mathbb{Z} \tag{21}$$

(that is, the static factor model (18)) or, denoting by $\mathbf{e}_{ni}$ the ith column of the $n \times n$ identity matrix $\mathbf{I}_n$, under the column-vector form (temporal projection)

$$\text{(II)} \qquad \boldsymbol{z}^i_T = (\mathbf{A}_n(L)\mathbf{X}'_{nT})'\mathbf{e}_{ni} = \boldsymbol{\mathcal{U}}_T\mathbf{R}_i + \boldsymbol{\varphi}^i_T, \qquad i \in \mathbb{N},\ T \in \mathbb{N}. \tag{22}$$

These two forms are kind of dual static factor model representations, with the time- and cross-sectional-dimensions changing roles—that is, $\mathbf{X}'_{nT}$ replacing $\mathbf{X}_{nT}$. In (21), time-indexed q-dimensional random vectors $\mathbf{u}_t$ are deterministically loaded at time t while, in (22), cross-sectionally indexed q-dimensional deterministic vectors $\mathbf{R}_i$ are randomly loaded by cross-sectional item i; the role of the matrix of common components is played by $\boldsymbol{\Psi}_{nT}$ in (21), by $\boldsymbol{\Psi}'_{nT}$ in (22); the role of the matrix of idiosyncratic components is played by $\boldsymbol{\Phi}_{nT}$ in (21), by $\boldsymbol{\Phi}'_{nT}$ in (22).

As we shall see, both representations have their advantages, and both will be used in the sequel. An essential difference remains, however: $\mathbf{u}_t$ in (21) is random, while $\mathbf{R}_i$ in (22) so far is deterministic. As a consequence, while, under second-order stationarity (Assumption (A0a)), the law of large numbers implies that, for any fixed $n \in \mathbb{N}$, $\frac{1}{T}\boldsymbol{\Phi}'_{nT}\boldsymbol{\Phi}_{nT}$ converges in probability to some $n\times n$ matrix $\boldsymbol{\Gamma}^{\phi}_n$ as $T \to \infty$, no such property holds for $\frac{1}{n}\boldsymbol{\Phi}_{nT}\boldsymbol{\Phi}'_{nT}$ as $n \to \infty$. The duality between (21) and (22), thus, is imperfect or incomplete.

3.2 A stochastic cross-section approach

This imperfect duality issue is easily palliated if a stochastically generated cross-section scheme is adopted. Under that approach, it is assumed that the stochastic process $\mathbf{x}$ is generated via a two-step random mechanism: (A) the stochastic selection, via some unspecified distribution $\mathbb{P}$, of the distributional features[4] of $\mathbf{x}$ as a time-indexed stochastic process, followed by (B) a realization over time of the selected process $\mathbf{x}$, of which a finite $n \times T$ realization is observed, and along which time-series asymptotics will be considered as $T \to \infty$.

The distribution $\mathbb{P}$ in step (A) of that mechanism plays the role of a nuisance. Statistical practice in such cases consists in conducting inference on the realization observed in step (B) conditional on the (unobserved) result of step (A): see, e.g., Lehmann and Romano (2006, Chapter 10), so that $\mathbb{P}$ needs no further description. Under such conditional approach, the distributional features of the stochastic process $\mathbf{x}$ of which the observed panel is a finite realization are treated as unknown but fixed, which is precisely what the deterministic approach is doing. An important feature of step (A), however, is that its result should be a cross-sectionally *exchangeable* process $\mathbf{x}$, i.e., the distributions of any of the resulting $n \times T$ subprocesses should remain invariant under cross-sectional permutations. The cross-sectional ordering, indeed, is completely arbitrary and should not play any role in the analysis.

This random cross-section approach is the one we are adopting in the sequel; the assumptions we are making (along, for some results, with Assumptions (S*(e)*), (S*(f1)*), and (S*(f2)*), which are postponed to Sections 3.3 and 5) under that approach are summarized as follows.

[4] That includes a spectral density $\boldsymbol{\Sigma}_n$ satisfying Assumption (A0), hence spectral densities $\boldsymbol{\Sigma}^{\chi}_n$ and $\boldsymbol{\Sigma}^{\xi}_n$, coefficients $b_{ij,k}$ and $\beta_{ij,k}$, the VAR filters $\mathbf{A}_n(L)$, a representation of the form (18), etc. The densities of $\mathbf{u}_t$ and $\boldsymbol{\eta}_t$, however, remain unspecified within the class of densities satisfying the requirements in Assumption (A1).

Assumption (**S**)(Random cross-section). *(a) Conditional on the result of the random mechanism generating the distributional features of* $\mathbf{x}$*, Assumptions (A0), (A1), and (A2) are satisfied;*

(b) for all n *and* T*, the distribution of* $\mathbf{X}_{nT}$*, hence also the distributions of* $\mathbf{Z}_{nT}$*,* $(\boldsymbol{\Psi}_{nT}, \boldsymbol{\Phi}_{nT})$*, etc. are cross-sectionally exchangeable, i.e., invariant under column permutations;*

(c) $\{\mathbf{R}_i := (R_{i1} \cdots R_{iq})' | \, i \in \mathbb{N}\}$ *is a* q*-dimensional i.i.d. stochastic process independent of* $\{\mathbf{u}_t\}$*, such that, for all* $i \in \mathbb{N}$ *and all* $j = 1, \ldots, q$*, , some* $p > 5$ *and some finite constant* M_R*,* $\mathbb{E}[|R_{ij}|^p] \leq M_R$*;*

(d1) for any fixed $n \in \mathbb{N}$*, there exists a positive definite* $n \times n$ *matrix* $\boldsymbol{\Gamma}_n^{\phi}$*, such that, as* $T \to \infty$*,*

$$\frac{1}{n}\left\|\frac{\boldsymbol{\Phi}'_{nT}\boldsymbol{\Phi}_{nT}}{T} - \boldsymbol{\Gamma}_n^{\phi}\right\| = O_{\mathrm{P}}\left(\frac{1}{\sqrt{T}}\right); \tag{23}$$

(d2) for any fixed $T \in \mathbb{N}$*, there exists a positive definite* $T \times T$ *matrix* $\boldsymbol{G}_T^{\phi}$ *with constant diagonal entries and constant off-diagonal entries such that, as* $n \to \infty$*,*

$$\frac{1}{T}\left\|\frac{\boldsymbol{\Phi}_{nT}\boldsymbol{\Phi}'_{nT}}{n} - \boldsymbol{G}_T^{\phi}\right\| = O_{\mathrm{P}}\left(\frac{1}{\sqrt{n}}\right). \tag{24}$$

Part *(d)* of the assumption is analogue to the requirements in (Bai, 2003, Assumption C). In particular, the existence (not the positiveness) of a limit matrix $\boldsymbol{\Gamma}_n^{\phi}$ in part *(d1)* is, under the randomly generated cross-section approach, quite natural and mild in view of stationarity and the existence of moments. As for part *(d2)*, the special form of $\boldsymbol{G}_T^{\phi}$ follows from the cross-sectional exchangeability for all n of $\boldsymbol{\Phi}_{nT}$. Note also that the linear rate of divergence of exploding eigenvalues in Assumption 2*(b)*, under this approach, is the only rate compatible with cross-sectional exchangeability.

Under Assumption (S), the duality between representations (21) and (22) is reinforced: both now have the form of static factor model representations, with random vectors $\mathbf{u}_t$ loaded at time t by cross-sectional item i via random loadings $\boldsymbol{\mathcal{R}}_n$ in (21) and random vectors $\mathbf{R}_i$ loaded by cross-sectional item i at time t via random loadings $\boldsymbol{\mathcal{U}}_T$ in (22). Both $\mathbf{u}_t$ and $\mathbf{R}_i$ are i.i.d. white noises, the only difference being that $\mathbf{u}_t$ is simply i.i.d. while $\mathbf{R}_i$ also is exchangeable.

Now, the differences between the random cross-section approach (Assumption (S)) and the deterministic cross-sectional approach is tenuous. If (on top of Assumptions (A0)-(A2)) we impose

(a) the deterministic sequence $\frac{1}{n}\sum_{i=1}^{n}\mathbf{R}_i\mathbf{R}_i'$ tends to a $q \times q$ positive definite matrix $\boldsymbol{\Sigma}^R$ as $n \to \infty$

and

(b) (23) and (24), but for the deterministic sequence $\boldsymbol{\Phi}_{nT}$, with $O\left(\frac{1}{\sqrt{T}}\right)$ and $O\left(\frac{1}{\sqrt{n}}\right)$ convergence instead of $O_{\mathrm{P}}\left(\frac{1}{\sqrt{T}}\right)$ and $O_{\mathrm{P}}\left(\frac{1}{\sqrt{n}}\right)$,

then, the random cross-section approach based on Assumption (S) yields, $\mathbb{P}$-a.s. conditionally on step (A) the same results as the deterministic approach based on Assumptions (A0)-(A3). The main benefit of Assumption (S), thus, is to provide a justification of Assumption (A3) (including the special form of $\boldsymbol{G}_T^{\phi}$) and the linear divergence of exploding eigenvalues which otherwise would be brutally imposed.

3.3 Common components

From part *(c)* of Assumption(A1) and the Weak Law of Large Numbers, it immediately follows that

$$\frac{1}{T}\sum_{t=1}^{T}\mathbf{u}_t\mathbf{u}_t' = \frac{\boldsymbol{\mathcal{U}}_T'\boldsymbol{\mathcal{U}}_T}{T} \longrightarrow_{\mathrm{P}} \boldsymbol{\Gamma}^u, \quad \text{as} \quad T \to \infty, \tag{25}$$

where $\mathbf{\Gamma}^u$ is a finite $q \times q$ positive definite matrix. This is the same as Assumption A in Bai (2003). Similarly, from part *(b)* of Assumption (A2), part *(c)* of Assumption (S), and the Weak Law of Large Numbers, there exists a finite $q \times q$ positive definite matrix $\mathbf{\Sigma}^R$ such that

$$\frac{1}{n}\sum_{i=1}^{n}\mathbf{R}_i\mathbf{R}_i' = \frac{\boldsymbol{\mathcal{R}}_n'\boldsymbol{\mathcal{R}}_n}{n} \longrightarrow_{\mathrm{P}} \mathbf{\Sigma}^R, \quad \text{as } n \to \infty, \tag{26}$$

which is the classical condition of factor pervasiveness made in static factor models; in particular, this is the same as Assumption B in Bai (2003), but in the case of random loadings. Moreover, the convergence rates in (25) and (26) are $\sqrt{T}$ and $\sqrt{n}$, respectively (see Lemma 1 in the Appendix).

Now, from Assumption (S*(d*$_1$*)*) and (25),

$$\frac{\boldsymbol{Z}_{nT}'\boldsymbol{Z}_{nT}}{T} = \frac{\boldsymbol{\mathcal{R}}_n\boldsymbol{\mathcal{U}}_T'\boldsymbol{\mathcal{U}}_T\boldsymbol{\mathcal{R}}_n'}{T} + \frac{\mathbf{\Phi}_{nT}'\mathbf{\Phi}_{nT}}{T} + o_{\mathrm{P}}(1) \longrightarrow_{\mathrm{P}} \boldsymbol{\mathcal{R}}_n\mathbf{\Gamma}^u\boldsymbol{\mathcal{R}}_n' + \mathbf{\Gamma}_n^\phi, \quad \text{as } T \to \infty. \tag{27}$$

Letting μ_{nj}^ψ denote the j largest eigenvalue of $\boldsymbol{\mathcal{R}}_n\mathbf{\Gamma}^u\boldsymbol{\mathcal{R}}_n'$, because of (26) and since $\mathbf{\Gamma}^u$ is positive definite, for all $j = 1, \dots, q$, there exist two positive reals α_j^ψ and β_j^ψ such that

$$\alpha_j^\psi \le \text{p-}\liminf_{n\to\infty}\frac{\mu_{nj}^\psi}{n} \le \text{p-}\limsup_{n\to\infty}\frac{\mu_{nj}^\psi}{n} \le \beta_j^\psi. \tag{28}$$

This is similar to Assumption 6 in Forni et al. (2017). Likewise, from Assumption (S*(d*$_2$*)*) and (26),

$$\frac{\boldsymbol{Z}_{nT}\boldsymbol{Z}_{nT}'}{n} = \frac{\boldsymbol{\mathcal{U}}_T\boldsymbol{\mathcal{R}}_n'\boldsymbol{\mathcal{R}}_n\boldsymbol{\mathcal{U}}_T'}{n} + \frac{\mathbf{\Phi}_{nT}\mathbf{\Phi}_{nT}'}{n} + o_{\mathrm{P}}(1) \longrightarrow_{\mathrm{P}} \boldsymbol{\mathcal{U}}_T\mathbf{\Sigma}^R\boldsymbol{\mathcal{U}}_T' + \boldsymbol{G}_T^\phi, \quad \text{as } n \to \infty. \tag{29}$$

Letting ν_{Tj}^ψ denote the j largest eigenvalue of $\boldsymbol{\mathcal{U}}_T\mathbf{\Sigma}^R\boldsymbol{\mathcal{U}}_T'$, because of (25) and since $\mathbf{\Sigma}^R$ is positive definite, for all $j = 1, \dots, q$, there exist two positive reals γ_j^ψ and δ_j^ψ such that

$$\gamma_j^\psi \le \text{p-}\liminf_{T\to\infty}\frac{\nu_{Tj}^\psi}{T} \le \text{p-}\limsup_{T\to\infty}\frac{\nu_{Tj}^\psi}{T} \le \delta_j^\psi. \tag{30}$$

In fact, by the Strong Law of Large Numbers, (25) and (26) hold also almost surely and weak convergence in Lemma 1 in the Appendix could be replaced by almost sure statements with convergence rates $T^{1/2-\epsilon}$ and $n^{1/2-\epsilon}$ for some $\epsilon > 0$.[5] As a consequence, the eigenvalue properties (28) and (30) could be shown to hold with probability one, as in the classical factor models literature.

Consistent estimation of eigenvectors, however, requires the usual assumption of asymptotic separation of eigenvalues—a slight reinforcement of (28) and (30).

Assumption (S*(e)*) (Random cross-section, continued). *With $\mathbb{P}$-probability one, conditional on the result of the random mechanism generating the distributional features of* $\mathbf{x}$,

(e) for $j = 1, \dots, (q-1)$, $\beta_{j+1}^\psi < \alpha_j^\psi$ and $\delta_{j+1}^\psi < \gamma_j^\psi$.

[5] More precisely, it is possible to show that $\left\|\frac{\boldsymbol{\mathcal{U}}_T'\boldsymbol{\mathcal{U}}_T}{T} - \mathbf{\Gamma}^u\right\| = O_{\text{a.s.}}\left(\frac{\log\log T}{\sqrt{T}}\right)$ and $\left\|\frac{\boldsymbol{\mathcal{R}}_n'\boldsymbol{\mathcal{R}}_n}{n} - \mathbf{\Sigma}^R\right\| = O_{\text{a.s.}}\left(\frac{\log\log n}{\sqrt{n}}\right)$.

3.4 Idiosyncratic components

Turning to idiosyncratic components, let $\mathbf{\Sigma}_n^{\phi}(\theta) := \mathbf{A}_n(e^{-\iota\theta})\mathbf{\Sigma}_n^{\xi}(\theta)\mathbf{A}_n'(e^{\iota\theta})$, and denote by $\mathbf{\Lambda}_n^{\phi}(\theta)$ the $n \times n$ diagonal matrix of the eigenvalues $\lambda_{nj}^{\phi}(\theta)$ of $\mathbf{\Sigma}_n^{\phi}(\theta)$. Let $\mathbf{P}^{\phi}(\theta)$, with (i,j)th entry $p_{ij}^{\phi}(\theta)$, be the corresponding $n \times n$ matrix of orthonormal eigenvectors: then, $\mathbf{\Sigma}_n^{\phi}(\theta) = \mathbf{P}_n^{\phi}(\theta)\mathbf{\Lambda}_n^{\phi}(\theta)\mathbf{P}_n^{\phi\dagger}(\theta)$ where $\mathbf{P}_n^{\phi\dagger}$ stands for the transposed complex-conjugate of $\mathbf{P}_n^{\phi}$. We have

$$\sup_{\theta\in[-\pi,\pi]}\sup_{n\in\mathbb{N}}\lambda_{n1}^{\phi}(\theta) \leq \sup_{\theta\in[-\pi,\pi]}\sup_{n\in\mathbb{N}}\lambda_{n1}^{\xi}(\theta)\lambda_{n1}^{A}(\theta) \leq B^{\xi}D^{\phi} =: B^{\phi}, \text{ say,} \tag{31}$$

where $\lambda_{n1}^{A}(\theta)$ is the largest eigenvalue of $\mathbf{A}_n(e^{-\iota\theta})\mathbf{A}_n'(e^{\iota\theta})$, which is finite because of Assumptions (A2a) and (A2c). Moreover, because of (31), the diagonal entries of $\mathbf{\Sigma}_n^{\phi}(\theta)$ are such that

$$\sup_{\theta\in[-\pi,\pi]}\sup_{i\in\mathbb{N}}\sigma_{ii}^{\phi}(\theta) = \sup_{\theta\in[-\pi,\pi]}\sup_{i\in\mathbb{N}}\sum_{j=1}^{n}|p_{ij}^{\phi}(\theta)|^2\lambda_{nj}^{\phi}(\theta) \leq B^{\phi} \tag{32}$$

since eigenvectors are normalized. Notice that $\sigma_{ii}^{\phi}(\theta)$ is real and positive. Similarly, the off-diagonal entries of $\mathbf{\Sigma}_n^{\phi}(\theta)$ satisfy

$$\sup_{\theta\in[-\pi,\pi]}\sup_{i,j\in\mathbb{N}}|\sigma_{ij}^{\phi}(\theta)| \leq \sup_{\theta\in[-\pi,\pi]}\sup_{i,j\in\mathbb{N}}\sum_{k=1}^{n}|p_{ik}^{\phi}(\theta)\bar{p}_{jk}^{\phi}(\theta)|\lambda_{nk}^{\phi}(\theta) \leq \sup_{\theta\in[-\pi,\pi]}\sup_{i,j\in\mathbb{N}}\sum_{k=1}^{n}|p_{ik}^{\phi}(\theta)|^2B^{\phi} \leq B^{\phi} \tag{33}$$

because of (31) and the Cauchy-Schwarz inequality.

Now, from Assumption (S*(d)*), we have that $\mathbf{\Gamma}_n^{\phi} = \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}[\boldsymbol{\phi}_{nt}\boldsymbol{\phi}_{nt}'] = \mathbb{E}[\boldsymbol{\phi}_{nt}\boldsymbol{\phi}_{nt}']$ because of stationarity, while $\boldsymbol{G}_T^{\phi} = \lim_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n}\mathbb{E}[\boldsymbol{\varphi}_T^{i}\boldsymbol{\varphi}_T^{i\prime}] = E[\boldsymbol{\varphi}_T^{i}\boldsymbol{\varphi}_T^{i\prime}]$ because of exchangeability. Therefore, for any $\boldsymbol{b}_n = (b_1\cdots b_n)'$ such that $\boldsymbol{b}_n'\boldsymbol{b}_n = 1$,

$$\begin{aligned}\sup_{n\in\mathbb{N}}\boldsymbol{b}_n'\mathbf{\Gamma}_n^{\phi}\boldsymbol{b}_n &= \sup_{n\in\mathbb{N}}\boldsymbol{b}_n'\mathbb{E}[\boldsymbol{\phi}_{nt}\boldsymbol{\phi}_{nt}']\boldsymbol{b}_n = \sup_{n\in\mathbb{N}}\sum_{i,j=1}^{n}b_ib_j\int_{-\pi}^{\pi}\sigma_{ij}^{\phi}(\theta)\mathrm{d}\theta\\ &\leq \sup_{n\in\mathbb{N}}\sum_{i,j=1}^{n}|b_ib_j|\int_{-\pi}^{\pi}|\sigma_{ij}(\theta)|\mathrm{d}\theta \leq \sup_{n\in\mathbb{N}}\sum_{i=1}^{n}|b_i|^2 2\pi B^{\phi} = 2\pi B^{\phi},\end{aligned} \tag{34}$$

and for any $\boldsymbol{c}_T = (c_1\cdots c_T)'$ such that $\boldsymbol{c}_T'\boldsymbol{c}_T = 1$,

$$\begin{aligned}\sup_{T\in\mathbb{N}}\boldsymbol{c}_T'\boldsymbol{G}_T^{\phi}\boldsymbol{c}_T &= \sup_{T\in\mathbb{N}}\boldsymbol{c}_T'\left\{\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[\boldsymbol{\varphi}_T^{i}\boldsymbol{\varphi}_T^{i\prime}]\right\}\boldsymbol{c}_T = \sup_{T\in\mathbb{N}}\sum_{t,s=1}^{T}c_tc_s\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\left\{\int_{-\pi}^{\pi}\sigma_{ii}^{\phi}(\theta)e^{\iota(t-s)\theta}\mathrm{d}\theta\right\}\\ &\leq \sup_{T\in\mathbb{N}}\sum_{t,s=1}^{T}|c_tc_s|\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\left\{\int_{-\pi}^{\pi}|\sigma_{ii}(\theta)|\,|e^{\iota(t-s)\theta}|\mathrm{d}\theta\right\}\\ &\leq \sup_{T\in\mathbb{N}}\sum_{t,s=1}^{T}|c_tc_s|\sup_{i\in\mathbb{N}}\left\{\int_{-\pi}^{\pi}|\sigma_{ii}(\theta)|\,|e^{\iota(t-s)\theta}|\mathrm{d}\theta\right\} \leq \sup_{T\in\mathbb{N}}\sum_{t=1}^{T}|c_t|^2 2\pi B^{\phi} = 2\pi B^{\phi}.\end{aligned} \tag{35}$$

This implies that the largest eigenvalues of $\mathbf{\Gamma}_n^{\phi}$ and $\boldsymbol{G}_T^{\phi}$ satisfy

$$\sup_{n\in\mathbb{N}} \|\mathbf{\Gamma}_n^{\phi}\| = \sup_{n\in\mathbb{N}} \max_{\substack{\boldsymbol{b}_n \\ \boldsymbol{b}_n'\boldsymbol{b}_n=1}} \boldsymbol{b}_n'\mathbf{\Gamma}_n^{\phi}\boldsymbol{b}_n \leq 2\pi B^{\phi} \text{ and} \tag{36}$$

$$\sup_{T\in\mathbb{N}} \|\boldsymbol{G}_T^{\phi}\| = \sup_{T\in\mathbb{N}} \max_{\substack{\boldsymbol{c}_T \\ \boldsymbol{c}_T'\boldsymbol{c}_T=1}} \boldsymbol{c}_T'\boldsymbol{G}_T^{\phi}\boldsymbol{c}_T \leq 2\pi B^{\phi}, \tag{37}$$

respectively. Following a similar reasoning, it is straightforward to show that also Assumptions C1 and C3 of Bai (2003) hold.

d column eigenvectors, associated with the q largest

4 Estimation

In order to estimate the common component we need to estimate the common filters, i.e., the impulse response functions, $\mathbf{B}_n(L) = [\mathbf{A}_n(L)]^{-1}\boldsymbol{\mathcal{R}}_n$, and the common factors $\mathbf{u}_t$. That estimation proceeds in two steps: first we estimate $\mathbf{A}_n(L)$ and then, by considering the static representation (18) of the GDFM, we estimate $\mathbf{u}_t$ and $\boldsymbol{\mathcal{R}}_n$ by a principal component analysis of the filtered data $\mathbf{z}_{nt} = \mathbf{A}_n(L)\mathbf{x}_{nt}$. This section describes the estimators while Section 5 is devoted to their asymptotic properties.

First, notice that we can consistently determine the number q of factors by applying the Hallin and Liška (2007) information criteria to the observed data matrix $\mathbf{X}_{nT}$. The resulting estimator $\widehat{q}$ converges in probability to q as $n, T \to \infty$. Since q is integer, this means that, for any $\epsilon > 0$, there exist $n^*(\epsilon)$ and $T^*(\epsilon)$ such that, $\mathrm{P}(\widehat{q} = q) > 1 - \epsilon$ for all $n > n^*(\epsilon)$ and $T > T^*(\epsilon)$. Hence, in this section and the next one, we can safely assume that q is known.

4.1 Estimation of $\mathbf{A}(L)$

Without loss of generality we keep assuming $n = m(q+1)$ for some finite integer m (we discuss below what to do in practice if this is not the case). To start with, we compute the lag-window estimator

$$\widehat{\mathbf{\Sigma}}_n(\theta_h) := \frac{1}{2\pi} \sum_{k=-T+1}^{T-1} \mathrm{K}\left(\frac{k}{B_T}\right) e^{-\iota k\theta_h} \widehat{\mathbf{\Gamma}}_{n,k}, \quad \theta_h = \frac{\pi h}{B_T}, \quad |h| \leq B_T, \tag{38}$$

of the spectral density matrix of the observables; here $\widehat{\mathbf{\Gamma}}_{n,k} := T^{-1}\sum_{t=|k|+1}^{T} \mathbf{x}_{nt}\mathbf{x}_{n,t-|k|}'$ is the usual lag-k sample autocovariance matrix and $\mathrm{K}(\cdot)$ is a suitable kernel with bandwidth B_T (see Assumption (K) in Section 5.1).

Then, we estimate the spectral density matrix of the common component by dynamic principal component analysis. Specifically, we collect the normalized column eigenvectors associated with the q largest eigenvalues of $\widehat{\mathbf{\Sigma}}_n(\theta_h)$ into the $n \times q$ matrix $\widehat{\mathbf{P}}_n(\theta_h)$, and the corresponding eigenvalues into the $q \times q$ diagonal matrix $\widehat{\mathbf{\Lambda}}_n(\theta_h)$. Our estimator of the spectral density matrix of the common component is defined as

$$\widehat{\mathbf{\Sigma}}_n^{\chi}(\theta_h) := \widehat{\mathbf{P}}_n(\theta_h)\widehat{\mathbf{\Lambda}}_n(\theta_h)\widehat{\mathbf{P}}_n^{\dagger}(\theta_h),$$

where $\widehat{\mathbf{P}}_n^{\dagger}(\theta_h)$ is the transposed complex-conjugate of $\widehat{\mathbf{P}}_n(\theta_h)$.

By computing the inverse Fourier transform of $\widehat{\mathbf{\Sigma}}_n^{\chi}(\theta_h)$, we can estimate the autocovariance matrices of the common component:

$$\widehat{\mathbf{\Gamma}}^{\chi}_{n,k} := \frac{\pi}{B_T} \sum_{h=-B_T}^{B_T} e^{\iota k \theta_h} \widehat{\mathbf{\Sigma}}^{\chi}_n(\theta_h), \qquad |k| \leq B_T.$$

Consider the m diagonal $(q+1) \times (q+1)$ blocks $\widehat{\mathbf{\Gamma}}^{\chi(s)}_k$ of the $\widehat{\mathbf{\Gamma}}^{\chi}_{n,k}$'s. For each block, estimate, via the Yule-Walker method, the coefficients of a $(q+1)$-dimensional VAR model (order determined via AIC or BIC). This yields, for the s-th diagonal block, an estimator $\widehat{\mathbf{A}}^{(s)}(L)$ of the autoregressive filter $\mathbf{A}^{(s)}(L)$ appearing in Assumption (A2).[6] By combining the m estimators for the m diagonal blocks $\mathbf{A}^{(1)}(L), \ldots, \mathbf{A}^{(m)}(L)$, we obtain an estimator $\widehat{\mathbf{A}}_n(L)$ of the VAR filter $\mathbf{A}_n(L)$ as defined in (15).

Three important remarks about estimation of $\mathbf{A}_n(L)$ are in order here.

Remark 1. The cross-sectional ordering of the panel has an impact on the selection of the diagonal blocks when estimating $\mathbf{A}_n(L)$. Each cross-sectional permutation of the panel, thus, would lead to distinct estimators—all sharing the same asymptotic properties. In line with the exchangeability property Assumption (S*(b)*), a Rao-Blackwell argument (see Forni et al., 2017 for details) suggests aggregating these estimators into a unique one by simple averaging (after obvious reordering of the cross-section) of the resulting estimated shocks. Although averaging over all $n!$ permutations is clearly unfeasible, as explained by Forni et al. (2017) and verified empirically also in Forni et al. (2018), a few of them are enough, in practice, to deliver stable averages, well-approximating the infeasible average over all $n!$ permutations.

Remark 2. Although we assumed for simplicity that $n = m(q+1)$ for some integer m, this might not be the case in practice. When n is not an integer multiple of $(q+1)$, we can consider $\lfloor n/(q+1) \rfloor - 1$ blocks of size $(q+1)$ and a last one of size $(q+1) + n - \lfloor n/(q+1) \rfloor (q+1)$ larger than $(q+1)$ but smaller than $2(q+1)$. Since the arguments from Forni et al. (2017) used in the next section apply to any partition into blocks of size $(q+1)$ or larger, nothing changes for the asymptotic theory that follows.

Remark 3. It is known that, as p_s increases, the estimation of a singular VAR via Yule-Walker methods may become unstable, since it requires inversion of a $p_s(q+1) \times p_s(q+1)$ Toeplitz matrix. To tame this potential issue, Hörmann and Nisol (2020) have proposed a regularized approach, aimed at stabilizing the estimates $\widehat{\mathbf{A}}^{(s)}(L)$. Empirically, this seems to be an important step—to be taken only when p_s is much larger than 1, though.

4.2 Estimation of $\boldsymbol{\mathcal{U}}_T$ and $\boldsymbol{\mathcal{R}}_n$

Letting $\widehat{\mathbf{Z}}'_{nT} := \widehat{\mathbf{A}}'_n(L)\mathbf{X}'_{nT}$, we propose to estimate the static model (18) by (static) principal component of $\widehat{\mathbf{Z}}_{nT}$ twice. The reason for this is that we aim at getting estimators of both $\boldsymbol{\mathcal{R}}_n$ and $\boldsymbol{\mathcal{U}}_T$ as linear projections, instead of getting one as a projection and the other as the normalized eigenvectors of a sample covariance matrix (as in Forni et al. (2015, 2017)). This is made possible by exploiting the duality between the two representations (21) and (22) of the static model. This double estimation procedure is the key to the derivation, in Section 5 below, of the asymptotic distributions of the estimators, while the asymptotics of normalized eigenvectors of sample covariance matrices are considerably more intricate.

[6]For example, in the VAR(1) case, i.e., $p_s = 1$, we have $\widehat{\mathbf{A}}^{(s)}(L) = \mathbf{I}_{q+1} - \widehat{\mathbf{A}}^{(s)} L$ with $\widehat{\mathbf{A}}^{(s)} := \widehat{\mathbf{\Gamma}}^{\chi(s)}_1 (\widehat{\mathbf{\Gamma}}^{\chi(s)}_0)^{-1}$.

Let us start with the estimation of $\boldsymbol{\mathcal{U}}_T$. Consider the $n \times n$ sample covariance matrix

$$\widehat{\mathbf{\Gamma}}_n^z := \frac{1}{T}\sum_{t=1}^T \widehat{\mathbf{z}}_{nt}\widehat{\mathbf{z}}_{nt}' = \frac{\widehat{\mathbf{Z}}_{nT}'\widehat{\mathbf{Z}}_{nT}}{T}.$$

of the $\widehat{\mathbf{z}}_{nt}$'s. Collect the normalized column eigenvectors associated with the q largest eigenvalues of $\widehat{\mathbf{\Gamma}}_n^z$ into the $n \times q$ matrix $\widehat{\mathbf{P}}_n^z$ and the corresponding eigenvalues into the $q \times q$ diagonal matrix $\widehat{\mathbf{\Lambda}}_n^z$. Then, for the estimation of $\boldsymbol{\mathcal{U}}_T$, construct a preliminary estimator of $\boldsymbol{\mathcal{R}}_n$ as

$$\check{\boldsymbol{\mathcal{R}}}_n = \left(\check{\mathbf{R}}_1 \cdots \check{\mathbf{R}}_n\right)' := \widehat{\mathbf{P}}_n^z\left(\widehat{\mathbf{\Lambda}}_n^z\right)^{1/2}. \tag{39}$$

Next consider the submatrix of $\check{\boldsymbol{\mathcal{R}}}_n$ consisting of a selection of $\bar{n}$ rows with $\bar{n} \leq n$. Without loss of generality, we can assume that the first $\bar{n}$ rows are selected, and define

$$\check{\boldsymbol{\mathcal{R}}}_{\bar{n}} = \left(\check{\mathbf{R}}_1 \cdots \check{\mathbf{R}}_{\bar{n}}\right)' := \widehat{\mathbf{P}}_{\bar{n}}^z\left(\widehat{\mathbf{\Lambda}}_n^z\right)^{1/2}, \tag{40}$$

where $\widehat{\mathbf{P}}_{\bar{n}}^z$ is the $\bar{n} \times q$ submatrix of $\widehat{\mathbf{P}}_n^z$'s first $\bar{n}$ rows. Note that each entry of $\check{\boldsymbol{\mathcal{R}}}_{\bar{n}}$ continues to be function of n and T only; in particular the matrix of eigenvalues $\widehat{\mathbf{\Lambda}}_n^z$ does not depend on $\bar{n}$.

Then, let $\widehat{\mathbf{Z}}_{\bar{n}T} = (\widehat{\boldsymbol{z}}_T^1 \cdots \widehat{\boldsymbol{z}}_T^{\bar{n}})$ be the $T \times \bar{n}$ matrix of $\widehat{\mathbf{Z}}_{nT}$'s first $\bar{n}$ columns. We estimate $\boldsymbol{\mathcal{U}}_T$ as the cross-sectional linear projection $\widehat{\boldsymbol{\mathcal{U}}}_T$ of the $\widehat{\boldsymbol{z}}_T^i$s onto $\check{\boldsymbol{\mathcal{R}}}_{\bar{n}}$: namely,

$$\begin{aligned}
\widehat{\boldsymbol{\mathcal{U}}}_T = (\widehat{\mathbf{u}}_1 \cdots \widehat{\mathbf{u}}_t \cdots \widehat{\mathbf{u}}_T)' &:= \widehat{\mathbf{Z}}_{\bar{n}T}\check{\boldsymbol{\mathcal{R}}}_{\bar{n}}\left(\check{\boldsymbol{\mathcal{R}}}_{\bar{n}}'\check{\boldsymbol{\mathcal{R}}}_{\bar{n}}\right)^{-1} \\
&= \widehat{\mathbf{Z}}_{\bar{n}T}\widehat{\mathbf{P}}_{\bar{n}}^z\left(\widehat{\mathbf{\Lambda}}_n^z\right)^{1/2}\left(\left(\widehat{\mathbf{\Lambda}}_n^z\right)^{1/2}\widehat{\mathbf{P}}_{\bar{n}}^{z\prime}\widehat{\mathbf{P}}_{\bar{n}}^z\left(\widehat{\mathbf{\Lambda}}_n^z\right)^{1/2}\right)^{-1} \\
&= \widehat{\mathbf{Z}}_{\bar{n}T}\widehat{\mathbf{P}}_{\bar{n}}^z\left(\widehat{\mathbf{\Lambda}}_n^z\right)^{-1/2}.
\end{aligned} \tag{41}$$

This (not $\check{\boldsymbol{\mathcal{U}}}_T$ defined in (42) below) is the estimator we are proposing for $\boldsymbol{\mathcal{U}}_T$.

Turning to the estimation of $\boldsymbol{\mathcal{R}}_n$, consider the $T \times T$ sample covariance matrix

$$\widehat{\boldsymbol{G}}_T^z := \frac{1}{n}\sum_{i=1}^n \widehat{\boldsymbol{z}}_T^i\widehat{\boldsymbol{z}}_T^{i\prime} = \frac{\widehat{\mathbf{Z}}_{nT}\widehat{\mathbf{Z}}_{nT}'}{n}.$$

of the $\widehat{\boldsymbol{z}}_T^i$'s. Collect the normalized column eigenvectors associated with the q largest eigenvalues of $\widehat{\boldsymbol{G}}_T^z$ into the $n \times q$ matrix $\widehat{\mathbf{\Pi}}_T^z$, and the corresponding eigenvalues into the $q \times q$ diagonal matrix $\widehat{\boldsymbol{L}}_T^z$. Then, for the estimation of $\boldsymbol{\mathcal{R}}_n$, construct a preliminary estimator $\check{\boldsymbol{\mathcal{U}}}_T$ of $\boldsymbol{\mathcal{U}}_T$ as

$$\check{\boldsymbol{\mathcal{U}}}_T = (\check{\mathbf{u}}_1 \cdots \check{\mathbf{u}}_T)' := \widehat{\mathbf{\Pi}}_T^z\left(\widehat{\boldsymbol{L}}_T^z\right)^{1/2}. \tag{42}$$

Next consider the submatrix of $\check{\boldsymbol{\mathcal{U}}}_T$ consisting of a selection of $\bar{T}$ rows with $\bar{T} \leq T$. Without loss of generality, we can assume that the first $\bar{T}$ rows are selected, and define

$$\check{\boldsymbol{\mathcal{U}}}_{\bar{T}} = (\check{\mathbf{u}}_1 \cdots \check{\mathbf{u}}_{\bar{T}})' := \widehat{\mathbf{\Pi}}_{\bar{T}}^z\left(\widehat{\boldsymbol{L}}_T^z\right)^{1/2},$$

where $\widehat{\mathbf{\Pi}}^z_{\bar{T}}$ is the $\bar{T} \times q$ submatrix of $\widehat{\mathbf{\Pi}}^z_T$'s first $\bar{T}$ rows. Note that each entry of $\check{\boldsymbol{\mathcal{U}}}_{\bar{T}}$ continues to be function of n and T only; in particular the matrix of eigenvalues $\widehat{\boldsymbol{L}}^z_T$ does not depend on $\bar{T}$.

Then, let $\widehat{\mathbf{Z}}_{n\bar{T}} = (\widehat{\mathbf{z}}_{n1} \cdots \widehat{\mathbf{z}}_{n\bar{T}})'$ be the $\bar{T} \times n$ matrix of $\widehat{\mathbf{Z}}_{nT}$'s first $\bar{T}$ rows. We estimate $\boldsymbol{\mathcal{R}}_n$ as the time-series linear projection $\widehat{\boldsymbol{\mathcal{R}}}_n$ of the $\widehat{\mathbf{z}}_{nt}$'s onto $\check{\boldsymbol{\mathcal{U}}}_T$: namely,

$$
\begin{aligned}
\widehat{\boldsymbol{\mathcal{R}}}_n = \left(\widehat{\mathbf{R}}_1 \cdots \widehat{\mathbf{R}}_i \cdots \widehat{\mathbf{R}}_n\right)' &:= \widehat{\mathbf{Z}}'_{n\bar{T}} \check{\boldsymbol{\mathcal{U}}}_{\bar{T}} \left(\check{\boldsymbol{\mathcal{U}}}'_{\bar{T}} \check{\boldsymbol{\mathcal{U}}}_{\bar{T}}\right)^{-1} \\
&= \widehat{\mathbf{Z}}'_{n\bar{T}} \widehat{\mathbf{\Pi}}^z_{\bar{T}} \left(\widehat{\boldsymbol{L}}^z_T\right)^{1/2} \left(\left(\widehat{\boldsymbol{L}}^z_T\right)^{1/2} \widehat{\mathbf{\Pi}}^{z\prime}_{\bar{T}} \widehat{\mathbf{\Pi}}^z_{\bar{T}} \left(\widehat{\boldsymbol{L}}^z_T\right)^{1/2}\right)^{-1} \\
&= \widehat{\mathbf{Z}}'_{\bar{T}} \widehat{\mathbf{\Pi}}^z_{\bar{T}} \left(\widehat{\boldsymbol{L}}^z_T\right)^{-1/2}. \qquad (43)
\end{aligned}
$$

This (not $\check{\boldsymbol{\mathcal{R}}}_n$ defined in (39) above) is the estimator we are proposing for $\boldsymbol{\mathcal{R}}_n$.

Summing up, we have two sets of estimators for $\boldsymbol{\mathcal{U}}_T$, namely $\widehat{\boldsymbol{\mathcal{U}}}_T$ and $\check{\boldsymbol{\mathcal{U}}}_T$, and two sets of estimators for $\boldsymbol{\mathcal{R}}_n$, namely $\widehat{\boldsymbol{\mathcal{R}}}_n$ and $\check{\boldsymbol{\mathcal{R}}}_n$. For the purpose of inference, we will consider $\widehat{\boldsymbol{\mathcal{U}}}_T$ and $\widehat{\boldsymbol{\mathcal{R}}}_n$ because they are constructed as least squares projections, thus involve averaging, which allows for deriving asymptotic distributions (see Section 5). Instead, $\check{\boldsymbol{\mathcal{R}}}_n$ and $\check{\boldsymbol{\mathcal{U}}}_T$ are (rescaled) eigenvector matrices the asymptotic distributions of which are less obvious due to the latent nature of $\mathbf{z}_{nT}$.

Combining these estimators yields two different estimators of the elements ψ_{it} of the static common component $\mathbf{\Psi}_{nT}$, such as $\widehat{\mathbf{R}}'_i \check{\mathbf{u}}_t$ and $\check{\mathbf{R}}'_i \widehat{\mathbf{u}}_t$. However, as discussed in Section 5.4 below, more efficient estimators are convex linear combinations of the form

$$
\widehat{\psi}_{it} := \omega_{nT} \check{\mathbf{R}}'_i \widehat{\mathbf{u}}_t + (1 - \omega_{nT}) \widehat{\mathbf{R}}'_i \check{\mathbf{u}}_t, \quad i = 1, \ldots, n, \ t = 1, \ldots, T, \qquad (44)
$$

where the weights ω_{nT} are such that $\omega_{nT} = 1/2$ if $n = T$, $\omega_{nT} \uparrow 1$ if $n/T \downarrow 0$, and $\omega_{nT} \downarrow 0$ if $T/n \downarrow 0$.

5 Asymptotic properties

5.1 Asymptotics for $\mathbf{A}_n(L)$

The first step in our estimation procedure is the computation of a lag-window estimator (38) of the spectral density matrix $\mathbf{\Sigma}_n(\theta)$. This requires a kernel $K(\cdot)$ and a bandwidth B_T on which we make the following standard assumptions.

Assumption (K)(Lag-window estimation)**.** *(a) the kernel* K *is even, bounded, with support* $[-1, 1]$*, and*

(i) $|\mathrm{K}(u) - 1| = O(|u|^\kappa)$*, as* $u \to 0$*, for some positive real* κ*;*

(ii) $\int_{-1}^1 \mathrm{K}^2(u) \mathrm{d}u < \infty$*;*

(iii) $\sum_{j \in \mathbb{Z}} \sup_{|s-j| \le 1} |\mathrm{K}(jw) - \mathrm{K}(sw)| = O(1)$*, as* $w \to 0$*;*

(b) the bandwidth B_T *is such that* $c_1 T^\delta \le B_T \le c_2 T^\delta$ *for some* $0 < \delta < 1$ *and positive reals* c_1 *and* c_2*.*

Let $\sigma^\chi_{ij}(\theta)$ and $\widehat{\sigma}_{ij}(\theta)$, $i, j = 1, \ldots, n$, denote the (i,j)th entries of $\mathbf{\Sigma}^\chi(\theta)$ and $\widehat{\mathbf{\Sigma}}(\theta)$, respectively. Building on recent results on the estimation of large spectral density matrices (Wu and Zaffaroni, 2018; Zhang and Wu, 2021), Forni et al. (2017, Propositions 6 and 7) prove the following result (see also Barigozzi et al., 2021, Lemma 4 and Proposition 1).

Proposition 1. *Let* $\eta_{T;\kappa,p} := \max\left(\sqrt{\frac{B_T \log T}{T}}, \frac{T^{2/p} B_T (\log T)^{2+2/p}}{T}, \frac{1}{B_T^\kappa}\right)$*, where* p *is defined in parts (c) and (d) of Assumption (A1),* B_T *and* κ *in Assumption (K). Then, under Assumptions (S) and (K), for* *any $\epsilon > 0$, there exist $\eta(\epsilon)$, $T^*(\epsilon)$, and $n^*(\epsilon)$, all independent of i and j, such that*

$$\mathrm{P}\left(\max_{|h|\le B_T}\frac{|\widehat{\sigma}_{ij}(\theta_h)-\sigma_{ij}(\theta_h)|}{\max\left(\eta_{T;\kappa,p},\frac{1}{\sqrt{n}}\right)}\ge\eta(\epsilon)\right)\le\epsilon \tag{45}$$

for all $T > T^*(\epsilon)$ and all $n > n^*(\epsilon)$.

The rate $\eta_{T;\kappa,p}$ in (45) depends on (i) the kernel smoothness κ, (ii) the bandwidth B_T which, by Assumption (K), is such that $B_T \asymp T^\delta$, and (iii) the minimum number p of moments we allow to exist. Typical values for κ are 1 for the Bartlett kernel, and 2 for the Parzen, Daniell, General Tukey, Tukey-Hanning, Tukey-Hamming, and Bartlett-Priestley kernels (see Priestley, 1982, p. 463). To determine the optimal rate, notice that $\eta_{T;\kappa,p}$ is the maximum of three terms. The first one is larger than the third if $\delta \ge \frac{1}{2\kappa+1}$: hence, given the choice of a kernel among Bartlett, Parzen, Daniell, General Tukey, Tukey-Hanning, Tukey-Hamming, and Bartlett-Priestley, we need to set either $\delta \ge \frac{1}{3}$ or $\delta \ge \frac{1}{5}$. Moreover, the first term in $\eta_{T;\kappa,p}$ is always larger than the second one if $\delta \le 1-\frac{4}{p}$. For $p > 5$, as per Assumption (A1b) and (A1c), the choice of $\kappa = 2$ and $\delta = \frac{1}{5}$ yields a rate $\eta_{T;\kappa,p} = \frac{1}{T^{2/5}}$, while the choice of $\kappa = 1$ and $\delta = \frac{1}{3}$ yields $\eta_{T;\kappa,p} = \frac{1}{T^{1/3}}$. Hereafter, we define $\zeta_{n,T} := \max\left(\eta_{T;\kappa,p}, \frac{1}{\sqrt{n}}\right)$, dropping for simplicity the dependence on κ and p.

Let $\mathbf{A}^{[s]} := (\mathbf{A}_1^{(s)}\cdots\mathbf{A}_{p_s}^{(s)})$ and $\widehat{\mathbf{A}}^{[s]} := (\widehat{\mathbf{A}}_1^{(s)}\cdots\widehat{\mathbf{A}}_{p_s}^{(s)})$ for $s = 1,\ldots,m$. Then, Forni et al. (2017, Proposition 9) prove the following.

Proposition 2. *Under Assumptions (S) and (K), for any $s = 1,\ldots,m$, $\|\widehat{\mathbf{A}}^{[s]}-\mathbf{A}^{[s]}\| = O_{\mathrm{P}}(\zeta_{n,T})$ as $n,T\to\infty$. .*

5.2 Asymptotics for $\widehat{\boldsymbol{\mathcal{U}}}_T$

Considering the spectral decomposition

$$\boldsymbol{\mathcal{R}}_n\boldsymbol{\Gamma}^u\boldsymbol{\mathcal{R}}_n' = \mathbf{P}_n^\psi\boldsymbol{\Lambda}_n^\psi\mathbf{P}_n^{\psi\prime}, \tag{46}$$

where $\boldsymbol{\Lambda}_n^\psi$ is the $q\times q$ diagonal matrix of $\boldsymbol{\mathcal{R}}_n\boldsymbol{\Gamma}^u\boldsymbol{\mathcal{R}}_n'$'s eigenvalues and $\mathbf{P}_n^\psi$ the $n\times q$ matrix with columns the corresponding orthonormal eigenvectors, we make the following assumption.

Assumption (S*(f1)*)(Random cross-section, continued). *Let $\bar{n}\le n$ be such that $\frac{1}{\bar{n}}+\frac{\bar{n}}{n}\to 0$ as $n\to\infty$ (that is, $\bar{n}\to\infty$ and $\bar{n}/n\to 0$ as $n\to\infty$). Then, for any $t\in\mathbb{Z}$,*

$$\sqrt{\frac{n}{\bar{n}}}\mathbf{P}_{\bar{n}}^{\psi\prime}\boldsymbol{\phi}_{\bar{n}t}\longrightarrow_d\mathcal{N}(\mathbf{0}_q,\boldsymbol{\mathcal{P}}_t^u)\qquad \text{as } n\to\infty, \tag{47}$$

where $\boldsymbol{\mathcal{P}}_t^u := \lim_{n\to\infty}\frac{n}{\bar{n}}\mathbb{E}[\mathbf{P}_{\bar{n}}^{\psi\prime}\boldsymbol{\phi}_{\bar{n}t}\boldsymbol{\phi}_{\bar{n}t}'\mathbf{P}_{\bar{n}}^\psi]$ is positive definite, and $\mathbf{0}_q$ is a q-dimensional vector of zeros.

Note that $\boldsymbol{\mathcal{P}}_t^u$ is not $\lim_{n\to\infty}\frac{n}{\bar{n}}\mathbf{P}_{\bar{n}}^{\psi\prime}\mathbb{E}[\boldsymbol{\phi}_{\bar{n}t}\boldsymbol{\phi}_{\bar{n}t}']\mathbf{P}_{\bar{n}}^\psi$ since eigenvectors are random; so we must assume its existence. A similar assumption is made also in Bai (2003, Assumption F3) in the case of non-random eigenvectors.

Theorem 1. *Denote by $\widehat{\mathbf{W}}^z$ a $q\times q$ diagonal matrix, depending on n and T, with diagonal entries ±1. Then, for any $t = 1,\ldots,T$ and any $\bar{n}\le n$ such that $\frac{1}{\bar{n}}+\frac{\bar{n}}{n}\to 0$, as $n,T\to\infty$,*

(i) *under Assumptions (S), (S(e)), and (K),*

$$\left\| \widehat{\mathbf{u}}_t - \widehat{\mathbf{W}}^z \left(\boldsymbol{\Gamma}^u\right)^{-1/2} \mathbf{u}_t \right\| = O_{\mathrm{P}} \left(\max \left(\frac{1}{\sqrt{\bar{n}}}, \zeta_{nT} \right) \right);$$

(ii) *under Assumptions (S), (S(e)), (S(f1)), and (K), with $\bar{n}$ such that*

$$\frac{1}{\bar{n}} + \sqrt{\bar{n}}\, \zeta_{nT} \to 0, \quad \text{as } n, T \to \infty, \tag{48}$$

and $\widehat{\mathbf{W}}^z \longrightarrow_{\mathrm{P}} \mathbf{W}^u$, letting $\boldsymbol{\mathcal{M}}^u := \mathrm{plim}_{n\to\infty} \frac{n}{\bar{n}} \mathbf{P}_{\bar{n}}^{\psi\prime} \mathbf{P}_{\bar{n}}^{\psi}$ and $\boldsymbol{\mathcal{L}}^u = \mathrm{plim}_{n\to\infty} \frac{\boldsymbol{\Lambda}_n^{\psi}}{n}$,

$$\sqrt{\bar{n}} \left(\widehat{\mathbf{u}}_t - \widehat{\mathbf{W}}^z \left(\boldsymbol{\Gamma}^u\right)^{-1/2} \mathbf{u}_t \right) \longrightarrow_d \mathcal{N} \left(\mathbf{0}_q, \mathbf{W}^u \left(\boldsymbol{\mathcal{L}}^u\right)^{-1/2} \left(\boldsymbol{\mathcal{M}}^u\right)^{-1} \boldsymbol{\mathcal{P}}_t^u \left(\boldsymbol{\mathcal{M}}^u\right)^{-1} \left(\boldsymbol{\mathcal{L}}^u\right)^{-1/2} \mathbf{W}^u \right)$$

where $\boldsymbol{\mathcal{P}}_t^u$ is defined in in part (f1) of Assumption (S).

Remark 4. In terms of rate of convergence (part (i) of Theorem 1), for $\bar{n} = n$, we have the rate ζ_{nT}^{-1} as already derived by Forni et al. (2017, Proposition 11). In particular, ours and Bai (2003) estimators of $\mathbf{u}_t$ converge at the same rate $\sqrt{n}$ when $T/(B_T n) \downarrow 0$, whereas when $(nB_T)/T \downarrow 0$ we achieve a rate of convergence $\sqrt{T/B_T}$, which is slower than the rates $\sqrt{n}$ or T (depending on whether $\sqrt{n}/T \downarrow 0$ or $T/\sqrt{n} \downarrow 0$) in Bai (2003).

Remark 5. Condition (48) imposes only a marginally slower rate than ζ_{nT}, which is the consistency rate when $\bar{n} = n$. For example we can assume $\bar{n}$ of the form $\bar{n} = \zeta_{nT}^{-2} L^{-1}(\zeta_{nT}^{-1})$ for some slowly varying at infinity function $L(\cdot)$ (this implies that $\bar{n} \simeq n$, hence is a viable choice). Note that $\bar{n}$ then depends on both n and T. In fact, by inspection of the proof of part (i), we can see that consistency holds with a faster rate and we could relax (48) to $\frac{\bar{n}}{\sqrt{n}} \zeta_{nT} \to 0$. However, since for deriving the properties of the common component, we still need to impose (48), we stick with it also in Theorem 1.

Remark 6. When (48) does not hold, Theorem 1 states that, as $n, T \to \infty$,

$$\sqrt{\bar{n}} \left(\widehat{\mathbf{u}}_t - \widehat{\mathbf{W}}^z \left(\boldsymbol{\Gamma}^u\right)^{-1/2} \mathbf{u}_t \right) - \widehat{\mathbf{W}}^z \boldsymbol{\mathcal{X}}_t \longrightarrow_{\mathrm{P}} \mathbf{0}_q,$$

for some random vector $\boldsymbol{\mathcal{X}}_t \sim \mathcal{N}(\mathbf{0}_q, \left(\boldsymbol{\mathcal{L}}^u\right)^{-1/2} \left(\boldsymbol{\mathcal{M}}^u\right)^{-1} \boldsymbol{\mathcal{P}}_t^u \left(\boldsymbol{\mathcal{M}}^u\right)^{-1} \left(\boldsymbol{\mathcal{L}}^u\right)^{-1/2})$.

Remark 7. A consistent estimator of the asymptotic covariance matrix of $\sqrt{\bar{n}}(\widehat{\mathbf{u}}_t - \widehat{\mathbf{W}}^z \left(\boldsymbol{\Gamma}^u\right)^{-1/2} \mathbf{u}_t)$ is

$$\left(\frac{\widehat{\boldsymbol{\Lambda}}_n^z}{n} \right)^{-1/2} \left(\frac{n}{\bar{n}} \widehat{\mathbf{P}}_{\bar{n}}^{z\prime} \widehat{\mathbf{P}}_{\bar{n}}^{z} \right)^{-1} \widehat{\boldsymbol{\mathcal{P}}}_t^u \left(\frac{n}{\bar{n}} \widehat{\mathbf{P}}_{\bar{n}}^{z\prime} \widehat{\mathbf{P}}_{\bar{n}}^{z} \right)^{-1} \left(\frac{\widehat{\boldsymbol{\Lambda}}_n^z}{n} \right)^{-1/2},$$

where $\widehat{\boldsymbol{\mathcal{P}}}_t^u$ is a consistent estimator of $\boldsymbol{\mathcal{P}}_t^u$. This requires specific assumptions on the form of cross-sectional dependence of the $\{\phi_{it}\}$. For instance, when the latter are cross-sectionally independent, then the approach of Section 5(a) in Bai (2003) can be adapted, providing[7]

$$\widehat{\boldsymbol{\mathcal{P}}}_t^u = \sum_{i=1}^n \widehat{\mathbf{p}}_i^z \widehat{\mathbf{p}}_i^{z\prime} \left\{ \frac{1}{T} \sum_{t=1}^T \widehat{\phi}_{it}^2 \right\}, \tag{49}$$

where $\widehat{\phi}_{it} = \widehat{z}_{it} - \check{\mathbf{R}}_i' \widehat{\mathbf{u}}_t$ and $\widehat{\mathbf{p}}_i^{z\prime}$ is the ith row of $\widehat{\mathbf{P}}_n^z$ (defined in (40)).

5.3 Asymptotics for $\widehat{\boldsymbol{\mathcal{R}}}_n$

Thanks to the duality between (21) and (22), the asymptotics for $\widehat{\boldsymbol{\mathcal{R}}}_n$ follow along the same lines as for $\widehat{\boldsymbol{\mathcal{U}}}_T$. Consider the spectral decomposition

$$\boldsymbol{\mathcal{U}}_T \boldsymbol{\Sigma}^R \boldsymbol{\mathcal{U}}_T' = \boldsymbol{\Pi}_T^\psi \boldsymbol{L}_T^\psi \boldsymbol{\Pi}_T^{\psi\prime}, \tag{50}$$

where $\boldsymbol{L}_T^\psi$ is the $q \times q$ diagonal matrix of $\boldsymbol{\mathcal{U}}_T \boldsymbol{\Sigma}^R \boldsymbol{\mathcal{U}}_T'$'s eigenvalues and $\boldsymbol{\Pi}_T^\psi$ the $T \times q$ with columns the corresponding orthonormal eigenvectors. Similar to (47), we make the following assumption.

Assumption (**S**(*f2*))(Random cross-section, continued)**.** *Let $\bar{T} \leq T$ be such that $\frac{1}{\bar{T}} + \frac{\bar{T}}{T} \to 0$ as $T \to \infty$. Then,*

$$\sqrt{\frac{T}{\bar{T}}} \boldsymbol{\Pi}_{\bar{T}}^{\psi\prime} \boldsymbol{\varphi}_{\bar{T}}^i \longrightarrow_d \mathcal{N}(\mathbf{0}_q, \boldsymbol{\mathcal{P}}^R) \qquad \text{as } T \to \infty \tag{51}$$

where $\boldsymbol{\mathcal{P}}_i^R := \lim_{T\to\infty} \frac{T}{\bar{T}} \mathbb{E}[\boldsymbol{\Pi}_{\bar{T}}^{\psi\prime} \boldsymbol{\varphi}_{\bar{T}}^i \boldsymbol{\varphi}_{\bar{T}}^{i\prime} \boldsymbol{\Pi}_{\bar{T}}^\psi]$ is positive definite.

Here again, notice that $\boldsymbol{\mathcal{P}}_i^R$ is not $\lim_{T\to\infty} \frac{T}{\bar{T}} \boldsymbol{\Pi}_{\bar{T}}^{\psi\prime} \mathbb{E}[\boldsymbol{\varphi}_{\bar{T}}^i \boldsymbol{\varphi}_{\bar{T}}^{i\prime}] \boldsymbol{\Pi}_{\bar{T}}^\psi$ since eigenvectors are random; so we must assume its existence. If eigenvectors were not random, its existence would follow from of Lemma 18 in the Appendix, for all $T \in \mathbb{N}$. Moreover, $\boldsymbol{\mathcal{P}}_i^R$ is positive definite since it is a Toeplitz matrix containing all autocovariances of the ith idiosyncratic component. A similar assumption is made also in Bai (2003, Assumption F4); it is satisfied, for example, by all α-mixing processes.

The following then can be proved along the same lines as Theorem 1.

Theorem 2. *Denote by $\widehat{\boldsymbol{W}}^z$ a $q \times q$ diagonal matrix, depending on n and T, with diagonal entries ± 1. Then, for any $i = 1, \ldots, n$ and any $\bar{T} \leq T$ such that $\frac{1}{\bar{T}} + \frac{\bar{T}}{T} \to 0$, as $n, T \to \infty$,*

(i) under Assumptions (S), (S(e)), and (K),

$$\left\| \widehat{\mathbf{R}}_i - \widehat{\boldsymbol{W}}^z \left(\boldsymbol{\Sigma}^R\right)^{-1/2} \mathbf{R}_i \right\| = O_{\mathrm{P}}\left(\max\left(\frac{1}{\sqrt{\bar{T}}}, \zeta_{nT} \right) \right);$$

(ii) under Assumptions (S), (S(e)), (S(f2)), and (K), with $\bar{T}$ such that

$$\frac{1}{\bar{T}} + \sqrt{\bar{T}}\, \zeta_{nT} \to 0, \qquad \text{as } n, T \to \infty, \tag{52}$$

and $\widehat{\boldsymbol{W}}^z \longrightarrow_{\mathrm{P}} \mathbf{W}^R$, letting $\boldsymbol{\mathcal{M}}_R := \operatorname{plim}_{T\to\infty} \frac{T}{\bar{T}} \boldsymbol{\Pi}_{\bar{T}}^{\psi\prime} \boldsymbol{\Pi}_{\bar{T}}^\psi$, and $\boldsymbol{\mathcal{L}}_R := \operatorname{plim}_{T\to\infty} \frac{\boldsymbol{L}_T^\psi}{T}$,

$$\sqrt{\bar{T}} \left(\widehat{\mathbf{R}}_i - \widehat{\boldsymbol{W}}^z \left(\boldsymbol{\Sigma}^R\right)^{-1/2} \mathbf{R}_i \right) \longrightarrow_d \mathcal{N}\left(\mathbf{0}_q, \mathbf{W}^R \left(\boldsymbol{\mathcal{L}}^R\right)^{-1/2} \left(\boldsymbol{\mathcal{M}}^R\right)^{-1} \boldsymbol{\mathcal{P}}_i^R \left(\boldsymbol{\mathcal{M}}^R\right)^{-1} \left(\boldsymbol{\mathcal{L}}^R\right)^{-1/2} \mathbf{W}^R \right)$$

[7] Although $\widehat{\boldsymbol{\mathcal{P}}}_t^u$ does not depend on t we keep the index t to highlight the possibility of considering estimators of the asymptotic covariance that allow for heteroskedasticity.

where $\boldsymbol{\mathcal{P}}_i^R$ is defined in in part (f2) of Assumption (S).

Remark 8. In terms of rates of convergence (part (i) of Theorem 2), in the case $\bar{T} = T$, we obtain the rate ζ_{nT}^{-1} as already derived in Forni et al. (2017, Proposition 10). In particular, our estimator of $\mathbf{R}_i$ converges at rate $\sqrt{T/B_T}$ when $T/(B_T n) \downarrow 0$ whereas, when $(nB_T)/T \downarrow 0$, we achieve the rate of convergence $\sqrt{n}$. Both rates are slower than the rate $\sqrt{T}$ or n depending on whether $\sqrt{T}/n \downarrow 0$ or $n/\sqrt{T} \downarrow 0$ in Bai (2003) estimator, . This is because we need to estimate a spectral density before running PCA. The best rates we can achieve are $T^{2/5}$ if we choose a quadratic kernel, i.e. $\kappa = 2$, with optimal bandwidth $B_T = T^{1/5}$, or $T^{1/3}$ if we choose a Bartlett kernel, i.e. $\kappa = 1$, with optimal bandwidth $B_T = T^{1/3}$.

Remark 9. Condition (52) imposes only a marginally slower rate than ζ_{nT}, which is the consistency rate when $\bar{T} = T$. For example we can assume $\bar{T}$ of the form $\bar{T} = \zeta_{nT}^{-2} L^{-1}(\zeta_{nT}^{-1})$ for some slowly varying at infinity function $L(\cdot)$ (this implies that $\bar{T} \simeq T$ is a viable choice (neglecting the bandwidth dependence)). Note that $\bar{T}$ then depends on both n and T. In fact, by inspection of the proof of part (i), we can see that consistency holds with a faster rate and we could relax (52) to $\frac{\bar{T}}{\sqrt{T}} \zeta_{nT} \to 0$. However, since we still need to impose (52) for deriving the properties of the common component, we stick with it also in Theorem 2.

Remark 10. When (52) does not hold, Theorem 2 states that, as $n, T \to \infty$,

$$\sqrt{\bar{T}} \left(\widehat{\mathbf{R}}_i - \widehat{\boldsymbol{W}}^z \left(\boldsymbol{\Sigma}^R \right)^{-1/2} \mathbf{R}_i \right) - \widehat{\boldsymbol{W}}^z \boldsymbol{\mathcal{X}}_t \longrightarrow_{\mathrm{P}} \mathbf{0}_q,$$

for some $\boldsymbol{\mathcal{X}}_t \sim \mathcal{N}(\mathbf{0}_q, \left(\boldsymbol{\mathcal{L}}^R\right)^{-1/2} \left(\boldsymbol{\mathcal{M}}^R\right)^{-1} \boldsymbol{\mathcal{P}}_i^R \left(\boldsymbol{\mathcal{M}}^R\right)^{-1} \left(\boldsymbol{\mathcal{L}}^R\right)^{-1/2})$.

Remark 11. A consistent estimator of the asymptotic covariance matrix of $\sqrt{\bar{T}}(\widehat{\mathbf{R}}_i - \widehat{\mathbf{W}}_T^z (\boldsymbol{\Sigma}^R)^{-1/2} \mathbf{R}_i)$ is

$$\left(\frac{\widehat{\boldsymbol{L}}_T^z}{T} \right)^{-1/2} \left(\frac{T}{\bar{T}} \widehat{\mathbf{P}}_{\bar{T}}^{z\prime} \widehat{\mathbf{P}}_{\bar{T}}^z \right)^{-1} \widehat{\boldsymbol{\mathcal{P}}}_i^R \left(\frac{T}{\bar{T}} \widehat{\mathbf{P}}_{\bar{T}}^{z\prime} \widehat{\mathbf{P}}_{\bar{T}}^z \right)^{-1} \left(\frac{\widehat{\boldsymbol{L}}_T^z}{T} \right)^{-1/2},$$

where $\widehat{\boldsymbol{\mathcal{P}}}_i^R$ is a consistent estimator of $\boldsymbol{\mathcal{P}}_i^R$. If we assume that $\{\phi_{it}\}$ is not autocorrelated, we can use

$$\widehat{\boldsymbol{\mathcal{P}}}_i^R := \sum_{t=1}^{T} \widehat{\boldsymbol{\pi}}_t^z \widehat{\boldsymbol{\pi}}_t^{z\prime} \widehat{\phi}_{it}^2, \tag{53}$$

where $\widehat{\phi}_{it} = \widehat{z}_{it} - \check{\mathbf{u}}_t' \widehat{\mathbf{R}}_i$ and $\widehat{\boldsymbol{\pi}}_t^{z\prime}$ is the tth row of $\widehat{\mathbf{\Pi}}_T^z$. To address idiosyncratic autocorrelation, a natural choice is the usual HAC estimator used also in Bai (2003, Section 5(b)).

5.4 Asymptotics for the static common component $\widehat{\psi}_{it}$

Using the estimates of the loadings $\mathbf{R}_i$ and the common shocks $\mathbf{u}_t$ developed in the previous sections, one can construct estimates of the static common components ψ_{it}. Several approaches are possible. Both $\check{\mathbf{R}}_i' \widehat{\mathbf{u}}_t$ and $\widehat{\mathbf{R}}_i' \check{\mathbf{u}}_t$, in fact, are consistent estimators of ψ_{it}. In principle, one can also consider $\widehat{\mathbf{R}}_i' \widehat{\mathbf{u}}_t$, although identification is not warranted due to the presence of the product of the two rotation matrices $\widehat{\mathbf{W}}^z(\mathbf{\Gamma}^u)^{-1/2}$ and $\widehat{\boldsymbol{W}}^z(\mathbf{\Sigma}^R)^{-1/2}$ which are not necessarily identical. In contrast with this, $\check{\mathbf{R}}_i'\widehat{\mathbf{u}}_t$ and $\widehat{\mathbf{R}}_i'\check{\mathbf{u}}_t$ both involve $(\widehat{\mathbf{W}}^z)^2$ hence a product $\mathbf{I}_q$. This is why the estimators $\widehat{\psi}_{it}$ we are proposing for ψ_{it} are of the form (44). These estimators could achieve an asymptotic efficiency gain with respect to both $\check{\mathbf{R}}_i'\widehat{\mathbf{u}}_t$ and $\widehat{\mathbf{R}}_i'\check{\mathbf{u}}_t$. Moreover, by setting $\bar{n} = \bar{T} =: \bar{h}$, they avoid the slight technical difficulty of combining estimators with different rates of convergence (see Bai, 2003, proof of Theorem 3).

Theorem 3. *Set $\bar{n} = \bar{T} = \bar{h}$. Then, for any $i = 1, \dots, n$ and $t = 1, \dots, T$, and any $\bar{h} < \min(n, T)$ such that $\frac{1}{\bar{h}} + \frac{\bar{h}}{\min(n,T)} \to 0$, as $n, T \to \infty$,*

(i) under Assumptions (S), (S(e)), and (K),

$$\left\| \widehat{\psi}_{it} - \psi_{it} \right\| = O_{\mathrm{P}} \left(\max \left(\frac{1}{\sqrt{\bar{h}}}, \zeta_{nT} \right) \right);$$

(ii) if also Assumptions (S(f1)) and (S(F2)) hold, and if $\bar{h}$ is such that

$$\frac{1}{\bar{h}} + \sqrt{\bar{h}}\zeta_{nT} \to 0, \qquad \text{as } n, T \to \infty, \tag{54}$$

and $\widehat{\mathbf{W}}^z \to_{\mathrm{P}} \mathbf{W}^u$, $\widehat{\boldsymbol{W}}^z \to_{\mathrm{P}} \mathbf{W}^R$, then

$$\sqrt{\bar{h}} \left(\widehat{\psi}_{it} - \psi_{it} \right) \longrightarrow_d \mathcal{N} \left(0, \boldsymbol{\omega}' \begin{pmatrix} V_{it}^u & C_{it} \\ C_{it} & V_{it}^R \end{pmatrix} \boldsymbol{\omega} \right),$$

where $\boldsymbol{\omega} := \lim_{nT \to \infty} \begin{pmatrix} \omega_{nT} \\ 1 - \omega_{nT} \end{pmatrix}$,

$$\begin{aligned}
V_{it}^u &:= \mathbf{R}_i' (\mathbf{\Gamma}^u)^{1/2} \mathbf{W}^u (\boldsymbol{\mathcal{L}}^u)^{-1/2} (\boldsymbol{\mathcal{M}}^u)^{-1} \boldsymbol{\mathcal{P}}_t^u (\boldsymbol{\mathcal{M}}^u)^{-1} (\boldsymbol{\mathcal{L}}^u)^{-1/2} \mathbf{W}^u (\mathbf{\Gamma}^u)^{1/2} \mathbf{R}_i, \\
V_{it}^R &:= \mathbf{u}_t' \left(\mathbf{\Sigma}^R\right)^{1/2} \mathbf{W}^R \left(\boldsymbol{\mathcal{L}}^R\right)^{-1/2} \left(\boldsymbol{\mathcal{M}}^R\right)^{-1} \boldsymbol{\mathcal{P}}_i^R \left(\boldsymbol{\mathcal{M}}^R\right)^{-1} \left(\boldsymbol{\mathcal{L}}^R\right)^{-1/2} \mathbf{W}^R \left(\mathbf{\Sigma}^R\right)^{1/2} \mathbf{u}_t, \\
C_{it} &:= \mathbf{R}_i' (\mathbf{\Gamma}^u)^{1/2} \mathbf{W}^u (\boldsymbol{\mathcal{L}}^u)^{-1/2} (\boldsymbol{\mathcal{M}}^u)^{-1} \mathbf{\Omega}_{it} \left(\boldsymbol{\mathcal{M}}^R\right)^{-1} \left(\boldsymbol{\mathcal{L}}^R\right)^{-1/2} \mathbf{W}^R \left(\mathbf{\Sigma}^R\right)^{1/2} \mathbf{u}_t,
\end{aligned}$$

with $\boldsymbol{\mathcal{P}}_t^u$, $\boldsymbol{\mathcal{M}}^u$, and $\boldsymbol{\mathcal{L}}^u$ as defined in Theorem 1, $\boldsymbol{\mathcal{P}}_i^R$, $\boldsymbol{\mathcal{M}}^R$, and $\boldsymbol{\mathcal{L}}^R$ as defined in Theorem 2, and $\mathbf{\Omega}_{it} := \lim_{n,T\to\infty} \left(\frac{\sqrt{nT}}{\bar{h}} \right) \mathbb{E}[\mathbf{P}_{\bar{n}}^{\psi\prime} \boldsymbol{\phi}_{\bar{n}t} \boldsymbol{\varphi}_{\bar{T}}^{i\prime} \mathbf{\Pi}_{\bar{T}}^{\psi}]$.

Remark 12. Notice that, consistently with (26), we can always write (see the proof of Theorem 1)

$$\mathbf{z}_{nt} = \mathbf{P}_n^{\psi} \left(\mathbf{\Lambda}_n^{\psi} \right)^{1/2} (\mathbf{\Gamma}^u)^{-1/2} \mathbf{u}_t + \boldsymbol{\phi}_{nt}, \qquad t = 1, \dots, T,$$

which implies $\mathbf{R}_i' = \mathbf{p}_i^{\psi\prime} (\mathbf{\Lambda}_n^{\psi})^{1/2} (\mathbf{\Gamma}^u)^{-1/2}$. Moreover, by definition $\check{\mathbf{R}}_i' = \widehat{\mathbf{p}}_i^{z\prime} (\widehat{\mathbf{\Lambda}}_n^z)^{1/2}$ which, as shown in the proof of Theorem 3, is a consistent estimator of $\mathbf{R}_i' (\mathbf{\Gamma}^u)^{1/2} \mathbf{W}^u$. Therefore, a natural estimator of V_{it}^u is, with $\widehat{\boldsymbol{\mathcal{P}}}_t^u$ defined in (49),

$$\begin{aligned}
\widehat{V}_{it}^u &:= \check{\mathbf{R}}_i' \left(\frac{\widehat{\mathbf{\Lambda}}_n^z}{n} \right)^{-1/2} \left(\frac{n}{\bar{n}} \widehat{\mathbf{P}}_{\bar{n}}^{z\prime} \widehat{\mathbf{P}}_{\bar{n}}^z \right)^{-1} \widehat{\boldsymbol{\mathcal{P}}}_t^u \left(\frac{n}{\bar{n}} \widehat{\mathbf{P}}_{\bar{n}}^{z\prime} \widehat{\mathbf{P}}_{\bar{n}}^z \right)^{-1} \left(\frac{\widehat{\mathbf{\Lambda}}_n^z}{n} \right)^{-1/2} \check{\mathbf{R}}_i \\
&= n\, \widehat{\mathbf{p}}_i^{z\prime} \left(\frac{n}{\bar{n}} \widehat{\mathbf{P}}_{\bar{n}}^{z\prime} \widehat{\mathbf{P}}_{\bar{n}}^z \right)^{-1} \widehat{\boldsymbol{\mathcal{P}}}_t^u \left(\frac{n}{\bar{n}} \widehat{\mathbf{P}}_{\bar{n}}^{z\prime} \widehat{\mathbf{P}}_{\bar{n}}^z \right)^{-1} \widehat{\mathbf{p}}_i^z,
\end{aligned}$$

which does not depend on the unknown matrix $\mathbf{\Gamma}^u$ nor on the sign matrix $\mathbf{W}^u$. Similarly, we can always write

$$\boldsymbol{z}_T^i = \mathbf{\Pi}_T^\psi \left(\boldsymbol{L}_T^\psi\right)^{1/2} \left(\mathbf{\Sigma}^R\right)^{-1/2} \mathbf{R}_i + \boldsymbol{\varphi}_T^i, \qquad i = 1, \dots, n,$$

which implies $\mathbf{u}_t' = \boldsymbol{\pi}_t^{\psi\prime}(\boldsymbol{L}_T^\psi)^{1/2}(\mathbf{\Sigma}^R)^{-1/2}$ ($\widehat{\boldsymbol{\pi}}_t^{\psi\prime}$ the tth row of $\widehat{\mathbf{\Pi}}_T^\psi$). Moreover, by definition, we have that $\check{\mathbf{u}}_t' = \widehat{\boldsymbol{\pi}}_t^{z\prime}(\widehat{\boldsymbol{L}}_T^z)^{1/2}$, which, as shown in the proof of Theorem 3, is a consistent estimator of $\mathbf{u}_t'(\mathbf{\Sigma}^R)^{1/2}\mathbf{W}^R$. Therefore, a natural estimator of V_{it}^R is, with $\widehat{\boldsymbol{\mathcal{P}}}_i^R$ defined in (53),

$$\begin{aligned}
\widehat{V}_{it}^R &:= \check{\mathbf{u}}_t' \left(\frac{\widehat{\boldsymbol{L}}_T^z}{T}\right)^{-1/2} \left(\frac{T}{\bar{T}}\widehat{\mathbf{P}}_{\bar{T}}^{z\prime}\widehat{\mathbf{P}}_{\bar{T}}^z\right)^{-1} \widehat{\boldsymbol{\mathcal{P}}}_i^R \left(\frac{T}{\bar{T}}\widehat{\mathbf{P}}_{\bar{T}}^{z\prime}\widehat{\mathbf{P}}_{\bar{T}}^z\right)^{-1} \left(\frac{\widehat{\boldsymbol{L}}_T^z}{T}\right)^{-1/2} \check{\mathbf{u}}_t \\
&= T\, \widehat{\boldsymbol{\pi}}_t^{z\prime} \left(\frac{T}{\bar{T}}\widehat{\mathbf{P}}_{\bar{T}}^{z\prime}\widehat{\mathbf{P}}_{\bar{T}}^z\right)^{-1} \widehat{\boldsymbol{\mathcal{P}}}_i^R \left(\frac{T}{\bar{T}}\widehat{\mathbf{P}}_{\bar{T}}^{z\prime}\widehat{\mathbf{P}}_{\bar{T}}^z\right)^{-1} \widehat{\boldsymbol{\pi}}_t^z,
\end{aligned}$$

which does not depend on the unknown matrix $\mathbf{\Sigma}^R$ nor on the sign matrix $\mathbf{W}^R$.

5.5 Asymptotics for the dynamic common component $\widehat{\chi}_{it}$

Let $\mathbf{C}_n(L) := [\mathbf{A}_n(L)]^{-1}$ and notice that since $\mathbf{A}_n(L)$ is block-diagonal, then also $\mathbf{C}_n(L)$ is block-diagonal. Denote as $\mathcal{I}_s := \{\ell \,|\, \ell = (s-1)(q+1)+1, \dots, s(q+1)\}$, the set of integers indicating the series belonging to block s, where $s = 1, \dots, m$. Then, given a cross-sectional unit $i \in \mathcal{I}_s$ of $\mathbf{A}_n(L)$, its dynamic common component χ_{it} is defined as (see (17))

$$\begin{aligned}
\chi_{it} &= \mathbf{C}_n(L)\boldsymbol{\mathcal{R}}_n\mathbf{u}_t \\
&= \sum_{k=0}^{\infty}\sum_{j_s=1}^{q+1} c_{i,j_s,k}\mathbf{R}_{j_s}'\mathbf{u}_{t-k} = \sum_{k=0}^{\infty}\sum_{j_s=1}^{q+1} c_{i,j_s,k}\psi_{j_s,t-k}, \quad i \in \mathcal{I}_s,\ s = 1, \dots, m,\ t \in \mathbb{Z},
\end{aligned} \tag{55}$$

where $c_{i,j_s,k}$ is the (i, j_s)th entry of $\mathbf{C}_n(L)$ and j_s indicates the jth column of block s of $\mathbf{C}_n(L)$, i.e., the jth element of $\mathcal{I}_s$. Our estimator of χ_{it} is then

$$\widehat{\chi}_{it} = \sum_{k=0}^{K}\sum_{j_s=1}^{q+1} \widehat{c}_{i,j_s,k}\widehat{\psi}_{j_s,t-k}, \quad i \in \mathcal{I}_s,\ s = 1, \dots, m,\ t = K+1, \dots, T, \tag{56}$$

where K is a finite integer, $\widehat{c}_{i,j_s,k}$ is the (i, j_s)th entry of $\widehat{\mathbf{C}}_n(L) := [\widehat{\mathbf{A}}_n(L)]^{-1}$, and $\widehat{\psi}_{j_s,t-k}$ is the estimator of the static common component defined in (44). Notice that in (56) we sum only over a finite number of lags K, since the observed sample has always finite length. Moreover, since, by stationarity, the coefficients of $\mathbf{C}_n(L)$ are decaying geometrically, K can always be chosen in such a way that the contribution of the lags $k > K$ is negligible.

Theorem 4. *Set $\bar{n} = \bar{T} = \bar{h}$. Then, for any $s = 1, \dots, m$, $i \in \mathcal{I}_s$, and $t = 1, \dots, T$ and for any $\bar{h} \le \min(n, T)$, such that $\frac{1}{\bar{h}} + \frac{\bar{h}}{\min(n,T)} \to 0$, as $n, T \to \infty$*

(i) under Assumptions (S), (S(e)), and (K),

$$\|\widehat{\chi}_{it} - \chi_{it}\| = O_{\mathrm{P}}\left(\max\left(\frac{1}{\sqrt{\bar{h}}}, \zeta_{nT}\right)\right);$$

(ii) *if also Assumptions (S(f1)) and (S(f2)) hold, if* $\widehat{\mathbf{W}}^z \to_{\mathrm{P}} \mathbf{W}^u$, $\widehat{\boldsymbol{W}}^z \to_{\mathrm{P}} \mathbf{W}^R$ *and* $\bar{h}$ *is such that*

$$\frac{1}{\bar{h}} + \sqrt{\bar{h}}\zeta_{nT} \to 0, \qquad \text{as } n, T \to \infty, \tag{57}$$

then

$$\sqrt{\bar{h}}\left(\widehat{\chi}_{it} - \chi_{it}\right) \longrightarrow_d \mathcal{N}\left(0, \boldsymbol{\omega}'\begin{pmatrix} W_{it}^u & G_{it} \\ G_{it} & W_{it}^R \end{pmatrix}\boldsymbol{\omega}\right),$$

where $\boldsymbol{\omega} := \lim_{n,T\to\infty}\begin{pmatrix}\omega_{nT} \\ 1-\omega_{nT}\end{pmatrix}$ *and, letting* $\mathcal{I}_s = \{i_1, \ldots, i_{q+1}\}$, *for a given finite integer lag* K,

$$W_{it}^u := \boldsymbol{\iota}_{q+1}'\left\{\boldsymbol{\mathcal{C}}_i \odot \left[\boldsymbol{\iota}_{K+1}' \otimes \begin{pmatrix} \mathbf{R}_{i_1}'(\boldsymbol{\Gamma}^u)^{1/2} \\ \vdots \\ \mathbf{R}_{i_{q+1}}'(\boldsymbol{\Gamma}^u)^{1/2}\end{pmatrix}\right]\right\}\boldsymbol{\mathcal{V}}_{t\ldots t-K}^u\left\{\boldsymbol{\mathcal{C}}_i \odot \left[\boldsymbol{\iota}_{K+1}' \otimes \begin{pmatrix} \mathbf{R}_{i_1}'(\boldsymbol{\Gamma}^u)^{1/2} \\ \vdots \\ \mathbf{R}_{i_{q+1}}'(\boldsymbol{\Gamma}^u)^{1/2}\end{pmatrix}\right]\right\}'\boldsymbol{\iota}_{q+1},$$

$$W_{it}^R := \boldsymbol{\iota}_{K+1}'\left\{\boldsymbol{\mathcal{D}}_i \odot \left[\boldsymbol{\iota}_{q+1}' \otimes \begin{pmatrix} \mathbf{u}_t'(\boldsymbol{\Sigma}^R)^{1/2} \\ \vdots \\ \mathbf{u}_{t-K}'(\boldsymbol{\Sigma}^R)^{1/2}\end{pmatrix}\right]\right\}\boldsymbol{\mathcal{V}}_{i_1\ldots i_{q+1}}^R\left\{\boldsymbol{\mathcal{D}}_i \odot \left[\boldsymbol{\iota}_{q+1}' \otimes \begin{pmatrix} \mathbf{u}_t'(\boldsymbol{\Sigma}^R)^{1/2} \\ \vdots \\ \mathbf{u}_{t-K}'(\boldsymbol{\Sigma}^R)^{1/2}\end{pmatrix}\right]\right\}'\boldsymbol{\iota}_{K+1},$$

$$G_{it} := \boldsymbol{\iota}_{q+1}'\left\{\boldsymbol{\mathcal{C}}_i \odot \left[\boldsymbol{\iota}_{K+1}' \otimes \begin{pmatrix} \mathbf{R}_{i_1}'(\boldsymbol{\Gamma}^u)^{1/2} \\ \vdots \\ \mathbf{R}_{i_{q+1}}'(\boldsymbol{\Gamma}^u)^{1/2}\end{pmatrix}\right]\right\}\boldsymbol{\mathcal{O}}_{\substack{i_1\ldots i_{q+1}\\ t\ldots t-K}}\left\{\boldsymbol{\mathcal{D}}_i \odot \left[\boldsymbol{\iota}_{q+1}' \otimes \begin{pmatrix} \mathbf{u}_t'(\boldsymbol{\Sigma}^R)^{1/2} \\ \vdots \\ \mathbf{u}_{t-K}'(\boldsymbol{\Sigma}^R)^{1/2}\end{pmatrix}\right]\right\}'\boldsymbol{\iota}_{K+1},$$

with $\otimes$ *and* $\odot$ *the Kronecker and Hadamard products, respectively,* $\boldsymbol{\iota}_{K+1}$ *a* $(K+1)$*-dimensional vector of ones, and* $\boldsymbol{\iota}_{q+1}$ *a* $(q+1)$*-dimensional vector of ones,*

$$\boldsymbol{\mathcal{C}}_i := \begin{pmatrix} \boldsymbol{\iota}_q' \otimes c_{i,1,0} & \ldots & \boldsymbol{\iota}_q' \otimes c_{i,1,K} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\iota}_q' \otimes c_{i,q+1,0} & \ldots & \boldsymbol{\iota}_q' \otimes c_{i,q+1,K}\end{pmatrix}, \qquad \boldsymbol{\mathcal{D}}_i := \begin{pmatrix} \boldsymbol{\iota}_q' \otimes c_{i,1,0} & \ldots & \boldsymbol{\iota}_q' \otimes c_{i,q+1,0} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\iota}_q' \otimes c_{i,1,K} & \ldots & \boldsymbol{\iota}_q' \otimes c_{i,q+1,K}\end{pmatrix},$$

$$\begin{aligned}
\boldsymbol{\mathcal{V}}_{t\ldots t-K}^u &:= \left\{\mathbf{I}_{K+1} \otimes \left[\mathbf{W}^u\left(\boldsymbol{\mathcal{L}}^u\right)^{-1/2}\left(\boldsymbol{\mathcal{M}}^u\right)^{-1}\right]\right\}\boldsymbol{\mathcal{P}}_{t\ldots t-K}^u\left\{\mathbf{I}_{K+1} \otimes \left[\left(\boldsymbol{\mathcal{M}}^u\right)^{-1}\left(\boldsymbol{\mathcal{L}}^u\right)^{-1/2}\mathbf{W}^u\right]\right\}, \\
\boldsymbol{\mathcal{V}}_{i_1\ldots i_{q+1}}^R &:= \left\{\mathbf{I}_{q+1} \otimes \left[\mathbf{W}^R\left(\boldsymbol{\mathcal{L}}^R\right)^{-1/2}\left(\boldsymbol{\mathcal{M}}^R\right)^{-1}\right]\right\}\boldsymbol{\mathcal{P}}_{i_1\ldots i_{q+1}}^R\left\{\mathbf{I}_{q+1} \otimes \left[\left(\boldsymbol{\mathcal{M}}^R\right)^{-1}\left(\boldsymbol{\mathcal{L}}^R\right)^{-1/2}\mathbf{W}^R\right]\right\}, \\
\boldsymbol{\mathcal{O}}_{\substack{i_1\ldots i_{q+1}\\ t\ldots t-K}} &:= \left\{\mathbf{I}_{K+1} \otimes \left[\mathbf{W}^u\left(\boldsymbol{\mathcal{L}}^u\right)^{-1/2}\left(\boldsymbol{\mathcal{M}}^u\right)^{-1}\right]\right\}\boldsymbol{\Omega}_{\substack{i_1\ldots i_{q+1}\\ t\ldots t-K}}\left\{\mathbf{I}_{q+1} \otimes \left[\left(\boldsymbol{\mathcal{M}}^R\right)^{-1}\left(\boldsymbol{\mathcal{L}}^R\right)^{-1/2}\mathbf{W}^R\right]\right\}, \\
\boldsymbol{\mathcal{P}}_{t\ldots t-K}^u &:= \lim_{n\to\infty}\frac{n}{\bar{n}}\mathbb{E}\left[\left\{\mathbf{I}_{K+1} \otimes \mathbf{P}_{\bar{n}}^{\psi\prime}\right\}\begin{pmatrix}\boldsymbol{\phi}_{\bar{n}t} \\ \vdots \\ \boldsymbol{\phi}_{\bar{n}t-K}\end{pmatrix}\begin{pmatrix}\boldsymbol{\phi}_{\bar{n}t} \\ \vdots \\ \boldsymbol{\phi}_{\bar{n}t-K}\end{pmatrix}'\left\{\mathbf{I}_{K+1} \otimes \mathbf{P}_{\bar{n}}^{\psi\prime}\right\}'\right], \\
\boldsymbol{\mathcal{P}}_{i_1\ldots i_{q+1}}^R &:= \lim_{T\to\infty}\frac{T}{\bar{T}}\mathbb{E}\left[\left\{\mathbf{I}_{q+1} \otimes \boldsymbol{\Pi}_{\bar{T}}^{\psi\prime}\right\}\begin{pmatrix}\boldsymbol{\varphi}_{\bar{T}}^{i_1} \\ \vdots \\ \boldsymbol{\varphi}_{\bar{T}}^{i_{q+1}}\end{pmatrix}\begin{pmatrix}\boldsymbol{\varphi}_{\bar{T}}^{i_1} \\ \vdots \\ \boldsymbol{\varphi}_{\bar{T}}^{i_{q+1}}\end{pmatrix}'\left\{\mathbf{I}_{q+1} \otimes \boldsymbol{\Pi}_{\bar{T}}^{\psi\prime}\right\}'\right],
\end{aligned}$$

$\boldsymbol{\mathcal{M}}^u$ *and* $\boldsymbol{\mathcal{L}}^u$ *as defined in Theorem 1,* $\boldsymbol{\mathcal{M}}^R$ *and* $\boldsymbol{\mathcal{L}}^R$ *as defined in Theorem 2.*

Remark 13. To appreciate the formulas given in Theorem 4, let us consider a simple example. Let $q = 1$, $K = 1$, $s = 1$, so that $i = 1, 2$ and $j_s = 1, 2$. Then, if $n \ll T$ so that $\omega_{nT} \simeq 1$, from the proof of Theorem 4 we have

$$\sqrt{\bar{h}}(\widehat{\chi}_{1t} - \chi_{1t}) = \sqrt{\bar{h}} \sum_{k=0}^{1} \sum_{j_s=1}^{2} \left\{ c_{1,j_s,k} \mathbf{R}'_{j_s} \left(\widehat{\mathbf{u}}_{t-k} - \mathbf{u}_{t-k} \right) \right\} + o_{\mathrm{P}}(1), \tag{58}$$

which has asymptotic variance (notice that $\mathbf{u}_t$ and $\mathbf{R}_i$ are now scalars)

$$\begin{aligned} W^u_{it} = \lim_{n,T\to\infty} \bar{n} \Big(& c^2_{1,1,0} R^2_1 \mathrm{Var}(\widehat{u}_t - u_t) + c^2_{1,2,0} R^2_2 \mathrm{Var}(\widehat{u}_t - u_t) \\ & + c^2_{1,1,1} R^2_1 \mathrm{Var}(\widehat{u}_{t-1} - u_{t-1}) + c^2_{1,2,1} R^2_2 \mathrm{Var}(\widehat{u}_{t-1} - u_{t-1}) \\ & + 2c_{1,1,0}c_{1,2,0} R_1 R_2 \mathrm{Var}(\widehat{u}_t - u_t) + 2c_{1,1,1}c_{1,2,1} R_1 R_2 \mathrm{Var}(\widehat{u}_{t-1} - u_{t-1}) \\ & + 2c_{1,1,0}c_{1,1,1} R^2_1 \mathrm{Cov}\big((\widehat{u}_t - u_t), (\widehat{u}_{t-1} - u_{t-1})\big) \\ & + 2c_{1,2,0}c_{1,2,1} R^2_2 \mathrm{Cov}\big((\widehat{u}_t - u_t), (\widehat{u}_{t-1} - u_{t-1})\big) \\ & + 2c_{1,1,0}c_{1,2,1} R_1 R_2 \mathrm{Cov}\big((\widehat{u}_t - u_t), (\widehat{u}_{t-1} - u_{t-1})\big) \\ & + 2c_{1,2,0}c_{1,1,1} R_1 R_2 \mathrm{Cov}\big((\widehat{u}_t - u_t), (\widehat{u}_{t-1} - u_{t-1})\big) \Big). \end{aligned} \tag{59}$$

Similarly, if $T \ll n$ so that $\omega_{nT} \simeq 0$, from the same proof we have

$$\sqrt{\bar{h}}(\widehat{\chi}_{1t} - \chi_{1t}) = \sqrt{\bar{h}} \sum_{k=0}^{1} \sum_{j_s=1}^{2} \left\{ c_{1,j_s,k} \mathbf{u}'_{t-k} \left(\widehat{\mathbf{R}}_{j_s} - \mathbf{R}_{j_s} \right) \right\} + o_{\mathrm{P}}(1), \tag{60}$$

which has asymptotic variance

$$\begin{aligned} W^R_{it} = \lim_{n,T\to\infty} \bar{T} \Big(& c^2_{1,1,0} u^2_t \mathrm{Var}(\widehat{R}_1 - R_1) + c^2_{1,2,0} u^2_t \mathrm{Var}(\widehat{R}_2 - R_2) \\ & + c^2_{1,1,1} u^2_{t-1} \mathrm{Var}(\widehat{R}_1 - R_1) + c^2_{1,2,1} u^2_{t-1} \mathrm{Var}(\widehat{R}_2 - R_2) \\ & + 2c_{1,1,0}c_{1,1,1} u_t u_{t-1} \mathrm{Var}(\widehat{R}_1 - R_1) + 2c_{1,2,0}c_{1,2,1} u_t u_{t-1} \mathrm{Var}(\widehat{R}_2 - R_2) \\ & + 2c_{1,1,0}c_{1,2,0} u^2_t \mathrm{Cov}\big((\widehat{R}_1 - R_1), (\widehat{R}_2 - R_2)\big) \\ & + 2c_{1,1,1}c_{1,2,1} u^2_{t-1} \mathrm{Cov}\big((\widehat{R}_1 - R_1), (\widehat{R}_2 - R_2)\big) \\ & + 2c_{1,1,0}c_{1,2,1} u_t u_{t-1} \mathrm{Cov}\big((\widehat{R}_1 - R_1), (\widehat{R}_2 - R_2)\big) \\ & + 2c_{1,2,0}c_{1,1,1} u_t u_{t-1} \mathrm{Cov}\big((\widehat{R}_1 - R_1), (\widehat{R}_2 - R_2)\big) \Big). \end{aligned} \tag{61}$$

The variances in (59) and (61) are given in Theorems 1 and 2, respectively, and the covariances are easily derived along the same lines (for details, see the proof of Theorem 4). Clearly, if $n \simeq T$, we should also include covariances between the terms in (58) and those in (60), which contribute to the term G_{it} in the expression of the asymptotic variance.

6 Monte Carlo Simulations

We set $q = 1$ and we consider the data-generating process (a slightly modified version of the one used by Forni et al. (2017))

$$x_{it} = a_i \beta_i(L) u_t + \xi_{it}, \tag{62}$$

where $\beta_i(L) := (1 - \alpha_i L)^{-1} = (1 + \alpha_i L + \alpha_i^2 L^2 + \ldots)$. We generate u_t and ξ_{it} as i.i.d. standardized normal variables, a_i as normal variables with mean and variance both equal to one, and α_i as i.i.d. variables uniformly distributed over the interval $[0.1, 0.8]$. Finally, each idiosyncratic component ξ_{it} is rescaled so that the share of variance of x_{it} accounts for by ξ_{it} is $\frac{\theta}{(1-\theta)}$, with $\theta = 0.5$.

We simulate panels of size $n = T \in \{100, 200, 300, 400, 500\}$ and we consider a total of $B = 500$ Monte Carlo replications. At each replication b, we compute an estimator $\widehat{\chi}_{it}^{(b)}$ of the common component $\chi_{it}^{(b)}$ and its asymptotic variance. The main goal of this section is to check whether the asymptotic distributions derived in Theorem 4 is empirically confirmed. However, because we simulate panels with $n = T$, we estimate the dynamic common component as in (56), where the static common component is estimated with weight $\omega_{nT} = \frac{1}{2}$, which is slightly different from the estimator used by Forni et al. (2017), in which $\omega_{nT} = 1$. Therefore, we begin the simulation exercise by looking at the properties of the estimator in (56), setting $K = 2$, $p_s = 1$ for all s, and considering $\lfloor \sqrt{n} \rfloor$ permutations in building the m blocks.

Table 1 shows the Standardized Mean Squared Error (S-MSE)

$$\text{S-MSE} := \frac{\sum_{b=1}^{B} \sum_{i=1}^{n} \sum_{t=1}^{T} \left(\widehat{\chi}_{it}^{(b)} - \chi_{it}^{(b)} \right)^2}{\sum_{b=1}^{B} \sum_{i=1}^{n} \sum_{t=1}^{T} \left(\chi_{it}^{(b)} \right)^2} \tag{63}$$

of the estimator of χ_{it}. The results in Table 1 clearly show that the estimator in (56) works very well. As n and T increase, the S-MSE monotonically decreases, to the point that, for $n = T = 500$, the S-MSE is more than 70% lower than for $n = T = 100$.

Table 1: Standardize Mean Squared Errors
Common components

T	n	S-MSE
100	100	0.30
200	200	0.17
300	300	0.12
400	400	0.09
500	500	0.08

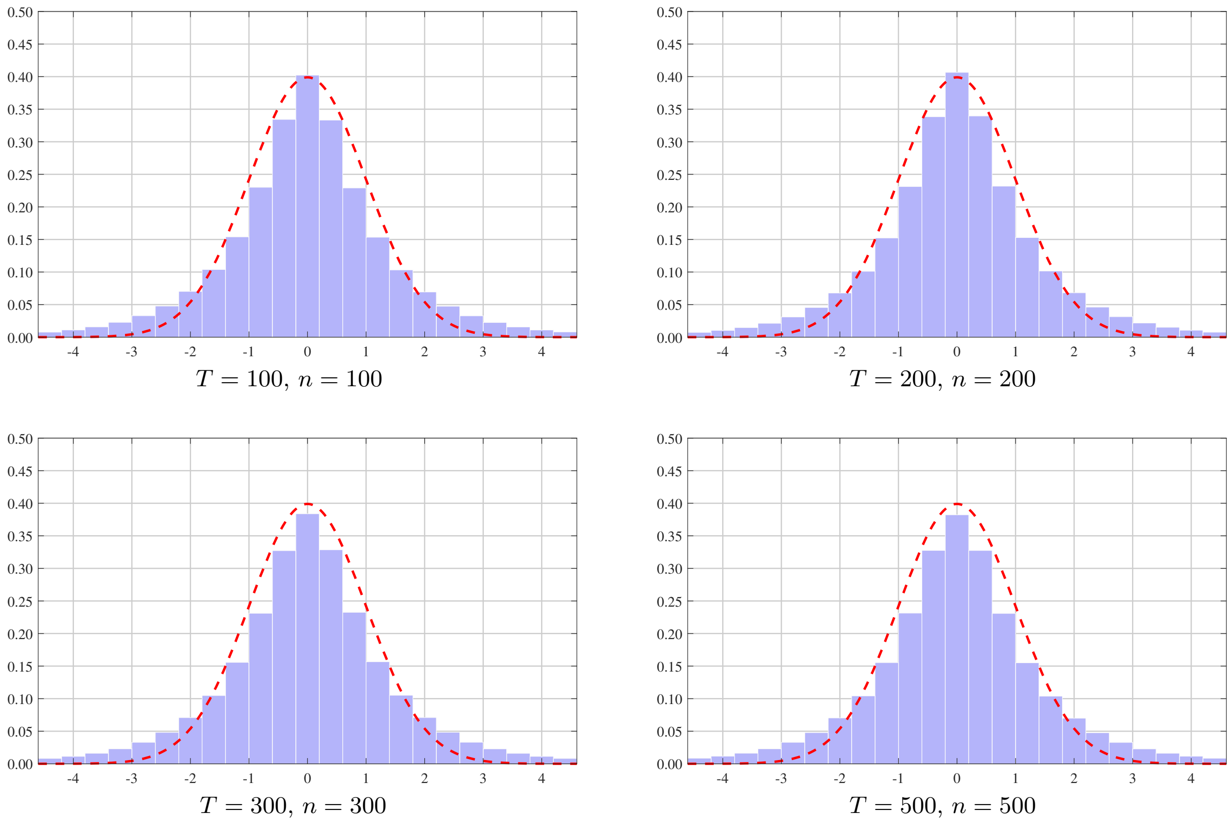
Next, we turn to the asymptotic distribution of the same common component. To this end, for each replication b and each (i, t), we compute

$$Z_{it}^{(b)} = \left(\frac{1}{4} \widehat{W}_{it}^{u} + \frac{1}{4} \widehat{W}_{it}^{R} \right)^{-1/2} \left(\widehat{\chi}_{it}^{(b)} - \chi_{it}^{(b)} \right) \tag{64}$$

which, according to Theorem 4, is asymptotically standard normal. Figure 1 shows, for four out of five of the (n, T) couples considered in Table 1, histograms of $\{Z_{it}^{(b)} : i = 1, \ldots, n, t = 1, \ldots, T, b = 1, \ldots B\}$. These histograms show that, while struggling a little bit in the tails, the empirical distribution of $Z_{it}^{(b)}$

is pretty close to the standard normal distribution (the red dashed line), well in line with Theorem 4. The fatter than Normal tails are the price we are paying for estimating $\mathbf{A}_n(L)$. That price is nil in the limit, but not for finite n and T.

Figure 1: HISTOGRAMS OF THE SIMULATED $Z_{it}^{(b)}$'S IN (64), FOR VARIOUS VALUES OF n AND T



7 Empirical Application: a "core" inflation indicator for the U.S.

Headline (or total) PCE price inflation, the measure chosen by the Federal Reserve to target its 2% target inflation objective, is highly volatile. Therefore, economists and policymakers have suggested alternative measures, which the literature calls "core" inflation indicators, to reduce the variance of the measured inflation, thus better distinguishing transitory from persistent movements. This Section uses the one-sided GDFM considered in this paper to estimate a new "core" inflation indicator for the U.S.[8]

Nowadays, the notion of core inflation in the U.S. is mainly associated with inflation excluding food and energy. The rationale for this indicator is that both food and energy prices are very volatile and often driven by idiosyncratic shocks (such as weather for food or OPEC decisions for energy). Thus, not only they do not provide a useful signal for inflation going forward, but also they are not controllable by the Federal Reserve (Blinder, 1997). However, the literature has proposed alternative ways of measuring core inflation, such as trimmed means and factor model-based estimates.[9]

[8]Altissimo et al. (2009) estimate a dynamic factor model, on disaggregated inflation data, that represents an over-simplified case of our setting, as it is assumed that the common components follow AR(1) processes with iid idiosyncratic components. This simplification allows to use a different estimation method. Unlike us, they estimate their model on euro area data.

[9]The rationale for the use of trimmed means as core inflation indexes is that a trimmed mean is a robust estimator of

The idea of considering (low-dimensional) factor models to estimate core inflation dates back to Bryan and Cecchetti (1993), while Cristadoro et al. (2005) and Amstad et al. (2017) more recently have used high-dimensional dynamic factor models, similar to the GDFM, with the same objective.[10] The rationale for considering factor models on the estimation of core inflation is that central banks are particularly interested in identifying movements in inflation that are driven by common (macroeconomic) shocks, so to avoid responding to changes in inflation due to sector-specific shocks, or, even worse, measurement error.

The dataset we are analyzing here consists of $n = 148$ PCE price inflation rates from January 1995 to December 2019 ($T = 300$).[11] Specifically, the dataset contains headline PCE price inflation, which is the target chosen by the Federal Reserve for their inflation stability objective, PCE price inflation excluding food and energy, and 146 disaggregated PCE prices. These 146 disaggregated PCE prices represent a particular disaggregation of PCE prices in which each disaggregated price index is constructed from a distinct data source. Indeed, most disaggregated PCE prices are measured using a corresponding index from the CPI, a few of them are measured using PPIs, and some others are imputed. As a result, some disaggregated PCE prices are based on the same CPI (or PPI) series, which means that some disaggregated PCE price indexes are identical (or nearly so). For the complete list of prices and detailed information on the data sources, we refer the reader to Luciani (2020).

The upper-left charts in Figures 2 and 3 show our estimate of core inflation based on the estimated common component of headline PCE price inflation, as defined in (56) (the red line), where the shaded area around our estimate is the $\pm$ one standard deviation confidence band, together with headline PCE price inflation (the black line).[12] Let P_t^h denote the headline PCE price index: Figure 2 shows month-over-month inflation in the PCE price index, i.e., $\pi_t^h = 100 \times (\frac{P_t^h}{P_{t-1}^h} - 1)$, while Figure 3 shows year-over-year inflation in the PCE price index, i.e., $\pi_t^h = 100 \times (\frac{P_t^h}{P_{t-12}^h} - 1)$. The former is the target of forecasters following inflation, and the latter is what policymaker care about and, consequently, what newspaper tends to comment on. Note that the model is estimated over month-over-month inflation rates, and then the estimated common component is computed by converting the month-over-month estimate into an year-over-year estimate.[13]

From simple visual inspection of the upper-left charts in Figures 2 and 3, we immediately see that our measure of core inflation is doing what it is supposed to do: tracking the trend of headline PCE price inflation while reducing the variance. Moreover, the confidence band seems to be quite well calibrated, as monthly headline PCE price inflation is outside the confidence band 27% of the time (as a reference, the $\pm$ one standard deviation interval of a standardized normal excludes 32% of the observations).

the location of a fat-tailed distribution, while a weighted mean (like the total inflation index, or the index excluding food and energy) typically is not.

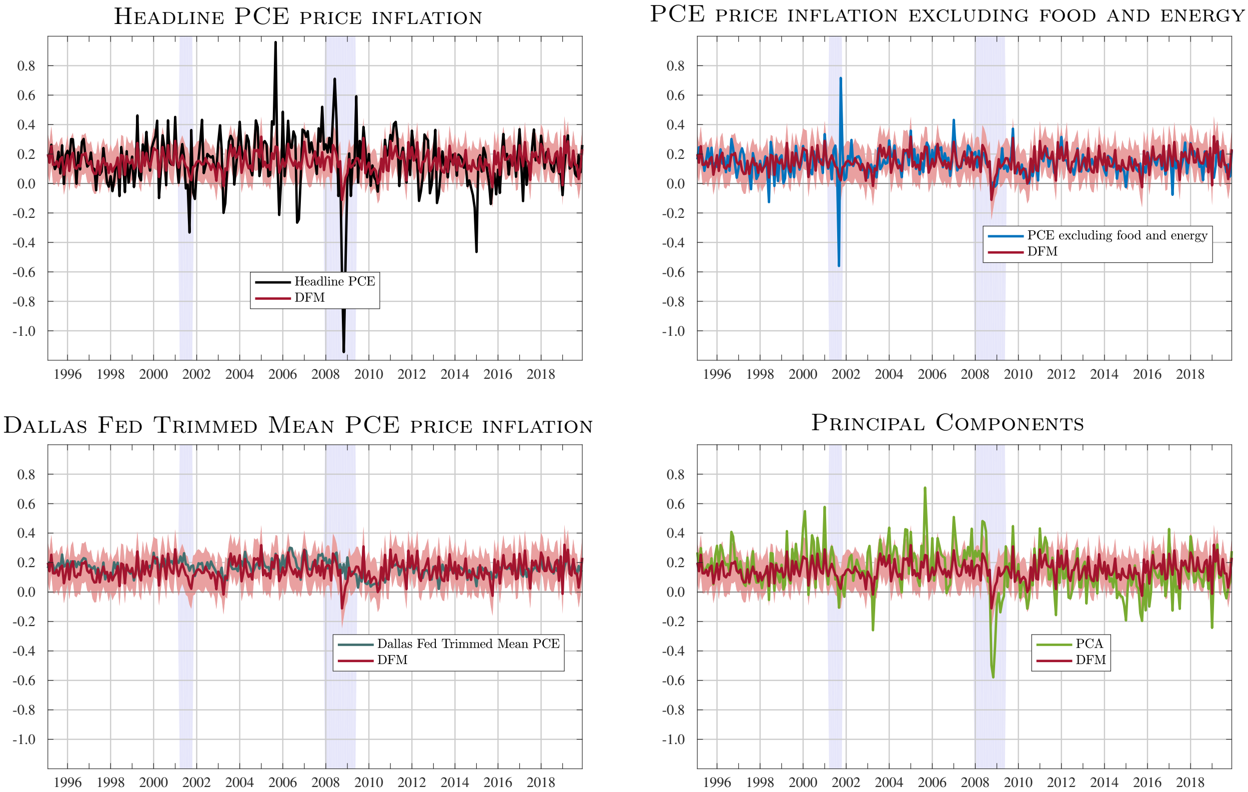
[10] Other papers have used high-dimensional factor models for constructing inflation indicators, though with a different goal. For example, Reis and Watson (2010) estimate an index of equiproportional changes in disaggregated PCE price inflation, while Luciani (2020) disentangles the effects of common versus idiosyncratic shocks in PCE price inflation excluding food and energy.

[11] Because $n \ll T$, to estimate $\boldsymbol{\mathcal{U}}_T$ and $\boldsymbol{\mathcal{R}}_n$, and therefore the common component and the asymptotic variances, we set $\omega_{nT} = 1$.

[12] The specification used in this section features one common shock, one lag in the VAR, and the number of autocovariances used to estimate the spectral density is set to $[T^{1/3}]$.

[13] As for the asymptotic variance, we took a shortcut for year-over-year estimates. Indeed we compute the variance for year-over-year estimates as $12\times$ the asymptotic variance over the month-over-month estimates. However, in doing so we are neglecting the autocorrelations, hence we can say that the confidence bands shown in Figure 3 are an approximation, which, most likely, are slightly tighter than they should be.

Figure 2: "Core" PCE price month-over-month inflation



In all charts, the red line is our estimate, while the shaded area is the ± one standard deviation confidence band.

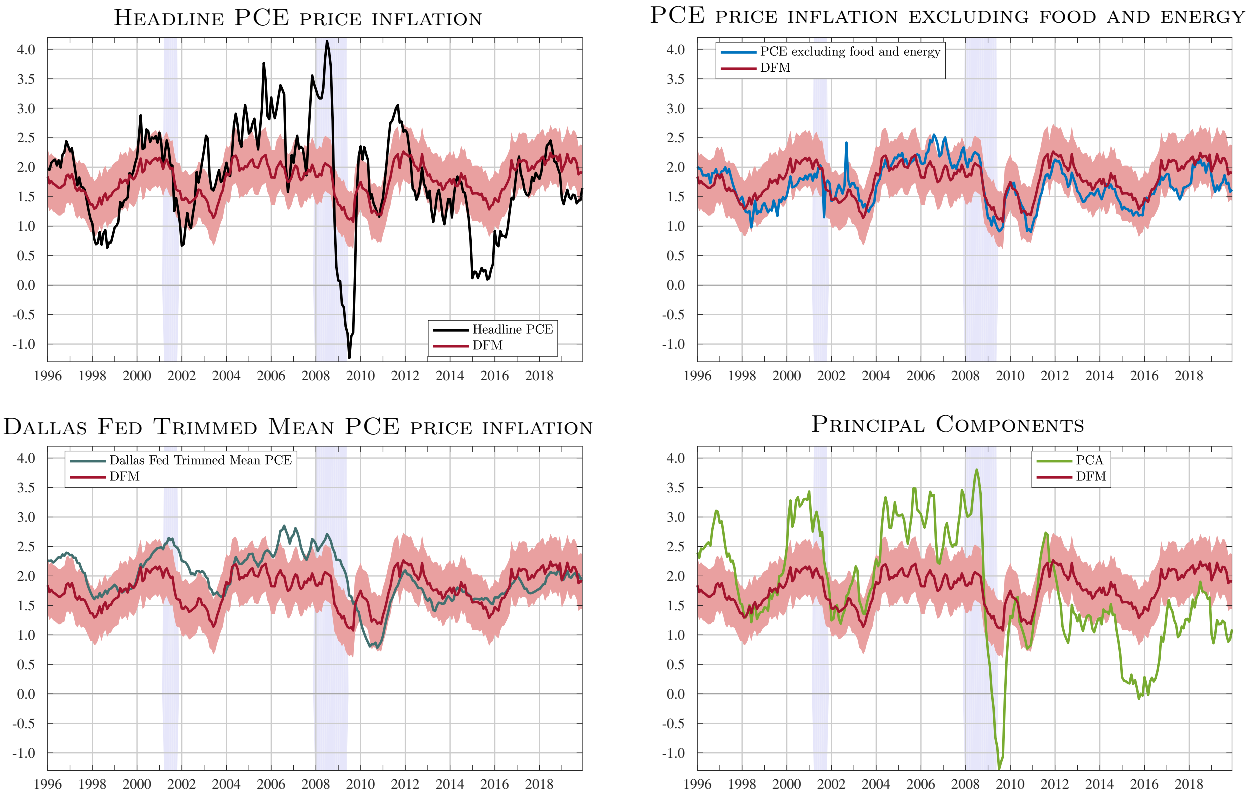
The other charts in Figures 2 and 3 compare our estimate with other core PCE price inflation estimates. Starting with the upper-right charts, our estimate of core inflation is quite similar to PCE price inflation excluding food and energy (the blue line), but less volatile. Indeed, our estimate is not affected by well-known idiosyncratic shocks such as the (down-up) spikes in September-October 2001 or the large decline in March 2017, which not surprisingly are 3 of the 15 (out of 300) dates in which PCE price inflation excluding food and energy is lying outside the confidence band of our estimate of core inflation.[14] Moreover, as shown in Figure 4, our estimate of core inflation captures primarily fluctuations with periods longer than six months, while a large share of fluctuations in PCE price inflation excluding food and energy is accounted for by fluctuations with periods shorter than six months. Finally, as can be clearly seen in Figure 3, our measure of core inflation points towards higher inflation at the end of the 1990s, which is in line with the literature indicating that the U.S. economy was very tight before the dot com bubble burst (see, e.g., Hasenzagl et al., 2020; Barigozzi and Luciani, 2020).

Next, the lower-left charts in Figures 2 and 3 compare our estimate of core inflation with the Dallas Fed Trimmed Mean PCE price inflation proposed by Dolmas (2005) (the slate-grey line), a measure that is highly considered by officials at the Federal Reserve and by newspapers.[15] Our measure and

[14]The 2001 swing in core PCE price inflation was driven by the price index for life insurance, which plunged 55 percent in September 2001 and jumped 121 percent in October 2001 as a result of the 9/11 terrorist attacks. The March 2017 decline in core PCE price inflation was largely due to the plunge in the price index for wireless telephone services (52% at an annual rate). The plunge was due to both a methodological change in the measurement of wireless services in the CPI and the fact that in late February of 2017 both Verizon and AT&T (which in March 2017 accounted for nearly 70% of wireless subscriptions in the US) brought back unlimited data plans.

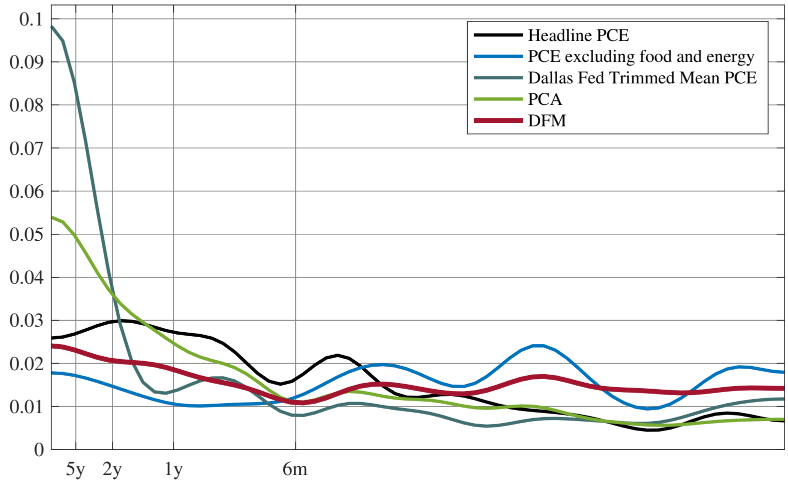
[15]The Dallas Fed Trimmed Mean PCE price inflation estimates core inflation by taking the weighted trimmed mean of

Figure 3: “Core” PCE price year-over-year inflation



In all charts, the red line is our estimate, while the shaded area is the ± one standard deviation confidence band.

Figure 4: Spectral density PCE price inflation



The spectral densities are standardized so that the integral below the curve is equal to one. The x-ticks stands for frequencies corresponding to periods of “5 years”, “2 years”, “1 year”, and “6 months.” Points on the right of a given x-tick denote fluctuations with period shorter than the x-tick.

the Dallas trimmed mean are remarkably similar, and they also capture similar frequencies. However, our measure performs better in capturing the decline in inflation during recessions, where the Dallas

a dataset of disaggregated PCE price inflation similar to the one used in this paper. As currently computed, this measure is computed by trimming out 24 percent from the lower tail of the distribution of monthly price changes and 31 percent from the upper tail.

trimmed mean is a bit lagging, as is evident when looking at Figure 3.

Finally, the lower-right charts in Figures 2 and 3 show the comparison with a principal component estimate. This is the estimate of core inflation that comes from a high-dimensional *static* factor model. By looking at the two charts, it is clear that a static factor model does not do a good job in estimating core inflation, as the estimate is very volatile, thus failing to achieve one of the goals a core inflation indicator is supposed to achieve. Even more so, the PCA estimate is very similar to the headline index itself. This demonstrates the importance of considering dynamic (GDFM) rather than static (DFM) loadings when constructing a core inflation indicator.

8 Conclusion

Factor models, in the past decades, have emerged as the most efficient tool in the analysis and prediction of high-dimensional time series (high-dimensional panel data). Several factor models have been proposed in the literature, the most flexible of which is the so-called *Generalized Dynamic Factor Model* (GDFM) where common shocks are loaded via filters—as opposed to the Dynamic Factor Model (DFM) where shocks are loaded in a static way. While complete results on the asymptotic behavior of DFM estimators are available (Bai, 2003), the corresponding theory for estimators of the GDFM is still incomplete. This paper fills that gap by deriving the asymptotic distributions of the GDFM estimators (common shocks, loadings, and common components).

Our results paves the way for inferential applications of the GDFM of great interest to macro and applied economists, such as asymptotic confidence intervals in prediction and in the construction of economic indicators. We illustrate the use of our methodology with an application to the construction of "core" inflation indicators for the U.S. economy. The GDFM-based indicator appears to provide much stable results than the current methods—it also outperforms its DFM-based counterpart, which appears to be much more volatile.

APPENDIX

This Appendix collects the proofs of the main results. For simplicity, we throughout assume that Assumptions (S) (from *(a)* through *(f)*) and (K) hold—even though most results are valid under a subset thereof.

A Proof of Theorem 1

A.1 Preliminary lemmas

Lemma 1. *As* $n, T \to \infty$,

$$\text{(i)}\quad \left\| \frac{\boldsymbol{\mathcal{U}}_T' \boldsymbol{\mathcal{U}}_T}{T} - \boldsymbol{\Gamma}^u \right\| = O_{\mathrm{P}}\left(\frac{1}{\sqrt{T}}\right), \ \text{as } T \to \infty; \qquad \text{(ii)}\quad \left\| \frac{\boldsymbol{\mathcal{R}}_n' \boldsymbol{\mathcal{R}}_n}{n} - \boldsymbol{\Sigma}^R \right\| = O_{\mathrm{P}}\left(\frac{1}{\sqrt{n}}\right), \ \text{as } n \to \infty.$$

Proof. Part (i) follows from parts *(a)* and *(c)* of Assumption (S) and (26); part (ii) follows from part *(a)* of Assumption (S) and (25). □

Lemma 2. *For any given t and any $\bar{n} \le n$ such that $\frac{1}{\bar{n}} + \frac{\bar{n}}{n} \to 0$ as $n \to \infty$,*

$$\frac{1}{\sqrt{\bar{n}}} \|\widehat{\mathbf{z}}_{\bar{n}t} - \mathbf{z}_{\bar{n}t}\| = O_{\mathrm{P}}(\zeta_{nT}) \qquad \text{as } n, T \to \infty.$$

Proof. Without loss of generality, set $\bar{n} = \bar{m}(q+1)$, implying $\bar{m} \sim c\bar{n}$. Then, because of Proposition 2,

$$\begin{aligned}
\|\widehat{\mathbf{z}}_{\bar{n}t} - \mathbf{z}_{\bar{n}t}\| &= \left\| \left(\widehat{\mathbf{A}}_{\bar{n}}(L) - \mathbf{A}_{\bar{n}}(L)\right) \mathbf{x}_{\bar{n}t} \right\| \le \sum_{r=0}^{p} \left(\sum_{i=1}^{\bar{m}} \mathbf{x}_{t-r}^{(i)\prime} \left(\widehat{\mathbf{A}}_r^{(i)} - \mathbf{A}_r^{(i)}\right)' \left(\widehat{\mathbf{A}}_r^{(i)} - \mathbf{A}_r^{(i)}\right) \mathbf{x}_{t-r}^{(i)} \right)^{1/2} \\
&\le \sum_{r=0}^{p} \left(\sum_{i=1}^{\bar{m}} \left(\mathbf{x}_{t-r}^{(i)\prime} \mathbf{x}_{t-r}^{(i)} \right)^2 \right)^{1/4} \left(\sum_{i=1}^{\bar{m}} \left(\sum_{j_i=1}^{q+1} \sum_{h_i=1}^{q+1} (\widehat{a}_{j_i,h_i,r} - a_{j_i,h_i,r})^2 \right)^2 \right)^{1/4} \\
&\le \sum_{r=0}^{p} \left(\sum_{i=1}^{\bar{m}} \left(\mathbf{x}_{t-r}^{(i)\prime} \mathbf{x}_{t-r}^{(i)} \right)^2 \right)^{1/4} \left((q+1)^3 \sum_{i=1}^{\bar{m}} \left\| \widehat{\mathbf{A}}_r^{(i)} - \mathbf{A}_r^{(i)} \right\|^4 \right)^{1/4} \\
&= O_{\mathrm{P}}\left(\sqrt{\bar{n}} \zeta_{nT} \right),
\end{aligned}$$

where $p = \max_{s=1,\dots,\bar{m}} p_s$, and $a_{j_i,h_i,r}$ and $\widehat{a}_{j_i,h_i,r}$ are the (j,h)th entries of $\mathbf{A}_r^{(i)}$ and of $\widehat{\mathbf{A}}_r^{(i)}$, respectively. See also (D.8) in the proof the Lemma 11 in Forni et al. (2017), which in turn follows from Lemmas 8 through 10, which entail uniformity over i for $\|\widehat{\mathbf{A}}_r^{(i)} - \mathbf{A}_r^{(i)}\|$. $\square$

Lemma 3. *Collect the q largest eigenvalues of $\widetilde{\mathbf{\Gamma}}_n^z := \frac{\mathbf{Z}_{nT}'\mathbf{Z}_{nT}}{T}$ in the $q \times q$ diagonal matrix $\widetilde{\mathbf{\Lambda}}_n^z$ and the corresponding normalized eigenvectors in $\widetilde{\mathbf{P}}_n^z$. Then, as $n, T \to \infty$,*

(i) $\frac{1}{n}\|\widetilde{\mathbf{\Lambda}}_n^z - \mathbf{\Lambda}_n^\psi\| = O_{\mathrm{P}}\left(\max\left(\frac{1}{\sqrt{T}}, \frac{1}{n}\right)\right)$;

(ii) there exists a $q \times q$ diagonal matrix $\widehat{\mathbf{W}}_1^z$ with entries ± 1 such that, for any $\bar{n} \le n$ such that $\frac{1}{\bar{n}} + \frac{\bar{n}}{n} \to 0$ as $n \to \infty$, $\|\widetilde{\mathbf{P}}_{\bar{n}}^z - \mathbf{P}_{\bar{n}}^\psi \widehat{\mathbf{W}}_1^z\| = O_{\mathrm{P}}\left(\frac{\bar{n}}{n}\max\left(\frac{1}{\sqrt{T}}, \frac{1}{n}\right)\right)$.

Proof. From Assumption (S*(d1)*), (36), and Lemma 1(i) it follows that, as $n, T \to \infty$,

$$\begin{aligned}
\frac{1}{n}\left\|\widetilde{\mathbf{\Gamma}}_n^z - \boldsymbol{\mathcal{R}}_n \mathbf{\Gamma}^u \boldsymbol{\mathcal{R}}_n'\right\| &= \frac{1}{n}\left\| \boldsymbol{\mathcal{R}}_n \frac{\boldsymbol{\mathcal{U}}_T'\boldsymbol{\mathcal{U}}_T}{T} \boldsymbol{\mathcal{R}}_n' + \frac{\mathbf{\Phi}_{nT}'\mathbf{\Phi}_{nT}}{T} - \boldsymbol{\mathcal{R}}_n \mathbf{\Gamma}^u \boldsymbol{\mathcal{R}}_n' \right\| \\
&\le \frac{1}{n}\left\| \boldsymbol{\mathcal{R}}_n \frac{\boldsymbol{\mathcal{U}}_T'\boldsymbol{\mathcal{U}}_T}{T} \boldsymbol{\mathcal{R}}_n' - \boldsymbol{\mathcal{R}}_n \mathbf{\Gamma}^u \boldsymbol{\mathcal{R}}_n' \right\|_F + \frac{1}{n}\left\| \frac{\mathbf{\Phi}_{nT}'\mathbf{\Phi}_{nT}}{T} \right\| \\
&= \frac{1}{n}\left\|\mathbf{\Gamma}_n^\phi\right\| + O_{\mathrm{P}}\left(\frac{1}{\sqrt{T}}\right) \le \frac{2\pi B^\phi}{n} + O_{\mathrm{P}}\left(\frac{1}{\sqrt{T}}\right) \\
&= O_{\mathrm{P}}\left(\max\left(\frac{1}{\sqrt{T}}, \frac{1}{n}\right)\right),
\end{aligned}$$

which implies

$$\frac{1}{n}\left\|\widetilde{\mathbf{\Lambda}}_n^z - \mathbf{\Lambda}_n^\psi\right\| \le \frac{1}{n}\left\|\widetilde{\mathbf{\Gamma}}_n^z - \boldsymbol{\mathcal{R}}_n \mathbf{\Gamma}^u \boldsymbol{\mathcal{R}}_n'\right\| = O_{\mathrm{P}}\left(\max\left(\frac{1}{\sqrt{T}}, \frac{1}{n}\right)\right).$$

hence part *(i)* of the claim. Turning to *(ii)*, by the Davis-Kahn sin-θ Theorem (see also Yu et al., 2015, Theorem 2) there exists a $q \times q$ diagonal matrix $\widehat{\mathbf{W}}_1^z$ with entries ± 1 such that

$$\left\|\widetilde{\mathbf{P}}_n^z - \mathbf{P}_n^\psi \widehat{\mathbf{W}}_1^z\right\| \le \frac{2^{3/2}\sqrt{q}\|\widetilde{\mathbf{\Gamma}}_n^z - \boldsymbol{\mathcal{R}}_n \mathbf{\Gamma}^u \boldsymbol{\mathcal{R}}_n'\|}{\min(\mu_{n0}^\psi - \mu_{n1}^\psi, \mu_{nq}^\psi - \mu_{n,q+1}^\psi)} = O_{\mathrm{P}}\left(\max\left(\frac{1}{\sqrt{T}}, \frac{1}{n}\right)\right),$$

where μ_{nj}^{ψ} are the eigenvalues of $\boldsymbol{\mathcal{R}}_n\boldsymbol{\Gamma}^u\boldsymbol{\mathcal{R}}_n'$ (satisfying (28) and Assumption (S*(e)*), $\mu_{n0}^{\psi} := \infty$, and $\mu_{n,q+1}^{\psi} = 0$. Similarly, for $\bar{n} \le n$,

$$\left\|\widetilde{\mathbf{P}}_{\bar{n}}^z - \mathbf{P}_{\bar{n}}^{\psi}\widehat{\mathbf{W}}_1^z\right\| \le \frac{2^{3/2}\sqrt{q}\|\widetilde{\boldsymbol{\Gamma}}_{\bar{n}}^z - \boldsymbol{\mathcal{R}}_{\bar{n}}\boldsymbol{\Gamma}^u\boldsymbol{\mathcal{R}}_{\bar{n}}'\|}{\min(\mu_{n0}^{\psi} - \mu_{n1}^{\psi}, \mu_{nq}^{\psi} - \mu_{n,q+1}^{\psi})} = O_{\mathrm{P}}\left(\frac{\bar{n}}{n}\max\left(\frac{1}{\sqrt{T}}, \frac{1}{n}\right)\right), \tag{65}$$

which completes the proof. □

Lemma 4. *Collect the q largest eigenvalues of $\widehat{\boldsymbol{\Gamma}}_n^z := \frac{\widehat{\boldsymbol{\mathcal{Z}}}_{nT}'\widehat{\boldsymbol{\mathcal{Z}}}_{nT}}{T}$ in the $q\times q$ diagonal matrix $\widehat{\boldsymbol{\Lambda}}_n^z$ and the corresponding normalized eigenvectors in $\widehat{\mathbf{P}}_n^z$. Then, as $n, T \to \infty$,*

(i) $\frac{1}{n}\|\widehat{\boldsymbol{\Lambda}}_n^z - \widetilde{\boldsymbol{\Lambda}}_n^z\| = O_{\mathrm{P}}(\zeta_{nT})$;

(ii) there exists a $q\times q$ diagonal matrix $\widehat{\mathbf{W}}_2^z$ with entries ± 1 such that, for any $\bar{n} \le n$ such that $\frac{1}{\bar{n}}+\frac{\bar{n}}{n} \to 0$ and $\bar{n} \to \infty$ as $n \to \infty$, $\|\widetilde{\mathbf{P}}_{\bar{n}}^z - \widehat{\mathbf{P}}_{\bar{n}}^z\widehat{\mathbf{W}}_2^z\| = O_{\mathrm{P}}(\frac{\bar{n}}{n}\zeta_{nT})$.

Proof. From Lemma 2 it immediately follows that $\frac{1}{n}\left\|\widehat{\boldsymbol{\Gamma}}_n^z - \widetilde{\boldsymbol{\Gamma}}_n^z\right\| = O_{\mathrm{P}}(\zeta_{nT})$, which implies

$$\frac{1}{n}\left\|\widehat{\boldsymbol{\Lambda}}_n^z - \widetilde{\boldsymbol{\Lambda}}_n^z\right\| \le \frac{1}{n}\left\|\widehat{\boldsymbol{\Gamma}}_n^z - \widetilde{\boldsymbol{\Gamma}}_n^z\right\| = O_{\mathrm{P}}(\zeta_{nT}),$$

hence part (i) of the claim. Now, from Lemma 3(i), with probability tending to one as $n, T \to \infty$, there exists a positive real c such that

$$\frac{1}{n}\left|\widetilde{\mu}_{nj}^z - \mu_{nj}^{\psi}\right| \le c\max\left(\frac{1}{\sqrt{T}}, \frac{1}{n}\right),\ j = 1,\ldots,q \quad \text{and} \quad \frac{1}{n}\left|\widetilde{\mu}_{nj}^z\right| \le c\max\left(\frac{1}{\sqrt{T}}, \frac{1}{n}\right),\ j = q+1,\ldots,n.$$

Thus, from (28), with probability tending to one as $n, T \to \infty$,

$$\widetilde{\mu}_{nj}^z \ge \mu_{nj}^{\psi} - c\max\left(\frac{1}{\sqrt{T}}, \frac{1}{n}\right) \ge n\alpha_j^{\psi} - c,\ j = 1,\ldots,q,$$

and $\widetilde{\mu}_{nj}^z \le c,\ j = q+1,\ldots,n$. Therefore, for $n \ge \frac{4c}{\alpha_j^{\psi}}$, with probability tending to one as $n, T \to \infty$, it holds that

$$\widetilde{\mu}_{nq}^z - \widetilde{\mu}_{n,q+1}^z \ge n\alpha_j^{\psi} - 2c = n\alpha_j^{\psi}\left(1 - \frac{2c}{\alpha_j^{\psi}n}\right) \ge n\frac{\alpha_j^{\psi}}{2}.$$

Then, by the Davis-Kahn sin-θ Theorem again, there exists a $q\times q$ diagonal matrix $\widehat{\mathbf{W}}_2^z$ with entries ± 1 such that

$$\left\|\widehat{\mathbf{P}}_n^z - \widetilde{\mathbf{P}}_n^z\widehat{\mathbf{W}}_2^z\right\| \le \frac{2^{3/2}\sqrt{q}\|\widehat{\boldsymbol{\Gamma}}_n^z - \widetilde{\boldsymbol{\Gamma}}_n^z\|}{\min(\widetilde{\mu}_{n0}^z - \widetilde{\mu}_{n1}^z, \widetilde{\mu}_{nq}^z - \widetilde{\mu}_{n,q+1}^z)} = O_{\mathrm{P}}(\zeta_{nT}),$$

where $\widetilde{\mu}_{nj}^z$ are the eigenvalues of $\widetilde{\boldsymbol{\Gamma}}_n^z$ and $\widetilde{\mu}_{n0}^z := \infty$. It follows that, for any $\bar{n} \le n$,

$$\left\|\widehat{\mathbf{P}}_{\bar{n}}^z - \widetilde{\mathbf{P}}_{\bar{n}}^z\widehat{\mathbf{W}}_2^z\right\| = O_{\mathrm{P}}\left(\frac{\bar{n}}{n}\zeta_{nT}\right).$$ □

Lemma 5. *As $n, T \to \infty$,*

(i) $\frac{1}{n}\|\widehat{\boldsymbol{\Lambda}}_n^z - \boldsymbol{\Lambda}_n^{\psi}\| = O_{\mathrm{P}}(\zeta_{nT})$;

(ii) for any $\bar{n} \le n$ such that $\frac{1}{\bar{n}} + \frac{\bar{n}}{n} \to 0$ as $n \to \infty$, $\|\widehat{\mathbf{P}}_{\bar{n}}^z - \mathbf{P}_{\bar{n}}^{\psi}\widehat{\mathbf{W}}^z\| = O_{\mathrm{P}}(\frac{\bar{n}}{n}\zeta_{nT})$, with $\widehat{\mathbf{W}}^z = \widehat{\mathbf{W}}_1^z\widehat{\mathbf{W}}_2^z$, where $\widehat{\mathbf{W}}_1^z$ is defined in Lemma 3 and $\widehat{\mathbf{W}}_2^z$ in Lemma 4.

Proof. From Lemmas 3(i) and 4(i) it holds that

$$\frac{1}{n}\left\|\widehat{\mathbf{\Lambda}}_n^z-\mathbf{\Lambda}_n^\psi\right\|\le\frac{1}{n}\left\|\widehat{\mathbf{\Lambda}}_n^z-\widetilde{\mathbf{\Lambda}}_n^z\right\|+\frac{1}{n}\left\|\widetilde{\mathbf{\Lambda}}_n^z-\mathbf{\Lambda}_n^\psi\right\|=O_{\mathrm{P}}(\zeta_{nT})+O_{\mathrm{P}}\left(\max\left(\frac{1}{\sqrt{T}},\frac{1}{n}\right)\right).$$

Part (i) of the claim follows, since $\frac{1}{\sqrt{T}}$ and $\frac{1}{n}$ are $O(\zeta_{nT})$. From Lemmas 3(ii) and 4(ii), and since $\|\widehat{\mathbf{W}}_2^z\|=1$, we obtain

$$\begin{aligned}\left\|\widehat{\mathbf{P}}_n^z-\mathbf{P}_n^\psi\widehat{\mathbf{W}}_1^z\widehat{\mathbf{W}}_2^z\right\|&\le\left\|\widehat{\mathbf{P}}_n^z-\widetilde{\mathbf{P}}_n^z\widehat{\mathbf{W}}_2^z\right\|+\left\|\widetilde{\mathbf{P}}_n^z\widehat{\mathbf{W}}_2^z-\mathbf{P}_n^\psi\widehat{\mathbf{W}}_1^z\widehat{\mathbf{W}}_2^z\right\|\\&\le O_{\mathrm{P}}(\zeta_{nT})+\left\|\widetilde{\mathbf{P}}_n^z-\mathbf{P}_n^\psi\widehat{\mathbf{W}}_1^z\right\|\,\|\widehat{\mathbf{W}}_2^z\|\\&=O_{\mathrm{P}}(\zeta_{nT})+O_{\mathrm{P}}\left(\max\left(\frac{1}{\sqrt{T}},\frac{1}{n}\right)\right).\end{aligned}$$

Since $\frac{1}{\sqrt{T}}$ and $\frac{1}{n}$ are $O(\zeta_{nT})$, this concludes the proof. $\square$

Lemma 6. *There exists a positive definite $q\times q$ diagonal matrix $\boldsymbol{\mathcal{L}}^u$ such that*

$$\frac{\mathbf{\Lambda}_n^\psi}{n}\longrightarrow_{\mathrm{P}}\boldsymbol{\mathcal{L}}^u\qquad\textit{as } n\to\infty.$$

Proof. The Lemma is an immediate consequence of (28). $\square$

Lemma 7. *(i)* $\left\|\left(\frac{\mathbf{\Lambda}_n^\psi}{n}\right)^{-1}\right\|=O_{\mathrm{P}}(1)$ *as $n\to\infty$; (ii)* $\left\|\left(\frac{\widehat{\mathbf{\Lambda}}_n^z}{n}\right)^{-1}\right\|=O_{\mathrm{P}}(1)$ *as $n,T\to\infty$.*

Proof. Part (i) follows from (28), part (ii) from Lemma 5(i) and part (i). $\square$

Lemma 8. *Denoting by $\mathbf{e}_{ni}$ the ith column of $\mathbf{I}_n$,*

$$\max_{i=1,\dots,n}\left\|\mathbf{e}_{ni}'\mathbf{P}_n^\psi\right\|=O_{\mathrm{P}}\left(\frac{1}{\sqrt{n}}\right)\qquad\textit{as } n\to\infty.$$

Proof. Since $\mathbf{P}_n^\psi=(\boldsymbol{\mathcal{R}}_n\mathbf{\Gamma}^u\boldsymbol{\mathcal{R}}_n')\mathbf{P}_n^\psi(\mathbf{\Lambda}_n^\psi)^{-1}$, we have

$$\max_{i=1,\dots,n}\left\|\mathbf{e}_{ni}'\mathbf{P}_n^\psi\right\|\le\max_{i=1,\dots,n}\left\|\mathbf{e}_{ni}'\boldsymbol{\mathcal{R}}_n\mathbf{\Gamma}^u\boldsymbol{\mathcal{R}}_n'\right\|\,\left\|\mathbf{P}_n^\psi\right\|\,\left\|(\mathbf{\Lambda}_n^\psi)^{-1}\right\|=O_{\mathrm{P}}\left(\frac{1}{\sqrt{n}}\right).$$

Indeed, $\|\mathbf{e}_{ni}'\boldsymbol{\mathcal{R}}_n\mathbf{\Gamma}^u\boldsymbol{\mathcal{R}}_n'\|=O_{\mathrm{P}}(\sqrt{n})$, $\|\mathbf{P}_n^\psi\|=1$, and $\|(\mathbf{\Lambda}_n^\psi)^{-1}\|=O_{\mathrm{P}}(n^{-1})$, because of Lemma 7(i) (which, actually, only requires Assumption (S)). $\square$

Lemma 9. *For any $\bar{n}\le n$ such that $\frac{1}{\bar{n}}+\frac{\bar{n}}{n}\to 0$ as $n\to\infty$,*

$$(i)\ \ \|\mathbf{P}_{\bar{n}}^\psi\|=O_{\mathrm{P}}\left(\sqrt{\frac{\bar{n}}{n}}\right)\quad\textit{and}\quad(ii)\ \ \|\mathbf{P}_{\bar{n}}^{\psi\prime}\mathbf{P}_{\bar{n}}^\psi\|=O_{\mathrm{P}}\left(\frac{\bar{n}}{n}\right).$$

Proof. It follows from Lemma 8 that

$$\left\|\mathbf{P}_{\bar{n}}^\psi\right\|^2\le\left\|\mathbf{P}_{\bar{n}}^\psi\right\|_F^2=\sum_{i=1}^{\bar{n}}\left\|\mathbf{e}_{ni}'\mathbf{P}_n^\psi\right\|^2\le\bar{n}\max_{i=1,\dots,\bar{n}}\left\|\mathbf{e}_{ni}'\mathbf{P}_n^\psi\right\|^2=O_{\mathrm{P}}\left(\frac{\bar{n}}{n}\right).$$

Moreover,

$$\left\|\mathbf{P}_{\bar n}^{\psi\prime}\mathbf{P}_{\bar n}^{\psi}\right\| \le \left\|\mathbf{P}_{\bar n}^{\psi}\right\|^2 = O_{\mathrm{P}}\left(\frac{\bar n}{n}\right). \qquad \square$$

Lemma 10. *For any $t \in \mathbb{Z}$ and any $\bar n \le n$ such that $\frac{1}{\bar n} + \frac{\bar n}{n} \to 0$ as $n \to \infty$,*

$$\sqrt{\frac{n}{\bar n}}\|\mathbf{P}_{\bar n}^{\psi\prime}\boldsymbol{\phi}_{\bar n t}\| = O_{\mathrm{P}}(1) \quad \text{as } n \to \infty.$$

Proof. Recall that $\frac{n}{\bar n}\|\mathbf{P}_{\bar n}^{\psi\prime}\mathbf{P}_{\bar n}^{\psi}\| = O_{\mathrm{P}}(1)$, because of Lemma 9(ii). Therefore, for the kth column of $\mathbf{P}_{\bar n}^{\psi}$, denoted as $\mathbf{p}_{\bar n k}^{\psi}$, it holds that $\frac{n}{\bar n}\mathbf{p}_{\bar n k}^{\psi\prime}\mathbf{p}_{\bar n k}^{\psi} = O_{\mathrm{P}}(1)$. Let $\widetilde{\mathbf{p}}_{\bar n k}^{\psi} := \mathbf{p}_{\bar n k}^{\psi}/\sqrt{\mathbf{p}_{\bar n k}^{\psi\prime}\mathbf{p}_{\bar n k}^{\psi}}$, so that $\widetilde{\mathbf{p}}_{\bar n k}^{\psi\prime}\widetilde{\mathbf{p}}_{\bar n k}^{\psi} = 1$. Let $\widetilde{p}_{ik}^{\psi}$ denote the ith entry of $\widetilde{\mathbf{p}}_{\bar n k}^{\psi}$ and let $\widetilde{\mathbf{P}}_{\bar n}^{\psi}$ be the matrix with columns $\widetilde{\mathbf{p}}_{\bar n 1}^{\psi}, \dots, \widetilde{\mathbf{p}}_{\bar n \bar n}^{\psi}$. Due to normalization of $\widetilde{\mathbf{p}}_{\bar n k}^{\psi}$ and Lemma 8, there exists a finite positive real $\bar c$ such that $\max_{i=1,\dots,n}\max_{j=1,\dots,q}|\widetilde{p}_{ij}^{\psi}| \le \frac{\bar c}{\sqrt{\bar n}}$ with probability one. Then, denoting by $\boldsymbol{\iota}_{\bar n}$ a $\bar n$-dimensional column vector of ones, for any $t \in \mathbb{Z}$,

$$\begin{aligned}
\mathbb{E}\left[\left\|\widetilde{\mathbf{P}}_{\bar n}^{\psi\prime}\boldsymbol{\phi}_{\bar n t}\right\|^2\right] &= \mathbb{E}\left[\sum_{k=1}^{q}\left(\sum_{i=1}^{\bar n}\widetilde{p}_{ik}^{\psi}\phi_{it}\right)^2\right] = \sum_{k=1}^{q}\sum_{i=1}^{\bar n}\sum_{j=1}^{\bar n}\mathbb{E}[\widetilde{p}_{ik}^{\psi}\widetilde{p}_{jk}^{\psi}\phi_{it}\phi_{jt}] \\
&\le \sum_{k=1}^{q}\sum_{i=1}^{\bar n}\sum_{j=1}^{\bar n}\frac{\bar c^2}{\bar n}\mathbb{E}[\phi_{it}\phi_{jt}] \\
&\le q\bar c^2 \max_{k=1,\dots,q}\frac{\boldsymbol{\iota}_{\bar n}'}{\sqrt{\bar n}}\mathbb{E}[\boldsymbol{\phi}_{\bar n t}\boldsymbol{\phi}_{\bar n t}']\frac{\boldsymbol{\iota}_{\bar n}}{\sqrt{\bar n}} \le q\bar c^2 \max_{k=1,\dots,q}\max_{\substack{\boldsymbol{b}_{\bar n}\\ \boldsymbol{b}_{\bar n}'\boldsymbol{b}_{\bar n}=1}}\boldsymbol{b}_{\bar n}'\boldsymbol{\Gamma}_{\bar n}^{\phi}\boldsymbol{b}_{\bar n} \\
&\le q\bar c^2 \max_{k=1,\dots,q}\sup_{\bar n\in\mathbb{N}}\max_{\substack{\boldsymbol{b}_{\bar n}\\ \boldsymbol{b}_{\bar n}'\boldsymbol{b}_{\bar n}=1}}\boldsymbol{b}_{\bar n}'\boldsymbol{\Gamma}_{\bar n}^{\phi}\boldsymbol{b}_{\bar n} \le q\bar c^2 2\pi B^{\phi},
\end{aligned}$$

in view of (36). Hence, it follows from Chebychev's inequality that $\|\widetilde{\mathbf{P}}_{\bar n}^{\psi\prime}\boldsymbol{\phi}_{\bar n t}\| = O_{\mathrm{P}}(1)$ and, therefore, $\|\mathbf{P}_{\bar n}^{\psi\prime}\boldsymbol{\phi}_{\bar n t}\|$ is $O_{\mathrm{P}}\left(\sqrt{\frac{\bar n}{n}}\right)$. $\square$

A.2 Proof of Theorem 1

Let $\mathbf{x}_{\bar n t}, \widehat{\mathbf{z}}_{\bar n t}, \mathbf{z}_{\bar n t}, \boldsymbol{\phi}_{\bar n t}$ denote the first $\bar n$ elements of $\mathbf{x}_{nt}, \widehat{\mathbf{z}}_{nt}, \mathbf{z}_{nt}, \boldsymbol{\phi}_{nt}$, respectively. Then, from (41)

$$\begin{aligned}
\widehat{\mathbf{u}}_t &= \left(\left(\widehat{\boldsymbol{\Lambda}}_n^z\right)^{1/2}\widehat{\mathbf{P}}_{\bar n}^{z\prime}\widehat{\mathbf{P}}_{\bar n}^{z}\left(\widehat{\boldsymbol{\Lambda}}_n^z\right)^{1/2}\right)^{-1}\left(\widehat{\boldsymbol{\Lambda}}_n^z\right)^{1/2}\widehat{\mathbf{P}}_{\bar n}^{z\prime}\widehat{\mathbf{z}}_{\bar n t} = \left(\widehat{\boldsymbol{\Lambda}}_n^z\right)^{-1/2}\left(\widehat{\mathbf{P}}_{\bar n}^{z\prime}\widehat{\mathbf{P}}_{\bar n}^{z}\right)^{-1}\widehat{\mathbf{P}}_{\bar n}^{z\prime}\widehat{\mathbf{z}}_{\bar n t} \\
&= \left(\left(\widehat{\boldsymbol{\Lambda}}_n^z\right)^{-1/2} - \left(\boldsymbol{\Lambda}_n^{\psi}\right)^{-1/2}\right)\left(\widehat{\mathbf{P}}_{\bar n}^{z\prime}\widehat{\mathbf{P}}_{\bar n}^{z}\right)^{-1}\widehat{\mathbf{P}}_{\bar n}^{z\prime}\widehat{\mathbf{z}}_{\bar n t} \\
&\quad + \left(\boldsymbol{\Lambda}_n^{\psi}\right)^{-1/2}\left(\left(\widehat{\mathbf{P}}_{\bar n}^{z\prime}\widehat{\mathbf{P}}_{\bar n}^{z}\right)^{-1} - \left(\widehat{\mathbf{W}}^z\mathbf{P}_{\bar n}^{\psi\prime}\mathbf{P}_{\bar n}^{\psi}\widehat{\mathbf{W}}^z\right)^{-1}\right)\widehat{\mathbf{P}}_{\bar n}^{z\prime}\widehat{\mathbf{z}}_{\bar n t} \\
&\quad + \left(\boldsymbol{\Lambda}_n^{\psi}\right)^{-1/2}\left(\widehat{\mathbf{W}}^z\mathbf{P}_{\bar n}^{\psi\prime}\mathbf{P}_{\bar n}^{\psi}\widehat{\mathbf{W}}^z\right)^{-1}\left(\widehat{\mathbf{P}}_{\bar n}^{z\prime} - \widehat{\mathbf{W}}^z\mathbf{P}_{\bar n}^{\psi\prime}\right)\widehat{\mathbf{z}}_{\bar n t} \\
&\quad + \left(\boldsymbol{\Lambda}_n^{\psi}\right)^{-1/2}\widehat{\mathbf{W}}^z\left(\mathbf{P}_{\bar n}^{\psi\prime}\mathbf{P}_{\bar n}^{\psi}\right)^{-1}\mathbf{P}_{\bar n}^{\psi\prime}\left(\widehat{\mathbf{A}}_{\bar n}(L) - \mathbf{A}_{\bar n}(L)\right)\mathbf{x}_{\bar n t} \\
&\quad + \left(\boldsymbol{\Lambda}_n^{\psi}\right)^{-1/2}\widehat{\mathbf{W}}^z\left(\mathbf{P}_{\bar n}^{\psi\prime}\mathbf{P}_{\bar n}^{\psi}\right)^{-1}\mathbf{P}_{\bar n}^{\psi\prime}\mathbf{z}_{\bar n t} \\
&= I + II + III + IV + V, \text{ say.}
\end{aligned} \tag{66}$$

For I, since

$$\left(\left(\widehat{\mathbf{\Lambda}}_n^z\right)^{-1/2}-\left(\mathbf{\Lambda}_n^\psi\right)^{-1/2}\right)=\left(\left(\widehat{\mathbf{\Lambda}}_n^z\right)^{-1}-\left(\mathbf{\Lambda}_n^\psi\right)^{-1}\right)\left(\left(\widehat{\mathbf{\Lambda}}_n^z\right)^{-1/2}+\left(\mathbf{\Lambda}_n^\psi\right)^{-1/2}\right)^{-1}$$
$$=\left(\widehat{\mathbf{\Lambda}}_n^z\right)^{-1}\left(\mathbf{\Lambda}_n^\psi-\widehat{\mathbf{\Lambda}}_n^z\right)\left(\mathbf{\Lambda}_n^\psi\right)^{-1}\left(\left(\widehat{\mathbf{\Lambda}}_n^z\right)^{-1/2}+\left(\mathbf{\Lambda}_n^\psi\right)^{-1/2}\right)^{-1}, \qquad (67)$$

and because of (67) and Lemmas 5(i) and 7, the norm of I is bounded from above by

$$\left\|\left(\widehat{\mathbf{\Lambda}}_n^z\right)^{-1}\right\|\left\|\mathbf{\Lambda}_n^\psi-\widehat{\mathbf{\Lambda}}_n^z\right\|\left\|\left(\mathbf{\Lambda}_n^\psi\right)^{-1}\right\|\left\|\left(\left(\widehat{\mathbf{\Lambda}}_n^z\right)^{-1/2}+\left(\mathbf{\Lambda}_n^\psi\right)^{-1/2}\right)^{-1}\right\|\left\|\left(\widehat{\mathbf{P}}_{\bar n}^{z\prime}\widehat{\mathbf{P}}_{\bar n}^z\right)^{-1}\widehat{\mathbf{P}}_{\bar n}^z\right\|\|\widehat{\mathbf{z}}_{\bar n t}\|$$
$$=O_{\mathrm{P}}\left(\frac{1}{n^2}\sqrt{n}\zeta_{nT}\sqrt{n}\frac{\sqrt{n}}{\sqrt{\bar n}}\sqrt{\bar n}\right) \qquad (68)$$

since $\|\widehat{\mathbf{z}}_{\bar n t}\|=O_{\mathrm{P}}(\sqrt{\bar n})$ by Lemma 2, and $\|(\widehat{\mathbf{P}}_{\bar n}^{z\prime}\widehat{\mathbf{P}}_{\bar n}^z)^{-1}\widehat{\mathbf{P}}_{\bar n}^z\|=O_{\mathrm{P}}(\sqrt{\frac{n}{\bar n}})$ by Lemma 9(i) and 9(ii). This yields $I=O_{\mathrm{P}}\left(\frac{\zeta_{nT}}{\sqrt{n}}\right)$.

For II, first notice that from Lemma 5(ii),

$$\left\|\mathbf{P}_{\bar n}^\psi\widehat{\mathbf{W}}^z-\widehat{\mathbf{P}}_{\bar n}^z\right\|=O_{\mathrm{P}}\left(\frac{\bar n}{n}\zeta_{nT}\right). \qquad (69)$$

Then,

$$\left(\left(\widehat{\mathbf{P}}_{\bar n}^{z\prime}\widehat{\mathbf{P}}_{\bar n}^z\right)^{-1}-\left(\widehat{\mathbf{W}}^z\mathbf{P}_{\bar n}^{\psi\prime}\mathbf{P}_{\bar n}^\psi\widehat{\mathbf{W}}^z\right)^{-1}\right)$$
$$=\left(\widehat{\mathbf{P}}_{\bar n}^{z\prime}\widehat{\mathbf{P}}_{\bar n}^z\right)^{-1}\left(\widehat{\mathbf{W}}^z\mathbf{P}_{\bar n}^{\psi\prime}\mathbf{P}_{\bar n}^\psi\widehat{\mathbf{W}}^z-\widehat{\mathbf{P}}_{\bar n}^{z\prime}\widehat{\mathbf{P}}_{\bar n}^z\right)\left(\widehat{\mathbf{W}}^z\mathbf{P}_{\bar n}^{\psi\prime}\mathbf{P}_{\bar n}^\psi\widehat{\mathbf{W}}^z\right)^{-1}$$
$$=\left(\widehat{\mathbf{P}}_{\bar n}^{z\prime}\widehat{\mathbf{P}}_{\bar n}^z\right)^{-1}\left(\widehat{\mathbf{W}}^z\mathbf{P}_{\bar n}^{\psi\prime}\left(\mathbf{P}_{\bar n}^\psi\widehat{\mathbf{W}}^z-\widehat{\mathbf{P}}_{\bar n}^z\right)+\left(\widehat{\mathbf{W}}^z\mathbf{P}_{\bar n}^{\psi\prime}-\widehat{\mathbf{P}}_{\bar n}^{z\prime}\right)\widehat{\mathbf{P}}_{\bar n}^z\right)\left(\widehat{\mathbf{W}}^z\mathbf{P}_{\bar n}^{\psi\prime}\mathbf{P}_{\bar n}^\psi\widehat{\mathbf{W}}^z\right)^{-1}$$

and, because of (69) and Lemma 9,

$$\left\|\left(\left(\widehat{\mathbf{P}}_{\bar n}^{z\prime}\widehat{\mathbf{P}}_{\bar n}^z\right)^{-1}-\left(\widehat{\mathbf{W}}^z\mathbf{P}_{\bar n}^{\psi\prime}\mathbf{P}_{\bar n}^\psi\widehat{\mathbf{W}}^z\right)^{-1}\right)\right\|=O_{\mathrm{P}}\left(\sqrt{\frac{n}{\bar n}}\frac{\bar n}{n}\zeta_{nT}\frac{n}{\bar n}\right)=O_{\mathrm{P}}\left(\sqrt{\frac{n}{\bar n}}\zeta_{nT}\right). \qquad (70)$$

Because of (70), and Lemmas 2, 7(i), and 9(i), the norm of II is bounded from above by

$$\left\|\left(\mathbf{\Lambda}_n^\psi\right)^{-1/2}\right\|\left\|\left(\left(\widehat{\mathbf{P}}_{\bar n}^{z\prime}\widehat{\mathbf{P}}_{\bar n}^z\right)^{-1}-\left(\widehat{\mathbf{W}}^z\mathbf{P}_{\bar n}^{\psi\prime}\mathbf{P}_{\bar n}^\psi\widehat{\mathbf{W}}^z\right)^{-1}\right)\right\|\left\|\widehat{\mathbf{P}}_{\bar n}^{z\prime}\right\|\|\widehat{\mathbf{z}}_{\bar n t}\|$$
$$=O_{\mathrm{P}}\left(\frac{1}{\sqrt{n}}\sqrt{\frac{n}{\bar n}}\zeta_{nT}\sqrt{\frac{\bar n}{n}}\sqrt{\bar n}\right)=O_{\mathrm{P}}\left(\sqrt{\frac{\bar n}{n}}\zeta_{nT}\right), \qquad (71)$$

yielding $II=O_{\mathrm{P}}\left(\sqrt{\frac{\bar n}{n}}\zeta_{nT}\right)$.

By (69) and Lemmas 2, 7(i), and 9(i), one immediately gets $III=O_{\mathrm{P}}\left(\sqrt{\frac{\bar n}{n}}\zeta_{nT}\right)$ and $IV=O_{\mathrm{P}}\left(\frac{\zeta_{nT}}{\sqrt{n}}\right)$.

Finally, consider term V. Recall that, from Assumption (S*(d1)*), (36), (46), and Lemma 1(i), for any $n \in \mathbb{N}$, as $T \to \infty$,

$$\frac{1}{T}\sum_{t=1}^{T} \mathbf{z}_{nt}\mathbf{z}_{nt}' \longrightarrow_{\mathrm{P}} \boldsymbol{\mathcal{R}}_n \boldsymbol{\Gamma}^u \boldsymbol{\mathcal{R}}_n' + \boldsymbol{\Gamma}_n^{\phi} = \mathbf{P}_n^{\psi}\boldsymbol{\Lambda}_n^{\psi}\mathbf{P}_n^{\psi\prime} + \boldsymbol{\Gamma}_n^{\phi}$$

(see also the proof of Lemma 3). Considering the upper-left $\bar{n} \times \bar{n}$ submatrix $\boldsymbol{\mathcal{R}}_{\bar{n}} \boldsymbol{\Gamma}^u \boldsymbol{\mathcal{R}}_{\bar{n}}' = \mathbf{P}_{\bar{n}}^{\psi}\boldsymbol{\Lambda}_n^{\psi}\mathbf{P}_{\bar{n}}^{\psi\prime}$ of $\boldsymbol{\mathcal{R}}_n \boldsymbol{\Gamma}^u \boldsymbol{\mathcal{R}}_n'$, it follows that $\mathbf{z}_{\bar{n}t} = \mathbf{P}_{\bar{n}}^{\psi}\left(\boldsymbol{\Lambda}_n^{\psi}\right)^{1/2}\left(\boldsymbol{\Gamma}^u\right)^{-1/2}\mathbf{u}_t + \boldsymbol{\phi}_{\bar{n}t}$. Collecting terms,

$$\widehat{\mathbf{u}}_t - \widehat{\mathbf{W}}^z\left(\boldsymbol{\Gamma}^u\right)^{-1/2}\mathbf{u}_t = I + II + III + IV + \widehat{\mathbf{W}}^z\left(\boldsymbol{\Lambda}_n^{\psi}\right)^{-1/2}\left(\mathbf{P}_{\bar{n}}^{\psi\prime}\mathbf{P}_{\bar{n}}^{\psi}\right)^{-1}\mathbf{P}_{\bar{n}}^{\psi\prime}\boldsymbol{\phi}_{\bar{n}t}. \tag{72}$$

Recalling that $\|\widehat{\mathbf{W}}^z\| = 1$, it follows from (72) that, in view of Lemmas 7(i), 9(ii), and 10,

$$\begin{aligned}
\left\|\widehat{\mathbf{u}}_t - \widehat{\mathbf{W}}^z\left(\boldsymbol{\Gamma}^u\right)^{-1/2}\mathbf{u}_t\right\| &\le \left\|\left(\boldsymbol{\Lambda}_n^{\psi}\right)^{-1/2}\right\|\left\|\left(\mathbf{P}_{\bar{n}}^{\psi\prime}\mathbf{P}_{\bar{n}}^{\psi}\right)^{-1}\right\|\left\|\mathbf{P}_{\bar{n}}^{\psi\prime}\boldsymbol{\phi}_{\bar{n}t}\right\| + O_{\mathrm{P}}\left(\sqrt{\frac{\bar{n}}{n}}\zeta_{nT}\right)\\
&= O_{\mathrm{P}}\left(\frac{1}{\sqrt{n}}\frac{n}{\bar{n}}\sqrt{\frac{\bar{n}}{n}}\right) + O_{\mathrm{P}}\left(\sqrt{\frac{\bar{n}}{n}}\zeta_{nT}\right) = O_{\mathrm{P}}\left(\max\left(\frac{1}{\sqrt{\bar{n}}}, \sqrt{\frac{\bar{n}}{n}}\zeta_{nT}\right)\right).
\end{aligned}$$

This proves consistency.

Now, by (28), there exists a $q \times q$ positive definite diagonal matrix $\boldsymbol{\mathcal{L}}_u$ such that $\frac{\boldsymbol{\Lambda}_n^{\psi}}{n} \to_{\mathrm{P}} \boldsymbol{\mathcal{L}}^u$ as $n \to \infty$. Similarly, by Lemma 9(ii), there exists a $q \times q$ positive definite matrix $\boldsymbol{\mathcal{M}}_u$ such that, as $n \to \infty$, $\frac{n}{\bar{n}}\mathbf{P}_{\bar{n}}^{\psi\prime}\mathbf{P}_{\bar{n}}^{\psi} \longrightarrow_{\mathrm{P}} \boldsymbol{\mathcal{M}}^u$. Therefore, by Assumption (S(f1)), as $n, T \to \infty$,

$$\begin{aligned}
\sqrt{\bar{n}}\left(\widehat{\mathbf{u}}_t - \widehat{\mathbf{W}}^z\left(\boldsymbol{\Gamma}^u\right)^{-1/2}\mathbf{u}_t\right) &= \sqrt{\bar{n}}\,\widehat{\mathbf{W}}^z\left(\boldsymbol{\Lambda}_n^{\psi}\right)^{-1/2}\left(\mathbf{P}_{\bar{n}}^{\psi\prime}\mathbf{P}_{\bar{n}}^{\psi}\right)^{-1}\mathbf{P}_{\bar{n}}^{\psi\prime}\boldsymbol{\phi}_{\bar{n}t} + o_{\mathrm{P}}(1)\\
&= \widehat{\mathbf{W}}^z\left(\frac{\boldsymbol{\Lambda}_n^{\psi}}{n}\right)^{-1/2}\left(\frac{n}{\bar{n}}\mathbf{P}_{\bar{n}}^{\psi\prime}\mathbf{P}_{\bar{n}}^{\psi}\right)^{-1}\sqrt{\frac{n}{\bar{n}}}\left(\mathbf{P}_{\bar{n}}^{\psi\prime}\boldsymbol{\phi}_{\bar{n}t}\right) + o_{\mathrm{P}}(1)\\
&\longrightarrow_d \mathcal{N}\left(\mathbf{0}_q, \mathbf{W}^u\left(\boldsymbol{\mathcal{L}}^u\right)^{-1/2}\left(\boldsymbol{\mathcal{M}}^u\right)^{-1}\boldsymbol{\mathcal{P}}_t^u\left(\boldsymbol{\mathcal{M}}^u\right)^{-1}\left(\boldsymbol{\mathcal{L}}^u\right)^{-1/2}\mathbf{W}^u\right),
\end{aligned}$$

since $\sqrt{\bar{n}}\zeta_{nT} \to 0$, because of (48). $\square$

B Proof of Theorem 2

B.1 Preliminary lemmas

Lemma 11. *Collect the q largest eigenvalues of* $\widetilde{\boldsymbol{G}}_T^z := \frac{\boldsymbol{Z}_{nT}\boldsymbol{Z}_{nT}'}{n}$ *in* $\widetilde{\boldsymbol{L}}_T^z$ *and the corresponding normalized eigenvectors in* $\widetilde{\boldsymbol{\Pi}}_T^z$. *As* $n, T \to \infty$,

(i) $\frac{1}{T}\left\|\widetilde{\boldsymbol{L}}_T^z - \boldsymbol{L}_T^{\psi}\right\| = O_{\mathrm{P}}\left(\max\left(\frac{1}{\sqrt{n}}, \frac{1}{T}\right)\right)$;

(ii) there exists a $q \times q$ *diagonal matrix* $\widehat{\boldsymbol{W}}_1^z$ *with entries* ± 1 *such that, for any* $\bar{T} \le T$ *such that* $\frac{1}{T} + \frac{\bar{T}}{T} \to 0$ *as* $T \to \infty$, $\left\|\widetilde{\boldsymbol{\Pi}}_{\bar{T}}^z - \boldsymbol{\Pi}_{\bar{T}}^{\psi}\widehat{\boldsymbol{W}}_1^z\right\| = O_{\mathrm{P}}\left(\frac{\bar{T}}{T}\max\left(\frac{1}{\sqrt{n}}, \frac{1}{T}\right)\right)$.

Proof. The claim follows along the same lines as for Lemma 3 but using Assumption (S(*d2*), (37), and Lemma 1(ii) instead of Assumption (S(*d1*), (36), and Lemma 1(i). $\square$

Lemma 12. *Collect the q largest eigenvalues of* $\widehat{\boldsymbol{G}}_T^z := \frac{\widehat{\boldsymbol{Z}}_{nT}\widehat{\boldsymbol{Z}}_{nT}'}{n}$ *in the* $q \times q$ *diagonal matrix* $\widehat{\boldsymbol{L}}_T^z$ *and the corresponding normalized eigenvectors in* $\widehat{\boldsymbol{\Pi}}_T^z$. *As* $n, T \to \infty$, *(i)* $\frac{1}{T}\|\widehat{\boldsymbol{L}}^z_T - \widetilde{\boldsymbol{L}}^z_T\| = O_{\mathrm{P}}(\zeta_{nT})$;

(ii) *there exists a* $q \times q$ *diagonal matrix* $\widehat{\boldsymbol{W}}^z_1$ *with entries* ± 1 *such that,* $\frac{1}{\bar{T}} + \frac{\bar{T}}{T} \to 0$*, as* $T \to \infty$*, such that* $\bar{T} \to \infty$, $\|\widetilde{\mathbf{\Pi}}^z_{\bar{T}} - \widehat{\mathbf{\Pi}}^z_{\bar{T}}\widehat{\boldsymbol{W}}^z_2\| = O_{\mathrm{P}}(\frac{\bar{T}}{T}\zeta_{nT})$.

Proof. The claim follows along the same lines as for Lemma 4 but using Lemma 12 and (30). □

Lemma 13. *As* $n, T \to \infty$*,*

(i) $\frac{1}{T}\|\widehat{\boldsymbol{L}}^z_T - \boldsymbol{L}^\psi_T\| = O_{\mathrm{P}}(\zeta_{nT})$;

(ii) *for any* $\bar{T} \leq T$ *such that* $\frac{1}{\bar{T}} + \frac{\bar{T}}{T} \to 0$ *as* $T \to \infty$, $\|\widehat{\mathbf{\Pi}}^z_{\bar{T}} - \mathbf{\Pi}^\psi_{\bar{T}}\widehat{\boldsymbol{W}}^z\| = O_{\mathrm{P}}(\frac{\bar{T}}{T}\zeta_{nT})$*, with* $\widehat{\boldsymbol{W}}^z = \widehat{\boldsymbol{W}}^z_1\widehat{\boldsymbol{W}}^z_2$*, where* $\widehat{\boldsymbol{W}}^z_1$ *is defined in Lemma 11 and* $\widehat{\boldsymbol{W}}^z_2$ *in Lemma 12.*

Proof. Same as Lemma 5 but using Lemmas 11 and 12. □

Lemma 14. *There exists a positive definite* $q \times q$ *diagonal matrix* $\boldsymbol{\mathcal{L}}_R$ *such that* $\frac{\boldsymbol{L}^\psi_T}{T} \longrightarrow_{\mathrm{P}} \boldsymbol{\mathcal{L}}_R$ *as* $T \to \infty$*.*

Proof. This Lemma is an immediate consequence of (30). □

Lemma 15. *(i)* $\left\|\left(\frac{\boldsymbol{L}^\psi_T}{T}\right)^{-1}\right\| = O_{\mathrm{P}}(1)$ *as* $T \to \infty$*;* *(ii)* $\left\|\left(\frac{\widehat{\boldsymbol{L}}^z_T}{T}\right)^{-1}\right\| = O_{\mathrm{P}}(1)$ *as* $n, T \to \infty$*.*

Proof. Part (i) follows from (30), part (ii) from Lemma 13(i) and part (i). □

Lemma 16. *Denoting by* $\mathbf{e}_{Tt}$ *the* t*th column of* $\mathbf{I}_T$*,*

$$\max_{t=1,\dots,T}\left\|\mathbf{e}'_{Tt}\mathbf{\Pi}^\psi_T\right\| = O_{\mathrm{P}}\left(\frac{1}{\sqrt{T}}\right) \qquad \text{as } T \to \infty.$$

Proof. Same as the proof of Lemma 8 but using Lemma 15(i). □

Lemma 17. *For any* $\bar{T} \leq T$ *such that* $\frac{1}{\bar{T}} + \frac{\bar{T}}{T} \to 0$ *as* $T \to \infty$*,*

$$(i)\ \|\mathbf{\Pi}^\psi_{\bar{T}}\| = O_{\mathrm{P}}\left(\sqrt{\frac{\bar{T}}{T}}\right); \qquad (ii)\ \|\mathbf{\Pi}^{\psi\prime}_{\bar{T}}\mathbf{\Pi}^\psi_{\bar{T}}\| = O_{\mathrm{P}}\left(\frac{\bar{T}}{T}\right).$$

Proof. Same as Lemma 9 but using Lemma 16. □

Lemma 18. *For any* $i \in \mathbb{N}$ *and any* $\bar{T} \leq T$ *such that* $\frac{1}{\bar{T}} + \frac{\bar{T}}{T} \to 0$ *as* $T \to \infty$*,*

$$\sqrt{\frac{T}{\bar{T}}}\|\mathbf{\Pi}^{\psi\prime}_{\bar{T}}\boldsymbol{\varphi}^i_{\bar{T}}\| = O_{\mathrm{P}}(1).$$

Proof. Recall that, in view of Lemma 17(ii), $\frac{T}{\bar{T}}\|\mathbf{\Pi}^{\psi\prime}_{\bar{T}}\mathbf{\Pi}^\psi_{\bar{T}}\| = O_{\mathrm{P}}(1)$. Therefore, for the kth column of $\mathbf{\Pi}^\psi_{\bar{T}}$, denoted as $\boldsymbol{\pi}^\psi_{\bar{T}k}$, it holds that $\frac{T}{\bar{T}}\boldsymbol{\pi}^{\psi\prime}_{\bar{T}k}\boldsymbol{\pi}^\psi_{\bar{T}k} = O_{\mathrm{P}}(1)$. Let $\widetilde{\boldsymbol{\pi}}^\psi_{\bar{T}k} := \boldsymbol{\pi}^\psi_{\bar{T}k}/\sqrt{\boldsymbol{\pi}^{\psi\prime}_{\bar{T}k}\boldsymbol{\pi}^\psi_{\bar{T}k}}$, so that $\widetilde{\boldsymbol{\pi}}^{\psi\prime}_{\bar{T}k}\widetilde{\boldsymbol{\pi}}^\psi_{\bar{T}k} = 1$. Let $\widetilde{\pi}^\psi_{ik}$ be the ith entry of $\widetilde{\boldsymbol{\pi}}^\psi_{\bar{T}k}$ and denote by $\widetilde{\mathbf{\Pi}}^\psi_{\bar{T}}$ the matrix with columns $\widetilde{\boldsymbol{\pi}}^\psi_{\bar{T}1}, \dots, \widetilde{\boldsymbol{\pi}}^\psi_{\bar{T}\bar{T}}$. Due to normalization of $\widetilde{\boldsymbol{\pi}}^\psi_{\bar{T}k}$ and Lemma 16, there exists a finite positive real $\bar{c}$ such that

$$\max_{t=1,\dots,T}\max_{j=1,\dots,q}|\widetilde{\pi}^\psi_{tj}| \leq \frac{\bar{c}}{\sqrt{\bar{T}}}$$

with probability one. Then, denoting by $\boldsymbol{\iota}_{\bar T}$ the $\bar T$-dimensional column vector of ones, for any $i \in \mathbb{N}$,

$$
\begin{aligned}
\mathbb{E}\left[\left\|\widetilde{\boldsymbol{\Pi}}_{\bar T}^{\psi\prime}\boldsymbol{\varphi}_{\bar T}^{i}\right\|^2\right] &= \mathbb{E}\left[\sum_{k=1}^{q}\left(\sum_{t=1}^{\bar T}\widetilde{p}_{tk}^{\psi}\phi_{it}\right)^2\right] = \sum_{k=1}^{q}\sum_{t=1}^{\bar T}\sum_{s=1}^{\bar T}\mathbb{E}[\widetilde{\pi}_{tk}^{\psi}\widetilde{\pi}_{sk}^{\psi}\phi_{it}\phi_{is}]\\
&\le \sum_{k=1}^{q}\sum_{t=1}^{\bar T}\sum_{s=1}^{\bar T}\frac{\bar c^2}{\bar T}\mathbb{E}[\phi_{it}\phi_{is}] \le q\bar c^2 \max_{k=1,\dots,q}\frac{\boldsymbol{\iota}_{\bar T}'}{\sqrt{\bar T}}\mathbb{E}[\boldsymbol{\varphi}_{\bar T}^{i}\boldsymbol{\varphi}_{\bar T}^{i\prime}]\frac{\boldsymbol{\iota}_{\bar T}}{\sqrt{\bar T}}\\
&\le q\bar c^2 \max_{k=1,\dots,q}\max_{\substack{\boldsymbol{c}_{\bar T}\\ \boldsymbol{c}_{\bar T}'\boldsymbol{c}_{\bar T}=1}}\boldsymbol{c}_{\bar T}'\mathbb{E}[\boldsymbol{\varphi}_{\bar T}^{i}\boldsymbol{\varphi}_{\bar T}^{i\prime}]\boldsymbol{c}_{\bar T}\\
&\le q\bar c^2 \max_{k=1,\dots,q}\max_{\substack{\boldsymbol{c}_{\bar T}\\ \boldsymbol{c}_{\bar T}'\boldsymbol{c}_{\bar T}=1}}\sup_{n\in\mathbb{N}}\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{c}_{\bar T}'\mathbb{E}[\boldsymbol{\varphi}_{\bar T}^{i}\boldsymbol{\varphi}_{\bar T}^{i\prime}]\boldsymbol{c}_{\bar T}\\
&\le q\bar c^2 \max_{k=1,\dots,q}\sup_{\bar T\in\mathbb{N}}\max_{\substack{\boldsymbol{c}_{\bar T}\\ \boldsymbol{c}_{\bar T}'\boldsymbol{c}_{\bar T}=1}}\boldsymbol{c}_{\bar T}'\boldsymbol{G}_{\bar T}^{\phi}\boldsymbol{c}_{\bar T} \le q\bar c^2 2\pi B^{\phi},
\end{aligned}
$$

because of (37) and since $\boldsymbol{G}_{\bar T}^{\phi} = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[\boldsymbol{\varphi}_{\bar T}^{i}\boldsymbol{\varphi}_{\bar T}^{i\prime}]$. From Chebychev's inequality, $\|\widetilde{\boldsymbol{\Pi}}_{\bar T}^{\psi\prime}\boldsymbol{\varphi}_{\bar T}^{i}\| = O_{\mathrm{P}}(1)$ and, therefore, $\|\boldsymbol{\Pi}_{\bar T}^{\psi\prime}\boldsymbol{\varphi}_{\bar T}^{i}\| = O_{\mathrm{P}}\left(\sqrt{\frac{\bar T}{T}}\right)$. □

B.2 Proof of Theorem 2

The proof is entirely the same as for Theorem 1, with Lemmas 11–18 replacing Lemmas 3–10. □

C Proof of Theorem 3

Proof. First, for any $i = 1,\dots,n$, we have, from the proof of Theorem 1, $\mathbf{R}_i' = \mathbf{p}_i^{\psi\prime}(\boldsymbol{\Lambda}_n^{\psi})^{1/2}(\boldsymbol{\Gamma}^u)^{-1/2}$. Therefore, from the definition of $\check{\mathbf{R}}_i'$ in (39),

$$
\begin{aligned}
\check{\mathbf{R}}_i' - \mathbf{R}_i'(\boldsymbol{\Gamma}^u)^{1/2}\widehat{\mathbf{W}}^z &= \widehat{\mathbf{p}}_i^{z\prime}\left(\widehat{\boldsymbol{\Lambda}}_n^{z}\right)^{1/2} - \mathbf{p}_i^{\psi\prime}\left(\boldsymbol{\Lambda}_n^{\psi}\right)^{1/2}\widehat{\mathbf{W}}^z = \widehat{\mathbf{p}}_i^{z\prime}\left(\widehat{\boldsymbol{\Lambda}}_n^{z}\right)^{1/2} - \mathbf{p}_i^{\psi\prime}\widehat{\mathbf{W}}^z\left(\boldsymbol{\Lambda}_n^{\psi}\right)^{1/2}\\
&= \mathbf{p}_i^{\psi\prime}\widehat{\mathbf{W}}^z\left(\widehat{\boldsymbol{\Lambda}}_n^{z} - \boldsymbol{\Lambda}_n^{\psi}\right)^{1/2} + \left(\widehat{\mathbf{p}}_i^{z\prime} - \mathbf{p}_i^{\psi\prime}\widehat{\mathbf{W}}^z\right)\left(\boldsymbol{\Lambda}_n^{\psi}\right)^{1/2} + \left(\widehat{\mathbf{p}}_i^{z\prime} - \mathbf{p}_i^{\psi\prime}\widehat{\mathbf{W}}^z\right)\left(\widehat{\boldsymbol{\Lambda}}_n^{z} - \boldsymbol{\Lambda}_n^{\psi}\right)^{1/2}\\
&= I + II + III, \quad \text{say.}
\end{aligned} \tag{73}
$$

Term I is $O_{\mathrm{P}}(\zeta_{nT})$ because of Lemmas 5(i) and 8, term II is $O_{\mathrm{P}}\left(\sqrt{\frac{\bar n}{n}}\zeta_{nT}\right)$ because of of Lemmas 5(ii) and 8 (see also the arguments in Lemma 6 in Forni et al., 2017), and term III is $o_{\mathrm{P}}(\zeta_{nT})$. From (73), we get

$$
\left\|\check{\mathbf{R}}_i' - \mathbf{R}_i'(\boldsymbol{\Gamma}^u)^{1/2}\widehat{\mathbf{W}}^z\right\| = O_{\mathrm{P}}(\zeta_{nT}) \tag{74}
$$

which, combined with Theorem 1(i), gives

$$
\left\|\check{\mathbf{R}}_i'\widehat{\mathbf{u}}_t - \mathbf{R}_i'\mathbf{u}_t\right\| = O_{\mathrm{P}}\left(\max\left(\frac{1}{\sqrt{\bar n}},\zeta_{nT}\right)\right) = O_{\mathrm{P}}\left(\max\left(\frac{1}{\sqrt{\bar h}},\zeta_{nT}\right)\right). \tag{75}
$$

Following a reasoning similar to (73), since, from the proof of Theorem 2, for any $t = 1,\dots,T$ we have $\mathbf{u}_t' = \boldsymbol{\pi}_t^{\psi\prime}(\boldsymbol{L}_T^{\psi})^{1/2}(\boldsymbol{\Sigma}^R)^{-1/2}$, the definition of $\check{\mathbf{u}}_t'$ in (42) and Lemmas 13 and 16 imply that

$$\left\|\check{\mathbf{u}}_t' - \mathbf{u}_t'\left(\boldsymbol{\Sigma}^R\right)^{1/2}\widehat{\boldsymbol{W}}^z\right\| = O_{\mathrm{P}}(\zeta_{nT}) \tag{76}$$

which, combined with Theorem 2(i), yields

$$\left\|\check{\mathbf{u}}_t'\widehat{\mathbf{R}}_i - \mathbf{u}_t'\mathbf{R}_i\right\| = O_{\mathrm{P}}\left(\max\left(\frac{1}{\sqrt{\bar{n}}}, \zeta_{nT}\right)\right) = O_{\mathrm{P}}\left(\max\left(\frac{1}{\sqrt{\bar{h}}}, \zeta_{nT}\right)\right). \tag{77}$$

Part (i) of the theorem filliws from (75) and (77).

Now, from the proof of Theorems 1 and 2 and using (74) and (76),

$$\begin{aligned}
\check{\mathbf{R}}_i'\widehat{\mathbf{u}}_t &= \mathbf{R}_i'\mathbf{u}_t + \mathbf{R}_i'\left(\boldsymbol{\Gamma}^u\right)^{1/2}\widehat{\mathbf{W}}^z\left(\widehat{\mathbf{u}}_t - \widehat{\mathbf{W}}^z\left(\boldsymbol{\Gamma}^u\right)^{-1/2}\mathbf{u}_t\right) + \left(\check{\mathbf{R}}_i' - \mathbf{R}_i'\left(\boldsymbol{\Gamma}^u\right)^{1/2}\widehat{\mathbf{W}}^z\right)\widehat{\mathbf{u}}_t \\
&= \mathbf{R}_i'\mathbf{u}_t + \mathbf{R}_i'\left(\boldsymbol{\Gamma}^u\right)^{1/2}\widehat{\mathbf{W}}^z\left(\boldsymbol{\Lambda}_n^{\psi}\right)^{-1/2}\left(\mathbf{P}_{\bar{n}}^{\psi\prime}\mathbf{P}_{\bar{n}}^{\psi}\right)^{-1}\mathbf{P}_{\bar{n}}^{\psi\prime}\boldsymbol{\phi}_{\bar{n}t} + \left(\check{\mathbf{R}}_i' - \mathbf{R}_i'\left(\boldsymbol{\Gamma}^u\right)^{1/2}\widehat{\mathbf{W}}^z\right)\widehat{\mathbf{u}}_t + O_{\mathrm{P}}(\zeta_{nT}) \\
&= \mathbf{R}_i'\mathbf{u}_t + \mathbf{R}_i'\left(\boldsymbol{\Gamma}^u\right)^{1/2}\widehat{\mathbf{W}}^z\left(\boldsymbol{\Lambda}_n^{\psi}\right)^{-1/2}\left(\mathbf{P}_{\bar{n}}^{\psi\prime}\mathbf{P}_{\bar{n}}^{\psi}\right)^{-1}\mathbf{P}_{\bar{n}}^{\psi\prime}\boldsymbol{\phi}_{\bar{n}t} + O_{\mathrm{P}}(\zeta_{nT}),
\end{aligned} \tag{78}$$

and

$$\begin{aligned}
\check{\mathbf{u}}_t'\widehat{\mathbf{R}}_i &= \mathbf{u}_t'\mathbf{R}_i + \mathbf{u}_t'\left(\boldsymbol{\Sigma}^R\right)^{1/2}\widehat{\boldsymbol{W}}^z\left(\widehat{\mathbf{R}}_i - \widehat{\boldsymbol{W}}^z\left(\boldsymbol{\Sigma}^R\right)^{1/2}\mathbf{R}_i\right) + \left(\check{\mathbf{u}}_t' - \mathbf{u}_t'\left(\boldsymbol{\Sigma}^R\right)^{1/2}\widehat{\boldsymbol{W}}^z\right)\widehat{\mathbf{R}}_i \\
&= \mathbf{u}_t'\mathbf{R}_i + \mathbf{u}_t'\left(\boldsymbol{\Sigma}^R\right)^{1/2}\widehat{\boldsymbol{W}}^z\left(\boldsymbol{L}_T^{\psi}\right)^{-1/2}\left(\boldsymbol{\Pi}_{\bar{T}}^{\psi\prime}\boldsymbol{\Pi}_{\bar{T}}^{\psi}\right)^{-1}\boldsymbol{\Pi}_{\bar{T}}^{\psi\prime}\boldsymbol{\varphi}_{\bar{T}}^{i} + \left(\check{\mathbf{u}}_t' - \mathbf{u}_t'\left(\boldsymbol{\Sigma}^R\right)^{1/2}\widehat{\boldsymbol{W}}^z\right)\widehat{\mathbf{R}}_i + O_{\mathrm{P}}(\zeta_{nT}) \\
&= \mathbf{u}_t'\mathbf{R}_i + \mathbf{u}_t'\left(\boldsymbol{\Sigma}^R\right)^{1/2}\widehat{\boldsymbol{W}}^z\left(\boldsymbol{L}_T^{\psi}\right)^{-1/2}\left(\boldsymbol{\Pi}_{\bar{T}}^{\psi\prime}\boldsymbol{\Pi}_{\bar{T}}^{\psi}\right)^{-1}\boldsymbol{\Pi}_{\bar{T}}^{\psi\prime}\boldsymbol{\varphi}_{\bar{T}}^{i} + O_{\mathrm{P}}(\zeta_{nT}),
\end{aligned} \tag{79}$$

since $\|\widehat{\mathbf{u}}_t\| = O_{\mathrm{P}}(1)$ and $\|\widehat{\mathbf{R}}_i\| = O_{\mathrm{P}}(1)$.

From Theorem 1, (78), and because of (57), as $n, T \to \infty$,

$$\begin{aligned}
\sqrt{n}\left(\check{\mathbf{R}}_i'\widehat{\mathbf{u}}_t - \mathbf{R}_i'\mathbf{u}_t\right) &= \sqrt{n}\,\mathbf{R}_i'\left(\boldsymbol{\Gamma}^u\right)^{1/2}\widehat{\mathbf{W}}^z\left(\boldsymbol{\Lambda}_n^{\psi}\right)^{-1/2}\left(\mathbf{P}_{\bar{n}}^{\psi\prime}\mathbf{P}_{\bar{n}}^{\psi}\right)^{-1}\mathbf{P}_{\bar{n}}^{\psi\prime}\boldsymbol{\phi}_{\bar{n}t} + o_{\mathrm{P}}(1) \\
&= \mathbf{R}_i'\left(\boldsymbol{\Gamma}^u\right)^{1/2}\widehat{\mathbf{W}}^z\left(\frac{\boldsymbol{\Lambda}_n^{\psi}}{n}\right)^{-1/2}\left(\frac{n}{\bar{n}}\mathbf{P}_{\bar{n}}^{\psi\prime}\mathbf{P}_{\bar{n}}^{\psi}\right)^{-1}\sqrt{\frac{n}{\bar{n}}}\left(\mathbf{P}_{\bar{n}}^{\psi\prime}\boldsymbol{\phi}_{\bar{n}t}\right) + o_{\mathrm{P}}(1) \\
&\longrightarrow_d \mathcal{N}\left(\mathbf{0}_q, \mathbf{R}_i'\left(\boldsymbol{\Gamma}^u\right)^{1/2}\mathbf{W}^u\left(\boldsymbol{\mathcal{L}}^u\right)^{-1/2}\left(\boldsymbol{\mathcal{M}}^u\right)^{-1}\boldsymbol{\mathcal{P}}_t^u\left(\boldsymbol{\mathcal{M}}^u\right)^{-1}\left(\boldsymbol{\mathcal{L}}^u\right)^{-1/2}\mathbf{W}^u\left(\boldsymbol{\Gamma}^u\right)^{1/2}\mathbf{R}_i\right),
\end{aligned}$$

where $\mathbf{W}^u = \mathrm{plim}_{n,T\to\infty}\widehat{\mathbf{W}}^z$ as defined in Theorem 1.

Likewise, from Theorem 2, (79), and because of (57), as $n, T \to \infty$,

$$\begin{aligned}
\sqrt{\bar{T}}\left(\check{\mathbf{u}}_t'\widehat{\mathbf{R}}_i - \mathbf{u}_t'\mathbf{R}_i\right) &= \sqrt{\bar{T}}\,\mathbf{u}_t'\left(\boldsymbol{\Sigma}^R\right)^{1/2}\widehat{\boldsymbol{W}}^z\left(\boldsymbol{L}_T^{\psi}\right)^{-1/2}\left(\boldsymbol{\Pi}_{\bar{T}}^{\psi\prime}\boldsymbol{\Pi}_{\bar{T}}^{\psi}\right)^{-1}\boldsymbol{\Pi}_{\bar{T}}^{\psi\prime}\boldsymbol{\varphi}_{\bar{T}}^{i} + o_{\mathrm{P}}(1) \\
&= \mathbf{u}_t'\left(\boldsymbol{\Sigma}^R\right)^{1/2}\widehat{\boldsymbol{W}}^z\left(\frac{\boldsymbol{L}_T^{\psi}}{T}\right)^{-1/2}\left(\frac{T}{\bar{T}}\boldsymbol{\Pi}_{\bar{T}}^{\psi\prime}\boldsymbol{\Pi}_{\bar{T}}^{\psi}\right)^{-1}\sqrt{\frac{T}{\bar{T}}}\left(\boldsymbol{\Pi}_{\bar{T}}^{\psi\prime}\boldsymbol{\varphi}_{\bar{T}}^{i}\right) + o_{\mathrm{P}}(1) \\
&\longrightarrow_d \mathcal{N}\left(\mathbf{0}_q, \mathbf{u}_t'\left(\boldsymbol{\Sigma}^R\right)^{1/2}\mathbf{W}^R\left(\boldsymbol{\mathcal{L}}^R\right)^{-1/2}\left(\boldsymbol{\mathcal{M}}^R\right)^{-1}\boldsymbol{\mathcal{P}}_i^R\left(\boldsymbol{\mathcal{M}}^R\right)^{-1}\left(\boldsymbol{\mathcal{L}}^R\right)^{-1/2}\mathbf{W}^R\left(\boldsymbol{\Sigma}^R\right)^{1/2}\mathbf{u}_t\right),
\end{aligned}$$

where $\mathbf{W}^R = \mathrm{plim}_{n,T\to\infty}\widehat{\boldsymbol{W}}^z$ as defined in Theorem 2.

Moreover, defining

$$\boldsymbol{\Omega}_{it} := \lim_{n,T\to\infty} \left(\frac{\sqrt{nT}}{\sqrt{\bar{n}\bar{T}}}\right) \mathbb{E}\left[\mathbf{P}_{\bar{n}}^{\psi\prime}\boldsymbol{\phi}_{\bar{n}t}\boldsymbol{\varphi}_{\bar{T}}^{i\prime}\boldsymbol{\Pi}_{\bar{T}}^{\psi}\right].$$

as $n, T \to \infty$, when $\bar{h} = \bar{n} = \bar{T}$,

$$\sqrt{\frac{n}{\bar{n}}}\mathbf{P}_{\bar{n}}^{\phi\prime}\boldsymbol{\phi}_{\bar{n}t} + \sqrt{\frac{T}{\bar{T}}}\boldsymbol{\Pi}_{\bar{T}}^{\phi\prime}\boldsymbol{\varphi}_{T}^{i} \to_d N\left(\mathbf{0}_q, \boldsymbol{\mathcal{P}}_t^u + \boldsymbol{\mathcal{P}}_i^R + \boldsymbol{\Omega}_{it} + \boldsymbol{\Omega}_{it}'\right).$$

Therefore, as $n, T \to \infty$,

$$\sqrt{\bar{h}}\left(\left(\omega_{nT}\check{\mathbf{R}}_i'\widehat{\mathbf{u}}_t + (1-\omega_{nT})\check{\mathbf{u}}_t'\widehat{\mathbf{R}}_i\right) - \mathbf{R}_i'\mathbf{u}_t\right) \to_d N\left(0, \omega^2 V_{it}^u + (1-\omega)^2 V_{it}^R + 2\omega(1-\omega)C_{it}\right),$$

where

$$\begin{aligned}
V_{it}^u &= \mathbf{R}_i'\left(\boldsymbol{\Gamma}^u\right)^{1/2}\mathbf{W}^u\left(\boldsymbol{\mathcal{L}}^u\right)^{-1/2}\left(\boldsymbol{\mathcal{M}}^u\right)^{-1}\boldsymbol{\mathcal{P}}_t^u\left(\boldsymbol{\mathcal{M}}^u\right)^{-1}\left(\boldsymbol{\mathcal{L}}^u\right)^{-1/2}\mathbf{W}^u\left(\boldsymbol{\Gamma}^u\right)^{1/2}\mathbf{R}_i,\\
V_{it}^R &= \mathbf{u}_t'\left(\boldsymbol{\Sigma}^R\right)^{1/2}\mathbf{W}^R\left(\boldsymbol{\mathcal{L}}^R\right)^{-1/2}\left(\boldsymbol{\mathcal{M}}^R\right)^{-1}\boldsymbol{\mathcal{P}}_i^R\left(\boldsymbol{\mathcal{M}}^R\right)^{-1}\left(\boldsymbol{\mathcal{L}}^R\right)^{-1/2}\mathbf{W}^R\left(\boldsymbol{\Sigma}^R\right)^{1/2}\mathbf{u}_t,\\
C_{it} &= \mathbf{R}_i'\left(\boldsymbol{\Gamma}^u\right)^{1/2}\mathbf{W}^u\left(\boldsymbol{\mathcal{L}}^u\right)^{-1/2}\left(\boldsymbol{\mathcal{M}}^u\right)^{-1}\boldsymbol{\Omega}_{it}\left(\boldsymbol{\mathcal{M}}^R\right)^{-1}\left(\boldsymbol{\mathcal{L}}^R\right)^{-1/2}\mathbf{W}^R\left(\boldsymbol{\Sigma}^R\right)^{1/2}\mathbf{u}_t.
\end{aligned} \qquad \square$$

D Proof of Theorem 4

Let $\mathbf{C}_n(L) := [\mathbf{A}_n(L)]^{-1}$ and $\widehat{\mathbf{C}}_n(L) := [\widehat{\mathbf{A}}_n(L)]^{-1}$. Then, for any $i = 1, \dots, n$ and $t = 1, \dots, T$,

$$\begin{aligned}
\widehat{\chi}_{it} - \chi_{it} &= \mathbf{e}_i'\left(\widehat{\mathbf{C}}_n(L)\widehat{\boldsymbol{\psi}}_{nt} - \mathbf{C}_n(L)\boldsymbol{\psi}_{nt}\right)\\
&= \mathbf{e}_i'\left(\widehat{\mathbf{C}}_n(L) - \mathbf{C}_n(L)\right)\boldsymbol{\psi}_{nt} + \mathbf{e}_i'\mathbf{C}_n(L)\left(\widehat{\boldsymbol{\psi}}_{nt} - \boldsymbol{\psi}_{nt}\right) + \mathbf{e}_i'\left(\widehat{\mathbf{C}}_n(L) - \mathbf{C}_n(L)\right)\left(\widehat{\boldsymbol{\psi}}_{nt} - \boldsymbol{\psi}_{nt}\right)\\
&= I + II + III, \quad \text{say,}
\end{aligned} \tag{80}$$

where $\mathbf{e}_i$ denotes the ith column of $\mathbf{I}_n$.

From Proposition 2, we have, for any $s = 1, \dots, m$, $j_s = 1, \dots, (q+1)$, and $h_s = 1, \dots, (q+1)$, as $n, T \to \infty$,

$$\max_{\ell=1,\dots,p_s}\ \max_{j_s,h_s=1,\dots,(q+1)} (\widehat{a}_{j_s,h_s,\ell} - a_{j_s,h_s,\ell})^2 \le \left\|\widehat{\mathbf{A}}^{[s]} - \mathbf{A}^{[s]}\right\|^2 = O_{\mathrm{P}}(\zeta_{n,T}^2), \tag{81}$$

where $a_{j_s,h_s,\ell}$ and $\widehat{a}_{j_s,h_s,\ell}$ are the (j,h)th entries of $\mathbf{A}_\ell^{(i)}$ and of $\widehat{\mathbf{A}}_\ell^{(i)}$, respectively.

Without loss of generality, let us assume $p_s = 1$ for all $s = 1, \dots, m$, so that $\mathbf{A}_n(L) = \mathbf{I}_n - \mathbf{A}_n L$ and $\widehat{\mathbf{A}}_n(L) = \mathbf{I}_n - \widehat{\mathbf{A}}_n L$. Thus, $\mathbf{C}_n(L) = \sum_{k=0}^\infty \mathbf{A}_n^k$ and $\widehat{\mathbf{C}}_n(L) = \sum_{k=0}^\infty \widehat{\mathbf{A}}_n^k$. Then, for any $i = 1, \dots, n$, there exists an $s \in \{1, \dots, m\}$ such that χ_{it} is an element of the sth $(q+1)$-dimensional subvector $\boldsymbol{\chi}_t^{(s)}$ of $\boldsymbol{\chi}_{nt}$. Let $c_{i,j_s,k}$ and $\widehat{c}_{i,j_s,k}$ denote the (i, j_s)th entries of $\mathbf{A}_n^k$ and $\widehat{\mathbf{A}}_n^k$, respectively (here j_s indicates the jth column of block s of $\mathbf{A}_n^k$ and $\widehat{\mathbf{A}}_n^k$).

Assumption (A2*(a)*), which holds with probability one in view of Assumption (S*(a)*) implies summability of the autoregressive coefficients, for any $i = 1, \dots, n$ and $t = 1, \dots, T$ and, for any $\epsilon > 0$ and $\eta > 0$, the existence of a constant $K = K(\epsilon, \eta)$ independent of i, j_s, s, and t such that

$$\mathrm{P}\left(\left| \sum_{j_s=1}^{q+1} \sum_{k=K+1}^{\infty} \left(\widehat{c}_{i,j_s,k} - c_{i,j_s,k}\right) \psi_{j_s,t-k} \right| > \eta \right) \leq \epsilon.$$

Hence, we can select K such that

$$\left| \sum_{j_s=1}^{q+1} \sum_{k=K+1}^{\infty} \left(\widehat{c}_{i,j_s,k} - c_{i,j_s,k}\right) \psi_{j_s,t-k} \right| = o_{\mathrm{P}}(\zeta_{nT}).$$

Then, the norm of I is such that

$$\begin{aligned}
\left| \mathbf{e}_i' \left(\widehat{\mathbf{C}}_n(L) - \mathbf{C}_n(L) \right) \boldsymbol{\psi}_{nt} \right| &\leq \sum_{k=0}^{K} \left(\sum_{j_s=1}^{q+1} \left(\widehat{c}_{i,j_s,k} - c_{i,j_s,k}\right)^2 \psi_{j_s,t-k}^2 \right)^{1/2} + o_{\mathrm{P}}(\zeta_{nT}) \\
&\leq \sum_{k=0}^{K} \left(\sum_{j_s=1}^{q+1} \psi_{j_s,t-k}^4 \right)^{1/4} \left(\sum_{j_s=1}^{q+1} \left(\widehat{c}_{i,j_s,k} - c_{i,j_s,k}\right)^4 \right)^{1/4} + o_{\mathrm{P}}(\zeta_{nT}) \\
&= O_{\mathrm{P}}(\zeta_{nT}),
\end{aligned} \tag{82}$$

because of (81) and the continuous mapping theorem.

Similarly, for the norm of II and because of Assumption (A2*(a)*), we can select K such that

$$\left| \sum_{j_s=1}^{q+1} \sum_{k=K+1}^{\infty} c_{i,j_s,k} \left(\widehat{\psi}_{j_s,t-k} - \psi_{j_s,t-k} \right) \right| = o_{\mathrm{P}}(\zeta_{nT}) \tag{83}$$

and, therefore, by Theorem 3, when $\bar{h} = \bar{n} = \bar{T}$,

$$\begin{aligned}
\left| \mathbf{e}_i' \mathbf{C}_n(L) \left(\widehat{\boldsymbol{\psi}}_{nt} - \boldsymbol{\psi}_{nt} \right) \right| &\leq \sum_{k=0}^{K} \left(\sum_{j_s=1}^{q+1} c_{i,j_s,k}^2 \left(\widehat{\psi}_{j_s,t-k} - \psi_{j_s,t-k} \right)^2 \right)^{1/2} + o_{\mathrm{P}}(\zeta_{nT}) \\
&\leq \sum_{k=0}^{K} \left(\sum_{j_s=1}^{q+1} c_{i,j_s,k}^4 \right)^{1/4} \left(\sum_{j_s=1}^{q+1} \left(\widehat{\psi}_{j_s,t-k} - \psi_{j_s,t-k} \right)^4 \right)^{1/4} + o_{\mathrm{P}}(\zeta_{nT}) \\
&= O_{\mathrm{P}}\left(\max\left(\frac{1}{\sqrt{h}}, \zeta_{nT} \right) \right) + o_{\mathrm{P}}(\zeta_{nT}).
\end{aligned} \tag{84}$$

Obviously $III = o_{\mathrm{P}}(\zeta_{nT})$. Therefore, substituting (82) and (84) into (80), we prove consistency.

Now, from Theorem 1, for any finite $k \in \mathbb{N}$ such that $k < T$ and any $t = k+1, \ldots, T$, as $n, T \to \infty$,

$$\sqrt{n} \left(\begin{pmatrix} \widehat{\mathbf{u}}_t \\ \vdots \\ \widehat{\mathbf{u}}_{t-k} \end{pmatrix} - \begin{pmatrix} \mathbf{u}_t \\ \vdots \\ \mathbf{u}_{t-k} \end{pmatrix} \right) \longrightarrow_d \mathcal{N}\left(\mathbf{0}_{qk}, \boldsymbol{\mathcal{V}}^u_{t\ldots t-k} \right), \tag{85}$$

where

$$\boldsymbol{\mathcal{V}}^u_{t\dots t-k} = \left\{\mathbf{I}_k \otimes \left[\mathbf{W}^u \left(\boldsymbol{\mathcal{L}}^u\right)^{-1/2} \left(\boldsymbol{\mathcal{M}}^u\right)^{-1}\right]\right\} \boldsymbol{\mathcal{P}}^u_{t\dots t-k} \left\{\mathbf{I}_k \otimes \left[\left(\boldsymbol{\mathcal{M}}^u\right)^{-1} \left(\boldsymbol{\mathcal{L}}^u\right)^{-1/2} \mathbf{W}^u\right]\right\}$$

and

$$\boldsymbol{\mathcal{P}}^u_{t\dots t-k} = \lim_{n\to\infty} \frac{n}{\bar{n}} \mathbb{E}\left[\left\{\mathbf{I}_k \otimes \mathbf{P}^{\psi\prime}_{\bar{n}}\right\} \begin{pmatrix} \boldsymbol{\phi}_{\bar{n}t} \\ \vdots \\ \boldsymbol{\phi}_{\bar{n}t-k} \end{pmatrix} \begin{pmatrix} \boldsymbol{\phi}_{\bar{n}t} \\ \vdots \\ \boldsymbol{\phi}_{\bar{n}t-k} \end{pmatrix}' \left\{\mathbf{I}_k \otimes \mathbf{P}^{\psi\prime}_{\bar{n}}\right\}'\right].$$

Similarly, from Theorem 2, for any finite $\ell \in \mathbb{N}$ such that $\{i_1, \dots, i_\ell\} \subset \{1, \dots, n\}$, as $n, T \to \infty$,

$$\sqrt{\bar{T}} \left(\begin{pmatrix} \widehat{\mathbf{R}}_{i_1} \\ \vdots \\ \widehat{\mathbf{R}}_{i_\ell} \end{pmatrix} - \begin{pmatrix} \mathbf{R}_{i_1} \\ \vdots \\ \mathbf{R}_{i_\ell} \end{pmatrix} \right) \longrightarrow_d \mathcal{N}\left(\mathbf{0}_{q\ell}, \boldsymbol{\mathcal{V}}^R_{i_1\dots i_\ell}\right), \tag{86}$$

where

$$\boldsymbol{\mathcal{V}}^R_{i_1\dots i_\ell} = \left\{\mathbf{I}_\ell \otimes \left[\mathbf{W}^R \left(\boldsymbol{\mathcal{L}}^R\right)^{-1/2} \left(\boldsymbol{\mathcal{M}}^R\right)^{-1}\right]\right\} \boldsymbol{\mathcal{P}}^R_{i_1\dots i_\ell} \left\{\mathbf{I}_\ell \otimes \left[\left(\boldsymbol{\mathcal{M}}^R\right)^{-1} \left(\boldsymbol{\mathcal{L}}^R\right)^{-1/2} \mathbf{W}^R\right]\right\}$$

and

$$\boldsymbol{\mathcal{P}}^R_{i_1\dots i_\ell} = \lim_{T\to\infty} \frac{T}{\bar{T}} \mathbb{E}\left[\left\{\mathbf{I}_\ell \otimes \boldsymbol{\Pi}^{\psi\prime}_{\bar{T}}\right\} \begin{pmatrix} \boldsymbol{\varphi}^{i_1}_{\bar{T}} \\ \vdots \\ \boldsymbol{\varphi}^{i_\ell}_{\bar{T}} \end{pmatrix} \begin{pmatrix} \boldsymbol{\varphi}^{i_1}_{\bar{T}} \\ \vdots \\ \boldsymbol{\varphi}^{i_\ell}_{\bar{T}} \end{pmatrix}' \left\{\mathbf{I}_\ell \otimes \boldsymbol{\Pi}^{\psi\prime}_{\bar{T}}\right\}'\right].$$

For any $i = 1, \dots, n$, define the $(q+1) \times q(K+1)$ matrix

$$\boldsymbol{\mathcal{C}}_i := \begin{pmatrix} \boldsymbol{\iota}_q' \otimes c_{i,1,0} & \dots & \boldsymbol{\iota}_q' \otimes c_{i,1,K} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\iota}_q' \otimes c_{i,q+1,0} & \dots & \boldsymbol{\iota}_q' \otimes c_{i,q+1,K} \end{pmatrix},$$

and the $(K+1) \times q(q+1)$ matrix

$$\boldsymbol{\mathcal{D}}_i := \begin{pmatrix} \boldsymbol{\iota}_q' \otimes c_{i,1,0} & \dots & \boldsymbol{\iota}_q' \otimes c_{i,q+1,0} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\iota}_q' \otimes c_{i,1,K} & \dots & \boldsymbol{\iota}_q' \otimes c_{i,q+1,K} \end{pmatrix},$$

where $\boldsymbol{\iota}_q$ is a q-dimensional vector of ones. For given $i = 1, \dots, n$, let $\mathbf{R}'_{i_{j_s}}$ be the row of $\boldsymbol{\mathcal{R}}_n$ corresponding to the j_sth series in block s, which is the block to which series i belongs. Then, from (80), (85), (86), and given K as defined in (83), for any $i = 1, \dots, n$ and $t = 1, \dots, T$, as $n, T \to \infty$,

$$
\begin{aligned}
\sqrt{\bar{h}}\,(\widehat{\chi}_{it} - \chi_{it}) &= \sqrt{\bar{h}} \sum_{k=0}^{K} \sum_{j_s=1}^{q+1} c_{i,j_s,k} \left(\widehat{\psi}_{j_s,t-k} - \psi_{j_s,t-k} \right) + o_{\mathrm{P}}(1) \\
&= \sqrt{\bar{h}} \sum_{k=0}^{K} \sum_{j_s=1}^{q+1} c_{i,j_s,k} \left(\omega_{nT} \mathbf{R}'_{i_{j_s}} \widehat{\mathbf{u}}_{t-k} + (1-\omega_{nT}) \mathbf{u}'_{t-k} \widehat{\mathbf{R}}_{i_{j_s}} - \mathbf{R}'_{i_{j_s}} \mathbf{u}_{t-k} \right) + o_{\mathrm{P}}(1) \\
&= \sqrt{\bar{h}} \sum_{k=0}^{K} \sum_{j_s=1}^{q+1} \left\{ \omega_{nT}\, c_{i,j_s,k} \mathbf{R}'_{i_{j_s}} \left(\widehat{\mathbf{u}}_{t-k} - \mathbf{u}_{t-k} \right) + (1-\omega_{nT})\, c_{i,j_s,k} \mathbf{u}'_{t-k} \left(\widehat{\mathbf{R}}_{i_{j_s}} - \mathbf{R}_{i_{j_s}} \right) \right\} + o_{\mathrm{P}}(1) \\
&\longrightarrow_d \mathcal{N}\left(0, \boldsymbol{\omega}' \begin{pmatrix} W^u_{it} & G_{it} \\ G_{it} & W^R_{it} \end{pmatrix} \boldsymbol{\omega} \right)
\end{aligned}
$$

where $\boldsymbol{\omega} = \lim_{nT\to\infty} \begin{pmatrix} \omega_{nT} \\ 1-\omega_{nT} \end{pmatrix}$,

$$
W^u_{it} = \boldsymbol{\iota}'_{q+1} \left\{ \boldsymbol{\mathcal{C}}_i \odot \left[\boldsymbol{\iota}'_{K+1} \otimes \begin{pmatrix} \mathbf{R}'_{i_1} (\boldsymbol{\Gamma}^u)^{1/2} \\ \vdots \\ \mathbf{R}'_{i_{q+1}} (\boldsymbol{\Gamma}^u)^{1/2} \end{pmatrix} \right] \right\} \boldsymbol{\mathcal{V}}^u_{t\dots t-k} \left\{ \boldsymbol{\mathcal{C}}_i \odot \left[\boldsymbol{\iota}'_{K+1} \otimes \begin{pmatrix} \mathbf{R}'_{i_1} (\boldsymbol{\Gamma}^u)^{1/2} \\ \vdots \\ \mathbf{R}'_{i_{q+1}} (\boldsymbol{\Gamma}^u)^{1/2} \end{pmatrix} \right] \right\}' \boldsymbol{\iota}_{q+1},
$$

$$
W^R_{it} = \boldsymbol{\iota}'_{K+1} \left\{ \boldsymbol{\mathcal{D}}_i \odot \left[\boldsymbol{\iota}'_{q+1} \otimes \begin{pmatrix} \mathbf{u}'_t (\boldsymbol{\Sigma}^R)^{1/2} \\ \vdots \\ \mathbf{u}'_{t-K} (\boldsymbol{\Sigma}^R)^{1/2} \end{pmatrix} \right] \right\} \boldsymbol{\mathcal{V}}^R_{i_1\dots i_{q+1}} \left\{ \boldsymbol{\mathcal{D}}_i \odot \left[\boldsymbol{\iota}'_{q+1} \otimes \begin{pmatrix} \mathbf{u}'_t (\boldsymbol{\Sigma}^R)^{1/2} \\ \vdots \\ \mathbf{u}'_{t-K} (\boldsymbol{\Sigma}^R)^{1/2} \end{pmatrix} \right] \right\}' \boldsymbol{\iota}_{K+1},
$$

$$
G_{it} = \boldsymbol{\iota}'_{q+1} \left\{ \boldsymbol{\mathcal{C}}_i \odot \left[\boldsymbol{\iota}'_{K+1} \otimes \begin{pmatrix} \mathbf{R}'_{i_1} (\boldsymbol{\Gamma}^u)^{1/2} \\ \vdots \\ \mathbf{R}'_{i_{q+1}} (\boldsymbol{\Gamma}^u)^{1/2} \end{pmatrix} \right] \right\} \boldsymbol{\mathcal{O}}_{\substack{i_1\dots i_{q+1} \\ t\dots t-K}} \left\{ \boldsymbol{\mathcal{D}}_i \odot \left[\boldsymbol{\iota}'_{q+1} \otimes \begin{pmatrix} \mathbf{u}'_t (\boldsymbol{\Sigma}^R)^{1/2} \\ \vdots \\ \mathbf{u}'_{t-K} (\boldsymbol{\Sigma}^R)^{1/2} \end{pmatrix} \right] \right\}' \boldsymbol{\iota}_{K+1},
$$

with

$$
\begin{aligned}
\boldsymbol{\mathcal{O}}_{\substack{i_1\dots i_{q+1} \\ t\dots t-K}} &= \left\{ \mathbf{I}_{K+1} \otimes \mathbf{W}^u \left(\boldsymbol{\mathcal{L}}^u \right)^{-1/2} \left(\boldsymbol{\mathcal{M}}^u \right)^{-1} \right\} \boldsymbol{\Omega}_{\substack{i_1\dots i_{q+1} \\ t\dots t-K}} \left\{ \mathbf{I}_{q+1} \otimes \mathbf{W}^R \left(\boldsymbol{\mathcal{L}}^R \right)^{-1/2} \left(\boldsymbol{\mathcal{M}}^R \right)^{-1} \right\}', \\
\boldsymbol{\Omega}_{\substack{i_1\dots i_{q+1} \\ t\dots t-K}} &= \lim_{n,T\to\infty} \frac{\sqrt{nT}}{\sqrt{\bar{n}\bar{T}}} \mathbb{E} \left[\left\{ \mathbf{I}_{K+1} \otimes \mathbf{P}^{\psi\prime}_{\bar{n}} \right\} \begin{pmatrix} \boldsymbol{\phi}_{\bar{n}t} \\ \vdots \\ \boldsymbol{\phi}_{\bar{n}t-K} \end{pmatrix} \begin{pmatrix} \boldsymbol{\varphi}^{i_1}_{\bar{T}} \\ \vdots \\ \boldsymbol{\varphi}^{i_{q+1}}_{\bar{T}} \end{pmatrix}' \left\{ \mathbf{I}_{q+1} \otimes \boldsymbol{\Pi}^{\psi\prime}_{\bar{T}} \right\}' \right],
\end{aligned}
$$

and $\boldsymbol{\iota}_{K+1}$ and $\boldsymbol{\iota}_{q+1}$ the vectors of ones with dimensions $K+1$ and $q+1$, respectively. □

References

Altissimo, F., B. Mojon, and P. Zaffaroni (2009). Can aggregation explain the persistence of inflation? *Journal of Monetary Economics 56*, 231–241.

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

Amstad, M., S. M. Potter, and R. W. Rich (2017). The New York Fed staff underlying inflation gauge (UIG). *Economic Policy Review 23*, 1–32.

Anderson, B. D. and M. Deistler (2008a). Generalized linear dynamic factor models-A structure theory. In *Proceedings of the 47th IEEE Conference on Decision and Control*, pp. 1980–1985.

Anderson, B. D. and M. Deistler (2008b). Properties of zero-free transfer function matrices. *SICE Journal of Control, Measurement and System Integration 1*, 284–292.

Bai, J. (2003). Inferential theory for factor models of large dimensions. *Econometrica 71*, 135–171.

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

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

Barigozzi, M., M. Hallin, S. Soccorsi, and R. von Sachs (2021). Time-varying general dynamic factor models and the measurement of financial connectedness. *Journal of Econometrics 222*, 324–343.

Barigozzi, M. and M. Luciani (2020). Measuring the output gap using large datasets. SSRN, id 3217816.

Blinder, A. S. (1997). Measuring short-run inflation for central bankers - commentary. *Review*, 157–160. Federal Reserve Bank of St. Louis.

Brillinger, D. R. (2001). *Time Series: Data Analysis and Theory.* Society for Industrial and Applied Mathematics.

Bryan, M. F. and S. G. Cecchetti (1993). The consumer price index as a measure of inflation. *Economic Review, Federal Reserve Bank Quarter 4*, 15–24.

Cristadoro, R., M. Forni, L. Reichlin, and G. Veronese (2005). A core inflation indicator for the Euro area. *Journal of Money, Credit and Banking 37*, 539–560.

Dolmas, J. (2005). Trimmed mean PCE inflation. Working Paper 506, Federal Reserve Bank of Dallas.

Fan, J., F. Han, H. Liu, and B. Vickers (2016). Robust inference of risks of large portfolios. *Journal of Econometrics 194*, 298–308.

Fan, J., Y. Ke, and Y. Liao (2021). Augmented factor models with applications to validating market risk factors and forecasting bond risk premia. *Journal of Econometrics 222*, 269–294.

Fan, J., Y. Liao, and M. Mincheva (2013). Large covariance estimation by thresholding principal orthogonal complements. *Journal of the Royal Statistical Society Series B, Statistical Methodology 75*, 603–680.

Fan, J., Y. Liao, and X. Shi (2015). Risks of large portfolios. *Journal of Econometrics 186*, 367–387.

Fan, J., L. Xue, and J. Yao (2017). Sufficient forecasting using factor models. *Journal of Econometrics 201*, 292–306.

Forni, M., A. Giovannelli, M. Lippi, and S. Soccorsi (2018). Dynamic factor model with infinite-dimensional factor space: forecasting. *Journal of Applied Econometrics 33*, 625–642.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2000). The generalized dynamic factor model: identification and estimation. *Review of Economics and Statistics 82*, 540–554.

Forni, M., M. Hallin, M. Lippi, and P. Zaffaroni (2015). Dynamic factor models with infinite-dimensional factor spaces: one-sided representations. *Journal of Econometrics 185*, 359–371.

Forni, M., M. Hallin, M. Lippi, and P. Zaffaroni (2017). Dynamic factor models with infinite-dimensional factor spaces: Asymptotic analysis. *Journal of Econometrics 199*, 74–92.

Forni, M. and M. Lippi (2001). The generalized dynamic factor model: representation theory. *Econometric Theory 17*, 1113–1141.

Hallin, M. and M. Lippi (2013). Factor models in high-dimensional time series. A time-domain approach. *Stochastic Processes and their Applications 123*, 2678–2695.

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

Hasenzagl, T., F. Pellegrino, L. Reichlin, and G. Ricco (2020). A model of the Fed's view on inflation. *Review of Economics and Statistics*, 1–45.

Hörmann, S. and G. Nisol (2020). Prediction of singular VARs and an application to generalized dynamic factor models. *Journal of Time Series Analysis 42*, 295–313.

Lehmann, E. L. and J. P. Romano (2006). *Testing Statistical Hypotheses.* Springer Science & Business Media.

Luciani, M. (2020). Common and idiosyncratic inflation. FEDS 2020-024, Board of Governors of the Federal Reserve System.

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

Priestley, M. B. (1982). *The Spectral Analysis of Time Series.* Academic Press.

Reis, R. and M. W. Watson (2010). Relative goods' prices, pure inflation, and the Phillips correlation. *American Economic Journal Macroeconomics 2*, 128–157.

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

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

Wu, W. B. (2005). Nonlinear system theory: Another look at dependence. *Proceedings of the National Academy of Sciences of the United States of America 102*, 14150–14154.

Wu, W. B. and P. Zaffaroni (2018). Asymptotic theory for spectral density estimates of general multivariate time series. *Econometric Theory 34*, 1–22.

Yu, Y., T. Wang, and R. J. Samworth (2015). A useful variant of the Davis–Kahan theorem for statisticians. *Biometrika 102*, 315–323.

Zhang, D. and W. B. Wu (2021). Convergence of covariance and spectral density estimates for high-dimensional locally stationary processes. *The Annals of Statistics 49*, 233–254.