



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

ARCHIVIO ISTITUZIONALE
DELLA RICERCA

Alma Mater Studiorum Università di Bologna Archivio istituzionale della ricerca

Inferential theory for generalized dynamic factor models

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

Published Version:

Barigozzi, M., Hallin, M., Luciani, M., Zaffaroni, P. (2023). Inferential theory for generalized dynamic factor models. JOURNAL OF ECONOMETRICS, on line first, 1-41 [10.1016/j.jeconom.2023.02.003].

Availability:

This version is available at: <https://hdl.handle.net/11585/920911> since: 2023-03-21

Published:

DOI: <http://doi.org/10.1016/j.jeconom.2023.02.003>

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>).
When citing, please refer to the published version.

(Article begins on next page)

This is the final peer-reviewed accepted manuscript of:

Barigozzi, M., Hallin, M., Luciani, M., & Zaffaroni, P. (2023). Inferential theory for generalized dynamic factor models. Journal of Econometrics.

The final published version is available online at:

<https://doi.org/10.1016/j.jeconom.2023.02.003>

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>)

When citing, please refer to the published version.

INFERENCEAL THEORY FOR GENERALIZED DYNAMIC FACTOR MODELS[•]

Matteo Barigozzi[†]

Marc Hallin[‡]

Matteo Luciani^{*}

Paolo Zaffaroni^{**}

February 18, 2023

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. By exploiting the duality between common shocks and dynamic loadings, we derive the asymptotic distribution and associated standard errors for a class of estimators for common shocks, dynamic loadings, common components, and impulse response functions. We present an empirical application aimed at constructing a “core” inflation indicator for the U.S. economy, which demonstrates the superiority of the GDFM-based indicator over the most common approaches, particularly 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 for the estimators recently proposed in Forni et al. (2015, 2017) for the so-called *General or Generalized Dynamic Factor Models* (GDFM) introduced by Forni et al. (2000). The GDFM is extremely popular in applied economic and finance applications (see, e.g., Cristadoro et al., 2005; Giannone and Matheson, 2007; Altissimo et al., 2010; Hallin et al., 2011; Luciani, 2014; Amstad et al., 2017; Barigozzi and Hallin, 2017; Forni et al., 2018; Barigozzi and Hallin, 2020; Peña et al., 2021; Hallin et al., 2022), but their theoretical analysis remains incomplete: essentially, only consistency results have been established so far, sometimes with rates (see Forni et al., 2017), with the limiting distributional properties remaining unavailable, which precludes the possibility of inferential procedures. This paper aims at filling this gap by providing a fully developed and operational asymptotic theory for the GDFM.

Our methodology permits to exploit the flexibility of GDFMs in terms of dynamics with the versatility offered by the *Dynamic Factor Models* (DFMs) of Stock and Watson (2002a,b) and Bai (2003), in

[†] Department of Economics, Università di Bologna, Italy matteo.barigozzi@unibo.it

[‡] ECARES, Université libre de Bruxelles, Belgium mhallin@ulb.ac.be

^{*} Amazon.com luciani.matteo@gmail.com

^{**} Imperial College Business School, London, UK and University of Rome La Sapienza, Italy p.zaffaroni@imperial.ac.uk

[•]Disclaimer: Matteo Luciani’s contribution was part of a continued collaboration based on work done prior to joining Amazon. This publication and its contents are not related to Amazon and do not reflect the position of the company and its subsidiaries.

terms of estimating the common shocks and their impulse response functions (IRFs). The possibility to draw inference on these quantities represents the crucial ingredient at the core of almost every empirical analysis in dynamic macroeconomics, and was hitherto available for DFMs but not for GDFMs. The results developed in this paper are meeting that need.

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

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

where $\mathbf{u}_t = (u_{1t} \dots u_{qt})'$ is an unobservable q -dimensional white-noise vector of mutually orthogonal *common shocks* or *dynamic factors* driving $\{\chi_{it}|t \in \mathbb{Z}, i \in \mathbb{N}\}$ and $b_{if}(L), i \in \mathbb{N}, f = 1, \dots, q$, are square-summable filters (L , as usual, stands for the lag operator). Detailed assumptions on this decomposition are formalized in the next section. The unobservable χ_{it} and ξ_{it} 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} .

In many applications, the space spanned by the common components χ_{it} in (1), is likely to be infinite-dimensional: e.g., when the $b_{ij}(L)$ yield plausible dynamic structures such as AR(1) filters, they cannot be recovered from a finite number of standard principal components. To solve this problem, 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 methods proposed by Hallin and Liška (2007), Onatski (2009), or Avarucci et al. (2022).¹ However, being based on dynamic principal components, the Forni et al. (2000) 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, precluding (due to their two-sidedness) evaluation of their impulse response functions (IRFs) and out-of-sample prediction exercises.

This contrasts with the versatility of DFMs, the validity of which, however, requires the crucial assumption that, for any given t , the space spanned by the common components $\{\chi_{it}|i \in \mathbb{N}\}$ is *finite-dimensional*—of dimension r , say, with r independent of t —whereas for the GDFM in (1) one has, in general, $r = \infty$. Under this finite-dimensional assumption, (1) can be rewritten as

$$x_{it} = \lambda_{i1}F_{1t} + \dots + \lambda_{ir}F_{rt} + \xi_{it}, \quad i \in \mathbb{N}, t \in \mathbb{Z}, \quad (2)$$

$$\mathbf{D}(L)\mathbf{F}_t = \mathbf{H}\mathbf{u}_t, \quad (3)$$

where $\mathbf{F}_t = (F_{1t} \dots F_{rt})'$, $\mathbf{D}(L) := (\mathbf{I}_r - \mathbf{D}_1L - \mathbf{D}_2L^2 - \dots - \mathbf{D}_pL^p)$, for some finite natural p , \mathbf{I}_r is the identity matrix of size $r \times r$, and the matrices \mathbf{D}_j ($j = 1, \dots, p$) and \mathbf{H} , of size $r \times r$ and $r \times q$, respectively, define the singular Vector Autoregression (VAR) in (3) for \mathbf{F}_t . Because the vector \mathbf{F}_t of r so-called *static factors* is loaded contemporaneously via the loadings λ_{ij} , we call (2)-(3) a *static representation* of the GDFM, which, as explained above, is possible only if r is finite. Sometimes in the literature (2) is also called a *static factor model* for x_{it} . As already mentioned, the very appealing

¹ In fact, assuming $q = 1$ and $b_{i1}(L) = (1 - \alpha_i L)^{-1}$, one gets $\chi_{it} = u_t + \alpha_i u_{t-1} + \alpha_i^2 u_{t-2} + \dots$, that is χ_{it} depends on the infinite number of present and lagged values of u_t . Furthermore, assuming that the α_i have an absolutely continuous distribution rules out that χ_{it} can itself be interpreted as a static factor.

consequence of the finite-dimensional DFM formulation (2)-(3) is that it readily permits to derive the IRFs of the common shocks \mathbf{u}_t , which can be interpreted as *structural* shocks by means of identification restrictions on the matrix \mathbf{H} (Forni et al., 2009; Stock and Watson, 2016), and to perform out-of-sample forecasting in real time (Stock and Watson, 2002a; Giannone et al., 2008).

In a DFM, consistent estimates of the factors \mathbf{F}_t and the loadings λ_{ij} in (2) can be obtained using the first r *standard* principal components analysis (hereafter PCA) (see Stock and Watson, 2002a,b and Bai, 2003; and also Fan et al., 2013, 2015, 2016, 2017, 2021 where, in a finance context, several refinements of the PCA approach are proposed). Bai (2003) formalized the inferential theory for this standard PCA estimator, where r can be obtained, for instance, from the information criteria proposed by Bai and Ng (2002) or from the test proposed by Onatski (2010). Estimation of (2) and inference via Quasi Maximum Likelihood, when possibly considering also (3), has been studied in Doz et al. (2012), Bai and Li (2016), and Barigozzi and Luciani (2019).

Forni et al. (2015, 2017) bring together the virtues of the (infinite-dimensional) GDFM in (1) and the convenience of the (finite-dimensional) DFM in (2)-(3). In particular, under the mild assumption of existence and *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 —Forni et al. (2015) prove that the vector $\boldsymbol{\chi}_{nt} := (\chi_{1t}, \chi_{2t}, \dots, \chi_{nt})'$ of common components in (1) admits a *unique singular VAR representation* of the form $\mathbf{A}_n(L)\boldsymbol{\chi}_{nt} = \mathbf{R}_n\mathbf{u}_t$, where $\mathbf{A}_n(L)$ is a $n \times n$ block-diagonal matrix polynomial in L of *finite* degree, and \mathbf{R}_n a $n \times q$ matrix of rank q , implying

$$\mathbf{A}_n(L)\mathbf{x}_{nt} = \mathbf{A}_n(L)(\boldsymbol{\chi}_{nt} + \boldsymbol{\xi}_{nt}) = \mathbf{R}_n\mathbf{u}_t + \boldsymbol{\phi}_{nt}, \quad t \in \mathbb{Z}, \quad (4)$$

with $\boldsymbol{\phi}_{nt} := \mathbf{A}_n(L)\boldsymbol{\xi}_{nt}$ where $\boldsymbol{\xi}_{nt} := (\xi_{1t} \cdots \xi_{nt})'$.² Representation (4) is key because it shows how to re-write the *infinite-dimensional* GDFM (1) as a *static factor model* of the form (2) for the filtered variables $\mathbf{z}_{nt} := \mathbf{A}_n(L)\mathbf{x}_{nt}$ by setting $\mathbf{D}_n(L) = \mathbf{I}_n$, $\mathbf{F}_t = \mathbf{u}_t$, $(\lambda_{i1} \cdots \lambda_{ir}) = \mathbf{R}'_i$, the i th row of \mathbf{R}_n , and thus having $r = q$. Moreover, it can be shown that, under our assumptions, $\boldsymbol{\phi}_{nt}$ is idiosyncratic in the sense of Forni et al. (2000).

Therefore, in (4), the static and dynamic factors coincide, suggesting that standard PCA can be used for estimation of the factors \mathbf{u}_t and loadings \mathbf{R}_i . This would be trivial and inference possible by means of the established asymptotic results for standard PCA, if the \mathbf{z}_{nt} were observed. However, this would require knowing the VAR filters $\mathbf{A}_n(L)$, which in practice are unspecified. Forni et al. (2017) show how to consistently pre-estimate the $\mathbf{A}_n(L)$ and, subsequently, the factors and the loadings.

Clearly, representation (4) of the GDFM readily yields the IRFs $\mathbf{A}_n^{-1}(L)\mathbf{R}_n$ of the common shocks and allows for performing out-of-sample forecasting of the x_{it} and their components χ_{it} and ξ_{it} . Suitable identification restrictions on \mathbf{R}_n permit to interpret the elements of \mathbf{u}_t as structural economic shocks. In macroeconomic analysis, these quantities are typically coupled with bands that reflect the sampling variability of the estimates, i.e., confidence and forecasting intervals. However, the existing inferential theory for DFM (Bai, 2003) cannot be used due to the sampling variability affecting the estimation of the filters $\mathbf{A}_n(L)$, and the consistency result of Forni et al. (2017) falls short. This is the challenge resolved in this paper, exploiting new insights regarding the combination of the GDFM parameters with

²In particular, $\mathbf{A}_n(L)$ has a block-diagonal structure consisting of $(q+1) \times (q+1)$ -dimensional blocks, mitigating the curse of dimensionality because q is finite and typically small, and thus the high-dimensional VAR filter $\mathbf{A}_n(L)$ is obtained by piecing together a set of $m := \lfloor n/(q+1) \rfloor$ low-dimensional VARs.

different convergence rates, and suitably combining the different representations of GDFMs, i.e., the time-series and cross-sectional projections, to improve efficiency of their parameter estimates. Finally, we show how our asymptotic distribution theory is unaffected by the fact that q , in practice, must be replaced by a consistent estimator. Therefore, by using the theoretical results established in this paper, one can make full use of the GDFM in (1), through its finite-dimensional representation (4), for the most common and important empirical macroeconomic analyses.

The estimation of the GDFM decomposition (1) mainly consists of three steps which can be summarized as follows. First, by means of the method of Hallin and Liška (2007), estimate q and by means of dynamic PCA estimate 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, allows us to switch from the dynamic to the static representation (4) of the GDFM. Third, exploiting the finite-dimensional nature of (4), apply static PCA to the estimated \mathbf{z}_{nt} .

Building on the *uniform* consistency results for high-dimensional spectral density matrices by Wu and Zaffaroni (2018) and Zhang and Wu (2021), and on the consistency results by Forni et al. (2017), we provide here the complete limiting distribution theory for the estimation procedure just described, allowing one to perform asymptotically correct inference on the GDFM and its VAR form (4) as n and T diverge. Noticeably, despite the increased generality and greater complexity, with respect to the DFM, of the GDFM, our estimators achieve the same rates of convergence as the traditional PCA estimators, which are valid in the static model context (3) only. A detailed Monte Carlo study corroborates our theory. Our simulation results confirm the findings of Forni et al. (2017), who show that (i) when the data are generated by (4), the estimation of the common component by our estimators are by far better than via classical PCA; (ii) when the data are generated by (3), our estimators still outperform, although by a slight margin, the PCA ones. We then present an empirical application that demonstrates the flexibility and enhancement of our theory, focusing on building an indicator of core inflation for the U.S. economy.

In Sections 2, 3, and 4 we formalize the GDFM and the assumptions needed for our inferential theory. The estimation procedure is described in Section 5 and its limiting statistical properties are established in Section 6. Section 7 presents the Monte Carlo experiments, and Section 8 illustrates an empirical application on U.S. inflation indexes. Section 9 concludes. Technical proofs are relegated to the final Appendix.

Notation. Throughout, for a generic real symmetric $N \times N$ matrix \mathbf{A} with j th largest eigenvalue μ_j^A , we make use of the norms $\|\mathbf{A}\| = \mu_1^A$ and $\|\mathbf{A}\|_F = \sqrt{\text{tr}(\mathbf{A}\mathbf{A}')}$.

2 GDFM: General Representation

Throughout, we study a double-indexed zero-mean stochastic process $\mathbf{x} := \{x_{it} | i \in \mathbb{N}, t \in \mathbb{Z}\}$, of which we observe a finite realization $\{x_{it} | i = 1, \dots, n, t = 1, \dots, T\}$.³ Assumption (A) is listing the assumptions we are making on \mathbf{x} . These assumptions, are borrowed from Forni et al. (2015, 2017); they are formalizing and reinforcing the general presentation of the introduction.

³All stochastic variables in this paper belong to the Hilbert space $L_2(\Omega, \mathcal{F}, P)$ where (Ω, \mathcal{F}, P) is some common probability space.

Denote 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}_n is $\mathbf{\Gamma}_{n,k} := \text{Cov}(\mathbf{x}_{nt}, \mathbf{x}_{n,t-k})$, with $\mathbf{\Gamma}_n := \mathbf{\Gamma}_{n,0} := \text{Var}(\mathbf{x}_{nt})$ for simplicity.

Assumption (A). *The process \mathbf{x} is stationary with respect to time and \mathbf{x}_n , for all $n \in \mathbb{N}$, admits the spectral density*

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

where $\iota = \sqrt{-1}$. There exists a finite natural number $q > 0$ such that (1) holds with:

(a) a common component χ_{it} of the form

$$\chi_{it} = \sum_{j=1}^q b_{ij}(L)u_{jt} = \mathbf{b}'_i(L)\mathbf{u}_t, \quad i \in \mathbb{N}, t \in \mathbb{Z}, \quad (5)$$

where

- (a-i) $\{\mathbf{u}_t := (u_{1t} \cdots u_{qt})' | t \in \mathbb{Z}\}$ is a q -dimensional i.i.d. zero-mean stochastic process, with positive definite covariance $\mathbb{E}[\mathbf{u}_t \mathbf{u}'_t] =: \mathbf{\Gamma}^u$; moreover, for some $\varepsilon > 0$ and some finite constant $M_u > 0$ independent of j and t , $\mathbb{E}[|u_{jt}|^{4+\varepsilon}] \leq M_u$ for all $j = 1, \dots, q$;
- (a-ii) for all $i \in \mathbb{N}$, $j = 1, \dots, q$, and $z \in \mathbb{C}$, $b_{ij}(z)$ is rational, that is, $b_{ij}(z) = c_{ij}(z)/d_{ij}(z)$ where
 - $c_{ij}(z) = \sum_{k=0}^{s_1} c_{ij,k} z^k$ for some positive integer s_1 , with $|c_{ij,k}| \leq M_c$ for some finite $M_c > 0$ independent of i and j and
 - $d_{ij}(z) = \sum_{k=0}^{s_2} d_{ij,k} z^k$ for some positive integer s_2 is such that all the roots of $d_{ij}(z) = 0$ satisfy $|z| \geq M_d > 1$ for some $M_d > 0$ independent of i and j ;
- (a-iii) for all $j = 1, \dots, q$, there exist two real strictly positive continuous functions $\theta \mapsto \underline{\lambda}_j^x(\theta)$ and $\theta \mapsto \bar{\lambda}_j^x(\theta)$ such that, for all $\theta \in [-\pi, \pi]$,

$$\underline{\lambda}_j^x(\theta) \leq \liminf_{n \rightarrow \infty} \frac{\lambda_{nj}^x(\theta)}{n} \leq \limsup_{n \rightarrow \infty} \frac{\lambda_{nj}^x(\theta)}{n} \leq \bar{\lambda}_j^x(\theta)$$

with $\bar{\lambda}_j^x(\theta) < \underline{\lambda}_{j-1}^x(\theta)$ for all $j = 2, \dots, q$;

(b) an idiosyncratic component ξ_{it} satisfying

$$\xi_{it} = \sum_{j=1}^{\infty} \sum_{k=0}^{\infty} \beta_{ij,k} \eta_{j,t-k}, \quad i \in \mathbb{N}, t \in \mathbb{Z}, \quad (6)$$

where

- (b-i) $\{\boldsymbol{\eta}_t := (\eta_{1t} \eta_{2t} \cdots)' | t \in \mathbb{Z}\}$ is an infinite-dimensional i.i.d. zero-mean stochastic process such that $\mathbb{E}[\eta_{it} \eta_{jt}] = 0$ for all $i \neq j$, $\mathbb{E}[\eta_{it}^2] = 1$, and $\mathbb{E}[|\eta_{it}|^{4+\varepsilon}] \leq M_\eta$ for some $\varepsilon > 0$ and some finite constant $M_\eta > 0$ independent of i and t ;
- (b-ii) 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 ;
- (c) $\{\mathbf{u}_t\}$ and $\{\boldsymbol{\eta}_t\}$ are mutually independent processes.

Remark 1. Denoting by $\lambda_{nj}(\theta)$ the j th largest eigenvalue of the spectral density matrix $\mathbf{\Sigma}_n(\theta)$, it has been shown (see, e.g., Forni and Lippi, 2001; Hallin and Lippi, 2013) that \mathbf{x} admits a GDFM

representation (1) with q common shocks if $\Sigma_n(\theta)$ has $q < \infty$ diverging eigenvalues, that is,

$$\lim_{n \rightarrow \infty} \lambda_{nq}(\theta) = \infty, \quad \theta\text{-a.e in } [-\pi, \pi] \quad \text{and} \quad \sup_{\theta \in [-\pi, \pi]} \sup_{n \in \mathbb{N}} \lambda_{n, q+1}(\theta) < \infty. \quad (7)$$

Now, under Assumption (A), the n -dimensional subprocesses

$$\chi_n := \{\chi_{1t}, \dots, \chi_{nt} | t \in \mathbb{Z}\} \quad \text{and} \quad \xi_n := \{\xi_{1t}, \dots, \xi_{nt} | t \in \mathbb{Z}\}$$

admit spectral density matrices $\Sigma_n^x(\theta)$ and $\Sigma_n^\xi(\theta)$ with j -th largest eigenvalues $\lambda_{nj}^x(\theta)$ and $\lambda_{nj}^\xi(\theta)$ satisfying

$$\lim_{n \rightarrow \infty} \lambda_{nq}^x(\theta) = \infty, \quad \theta\text{-a.e in } [-\pi, \pi] \quad \text{and} \quad \sup_{\theta \in [-\pi, \pi]} \sup_{n \in \mathbb{N}} \lambda_{n1}^\xi(\theta) < \infty, \quad (8)$$

respectively; this is a consequence of Assumption (A-*a-iii*) for $\lambda_{nj}^x(\theta)$, of Assumption (A-*b-ii*) for $\lambda_{n1}^\xi(\theta)$ (see Remark 3 for details). It then straightforwardly follows (see Remark 5) that (8) implies (7).⁴ Hence, under Assumption (A), \mathbf{x} admits a GDFM representation and the number q of diverging eigenvalues of $\Sigma_n(\theta)$ corresponds to the number of dynamic factors and common shocks.

Remark 2. Assumption (A) requires q to be finite, irrespective of n . That assumption is standard in all the factor model literature (Stock and Watson, 2002a,b; Bai and Ng, 2002; Bai, 2003; Doz et al., 2012; Forni et al., 2017; etc.). It sometimes has been argued that q itself should increase as $n \rightarrow \infty$. Remember, however, that the cross-sectional dimension, in practice, is fixed, and that letting $n \rightarrow \infty$ is a mathematical way to construct a tractable approximation to the actual finite- n problem. Adding factors or common shocks that are not present in the finite- n observation is unlikely to improve that approximation, while purposelessly making the inferential problem harder. An n -dependent q , moreover, is incompatible with the random cross-section approach adopted here.

Remark 3. Assumption (A-*a-ii*) implies that $\sigma_{ij}^x(\theta)$ is a rational spectral density matrix, bounded in θ uniformly in i and j . Assumption (A-*b-ii*) entails square-summability of the idiosyncratic filters for the idiosyncratic MA(∞) representation (6) in the i.i.d. innovations, both along the time and the cross-sectional dimensions. This, in turn, implies limited (lagged) cross-sectional dependence among idiosyncratic components. For example, Assumption (A-*b-ii*) is obviously satisfied in the *purely idiosyncratic* case $\xi_{it} = \eta_{it}$ and for finite *cross-section moving averages*, for example $\xi_{it} = \eta_{it} + \eta_{i+1t}$. Moreover, letting $\sigma_{ij}^\xi(\theta)$ denote the (i, j) th entry of the idiosyncratic spectral density matrix $\Sigma_n^\xi(\theta)$, it follows that

$$\sup_{\theta \in [-\pi, \pi]} \sup_{j \in \mathbb{N}} \sum_{i=1}^{\infty} |\sigma_{ij}^\xi(\theta)| \leq \sup_{\theta \in [-\pi, \pi]} \sup_{j \in \mathbb{N}} \frac{1}{(1-\rho)^2} \sum_{i=1}^{\infty} \sum_{s=0}^{\infty} B_{is} B_{js} \leq \frac{B^2}{(1-\rho)^2} = B^\xi, \text{ say.}$$

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

$$\sup_{\theta \in [-\pi, \pi]} \sup_{n \in \mathbb{N}} \lambda_{n1}^\xi(\theta) \leq B^\xi, \quad (9)$$

hence the ξ_{it} s are effectively idiosyncratic components since they satisfy (8).

⁴In fact, also the converse holds true, see Forni and Lippi (2001).

Remark 4. Assumption (A-*a-iii*) implies that each common shock u_{it} is pervasive in the sense that it affects almost all items of the cross-section as n increases. This implies that the common components χ_{it} are identified (see Chamberlain and Rothschild, 1983), that the number of dynamic factors q is unique, ruling out the possibility of a representation like (5) with a different number of dynamic factors (see Forni and Lippi, 2001). Linear divergence rates, moreover, are the only ones compatible with the fact that the cross-sectional ordering is completely arbitrary, hence should remain irrelevant—see the stochastic approach in Section 3 for further justification. In other words, we are only considering the case of strong factors, as opposed to the case of factors which have a weak effect on all series (see, e.g., Onatski, 2012) or only affect some sub-group of series (see, e.g., Hallin and Liška, 2011).

Remark 5. Having spelled out the second-order properties of the common component χ_{it} and ξ_{it} , the corresponding properties for the observables x_{it} follow. For instance, the cross-spectral densities $\sigma_{ij}^x(\theta) = \sigma_{ij}^{\chi}(\theta) + \sigma_{ij}^{\xi}(\theta)$ are bounded, in θ , uniformly in i and j . It also follows (see Forni et al., 2017, Proposition 2) that $\sigma_{ij}^x(\theta)$ possesses derivatives of any order and are of bounded variation uniformly in $i, j \in \mathbb{N}$.⁵ Furthermore, a simple application of Weyl's inequality entails linear divergence and separability of the largest q eigenvalues $\lambda_{nj}(\theta)$ of $\Sigma_n(\theta)$, i.e., for all $j = 1, \dots, q$, and all $\theta \in [-\pi, \pi]$,

$$\underline{\lambda}_j^x(\theta) \leq \liminf_{n \rightarrow \infty} \frac{\lambda_{nj}(\theta)}{n} \leq \limsup_{n \rightarrow \infty} \frac{\lambda_{nj}(\theta)}{n} \leq \bar{\lambda}_j^x(\theta), \text{ and } \sup_{\theta \in [-\pi, \pi]} \sup_{n \in \mathbb{N}} \lambda_{n, q+1}(\theta) \leq B^x \quad (10)$$

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

Remark 6. Finite $(4 + \varepsilon)$ th moments and i.i.d.-ness of the common and idiosyncratic shocks \mathbf{u}_t and $\boldsymbol{\eta}_t$ are needed in order to control the degree of *physical dependence* (Wu, 2005) of the common and idiosyncratic components, hence of each x_{it} . Note that Bai (2003, Assumption C) requires boundedness of the eighth moment of ξ_{it} s, which is a much stronger condition.

Remark 7. Let us observe that whereas the χ_{it} are identified, the \mathbf{u}_t and the polynomials $\mathbf{b}_i(L)$ are identified up to an invertible linear transformation, i.e., for any invertible $q \times q$ matrix \mathbf{H} the common component χ_{it} has the alternative representation $\chi_{it} = \mathbf{b}'_i(L)\mathbf{H}\mathbf{H}^{-1}\mathbf{u}_t = \mathbf{b}'_i^*(L)\mathbf{u}_t^*$. Further discussion on identification is elaborated in the next section, where we describe the VAR representation of the GDFM.

Example. AR(1) common and idiosyncratic components. We illustrate how the eigenvalue conditions derived from Assumption (A) are verified in a simple GDFM example. Let

$$x_{it} = \frac{u_t}{1 - d_i L} + \frac{\eta_{it}}{1 - \rho_i L}, \quad (11)$$

where $\{d_i\}$ and $\{\rho_i\}$ are i.i.d., with $-1 < -1/M_d \leq d_i \leq 1/M_d < 1$ and $-1 < -\rho \leq \rho_i \leq \rho < 1$, and $\{u_t\}$ and $\{\eta_{it}\}$ satisfy Assumptions (A-*a-i*) and (A-*b-i*), respectively.

Despite its simplicity, (11) does not admit the popular static representation (2)-(3) studied by Stock and Watson (2002a,b) and Bai (2003) as soon as the AR coefficients d_i are generic, i.e., are drawn from an absolutely continuous distribution (see footnote 1), thus severely limiting the applicability of the DFM approach. Let us show that it nevertheless enters the realm of Assumption (A)

⁵Namely, there exists $A^x > 0$ such that $\sum_{h=1}^{\nu} |\sigma_{ij}^x(\theta_h) - \sigma_{ij}^x(\theta_{h-1})| \leq A^x$, for all $i, j, \nu \in \mathbb{N}$ and all partitions of the form $-\pi = \theta_0 < \theta_1 < \dots < \theta_{\nu-1} < \theta_{\nu} = \pi$ of the interval $[-\pi, \pi]$.

Setting $c_{i1,0} = 1, c_{i1,k} = 0$ for $k > 0$, $d_{i1,0} = 1$, and $d_{i1,1} = -d_i$, Assumption (A-a-ii) follows. Assumption (A-b-ii) holds with $\beta_{ij,k} = 0$ whenever $i \neq j$ for every k , and $\beta_{ii,k} = \rho_i^k$. Finally, letting $B_{ij} = 0$ for $i \neq j$ and $B_{ii} = B = 1$ otherwise, Assumption (A-c) is satisfied provided that $\text{Cov}(u_s, \eta_{it}) = 0$ for every s, t , and i . Writing Γ^u for $\mathbb{V}\text{ar}(u_t)$, elementary algebra, yields, for any given $n \in \mathbb{N}$, the spectral density matrix

$$\begin{aligned} \Sigma_n^X(\theta) &= \frac{\Gamma^u}{2\pi} \begin{pmatrix} \frac{1}{1-d_1 e^{-i\theta}} \\ \vdots \\ \frac{1}{1-d_N e^{-i\theta}} \end{pmatrix} \begin{pmatrix} \frac{1}{1-d_1 e^{i\theta}} \cdots \frac{1}{1-d_N e^{i\theta}} \end{pmatrix} \\ &= \frac{\Gamma^u}{2\pi} \begin{pmatrix} \frac{1}{|1-d_1 e^{-i\theta}|^2} & \cdots & \frac{1}{(1-d_1 e^{-i\theta})(1-d_N e^{i\theta})} \\ \vdots & \ddots & \vdots \\ \frac{1}{(1-d_N e^{-i\theta})(1-d_1 e^{i\theta})} & \cdots & \frac{1}{|1-d_N e^{-i\theta}|^2} \end{pmatrix}, \end{aligned}$$

which has rank one for all θ and unique non-zero eigenvalue $\lambda_{n1}^X(\theta) = \frac{\Gamma^u}{2\pi} \sum_{i=1}^n \frac{1}{|1-d_i e^{-i\theta}|^2}$ satisfying

$$\frac{1}{2} \leq \frac{1}{1+1/M_d^2} \leq \frac{1}{n} \sum_{i=1}^n \frac{1}{1+d_i^2} \leq \frac{1}{n} \sum_{i=1}^n \frac{1}{|1-d_i e^{-i\theta}|^2} \leq \frac{1}{n} \sum_{i=1}^n \frac{1}{1-d_i^2} \leq \frac{1}{1-1/M_d^2} \leq 1,$$

hence

$$\lambda_1^X(\theta) := 1/2 \leq \lambda_{n1}^X(\theta)/n \leq 1 =: \bar{\lambda}_1^X(\theta).$$

Assumption (A-a-iii) thus holds. Finally, the spectral density matrix of $\boldsymbol{\xi}_n$, for any given $n \in \mathbb{N}$, is

$$\Sigma_n^\xi(\theta) = \frac{1}{2\pi} \begin{pmatrix} \frac{1}{|1-\rho_1 e^{-i\theta}|^2} & 0 & \cdots & 0 \\ 0 & \frac{1}{|1-\rho_2 e^{-i\theta}|^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{|1-\rho_N e^{-i\theta}|^2} \end{pmatrix},$$

with largest eigenvalue $\lambda_{n1}^\xi(\theta) \leq 1/(1-\rho^2) \leq 1$: all the eigenvalues of the spectral density matrix $\Sigma_n(\theta)$ of \mathbf{x}_n thus satisfy (10) in Remark 5.

3 GDFM: Singular Vector Autoregressive Representation

Let Assumption (A) hold. For any $s \in \mathbb{N}$ and $t \in \mathbb{Z}$, consider the $(q+1)$ -dimensional subvector of common components $\boldsymbol{\chi}_t^{(s)} := (\chi_{(s-1)(q+1)+1,t} \cdots \chi_{s(q+1),t})'$. 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 (A-a-ii): turning it into an assumption, thus, only places an extremely mild restriction on the actual data-generating process.

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

$$\mathbf{A}^{(s)}(L)\boldsymbol{\chi}_t^{(s)} = \mathcal{R}^{(s)}\mathbf{u}_t, \quad t \in \mathbb{Z} \tag{12}$$

where

- (a) $p_s \leq S := qs_1 + q^2s_2 < \infty$ and all the roots of the determinantal equation $\det(\mathbf{A}^{(s)}(z)) = 0$, $z \in \mathbb{C}$, are such that $|z| > 1$;
- (b) $\mathbf{R}^{(s)}$ has maximal rank q ;
- (c) denoting by \mathbf{C}_s^X the $S(q+1) \times S(q+1)$ covariance matrix of $(\boldsymbol{\chi}_t^{(s)'} \cdots \boldsymbol{\chi}_{t-S}^{(s)'})'$, there exists a finite real d such that $\det(\mathbf{C}_s^X) \geq d > 0$ for all $s \in \mathbb{N}$.

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 $\underline{\mathbf{R}} := (\mathbf{R}^{(1)'} \mathbf{R}^{(2)'} \cdots)'$ with $(q+1)$ columns and infinitely many rows. Considering, without loss of generality, n such that $n = m(q+1)$ for some integer m , let

$$\mathbf{A}_n(L) := \begin{pmatrix} \mathbf{A}^{(1)}(L) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}^{(2)}(L) & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{A}^{(m)}(L) \end{pmatrix} \quad \text{and} \quad \mathbf{R}_n := \begin{pmatrix} \mathbf{R}^{(1)} \\ \mathbf{R}^{(2)} \\ \vdots \\ \mathbf{R}^{(m)} \end{pmatrix} \quad (13)$$

denote the upper $n \times n$ and upper $n \times q$ sub-matrices of $\underline{\mathbf{A}}(L)$ and $\underline{\mathbf{R}}$, respectively. Then, from Assumption (B), the common component $\boldsymbol{\chi}_n$ admits the finite-order singular VAR representation

$$\mathbf{A}_n(L)\boldsymbol{\chi}_{nt} = \mathbf{R}_n\mathbf{u}_t, \quad n \in \mathbb{N}, t \in \mathbb{Z}, \quad (14)$$

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

$$\boldsymbol{\chi}_{nt} = \mathbf{B}_n(L)\mathbf{u}_t = \mathbf{C}_n(L)\mathbf{R}_n\mathbf{u}_t, \quad n \in \mathbb{N}, t \in \mathbb{Z}. \quad (15)$$

Here, $\mathbf{B}_n(L)$ are the non-identified IRFs of the common shocks \mathbf{u}_t .

Now, since $\mathbf{A}_n(L)$ is block-diagonal, $\mathbf{C}_n(L)$ also is block-diagonal. Therefore, denoting by $\mathcal{I}_s := \{\ell \mid \ell = (s-1)(\widehat{q}+1) + 1, \dots, s(\widehat{q}+1)\}$ the set of cross-sectional indices of the series belonging to block s , $s = 1, \dots, m$, each common component χ_{it} of $\boldsymbol{\chi}_{nt}$ in (15) satisfies

$$\chi_{it} = \sum_{k=0}^{\infty} \sum_{j_s=1}^{q+1} c_{i,j_s,k} \mathbf{R}'_{j_s} \mathbf{u}_{t-k}, \quad i \in \mathcal{I}_s, s = 1, \dots, m, t \in \mathbb{Z}, \quad (16)$$

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

The following example shows how to build a singular VAR representation of the common component.

Example. MA(1) common components. Consider a GDFM with $q = 1$ and MA(1) common components. Since $q = 1$, $q + 1 = 2$ and

$$\chi_{it} = c_{i,0}u_t + c_{i,1}u_{t-1}, \quad (17)$$

$$\chi_{jt} = c_{j,0}u_t + c_{j,1}u_{t-1}. \quad (18)$$

Excluding the non-generic⁶ subset of \mathbb{R}^4 in which $c_{i,0}c_{j,1} - c_{i,1}c_{j,0} = 0$, one obtains the solu-

⁶Several difstinct notions of "generic" exist in mathematics. In measure theory, a generic property is one that holds

tion $u_t = (c_{j,1}\chi_{it} - c_{i,1}\chi_{jt})/(c_{i,0}c_{j,1} - c_{i,1}c_{j,0})$. Taking this into account in (17)-(18), yields the desired VAR(1) representation

$$\chi_{it} = \frac{c_{i,1}c_{j,1}}{(c_{i,0}c_{j,1} - c_{i,1}c_{j,0})}\chi_{it-1} - \frac{(c_{i,1})^2}{(c_{i,0}c_{j,1} - c_{i,1}c_{j,0})}\chi_{jt-1} + c_{i,0}u_t, \quad (19)$$

$$\chi_{jt} = \frac{(c_{j,1})^2}{(c_{i,0}c_{j,1} - c_{i,1}c_{j,0})}\chi_{it-1} - \frac{c_{i,1}c_{j,1}}{(c_{i,0}c_{j,1} - c_{i,1}c_{j,0})}\chi_{jt-1} + c_{j,0}u_t. \quad (20)$$

Now, we can assume without loss of generality that n is an even number and that $i = 2l - 1$ and $j = 2l$ for some $l = 1, \dots, m$ and $m = n/2$. Then, the l th diagonal 2×2 block of $\mathbf{A}_n(L)$ and the corresponding rows of \mathcal{R}_n in (12) read as

$$\mathbf{A}^{(l)}(L) = \mathbf{I}_2 - \begin{pmatrix} \frac{c_{i,1}c_{j,1}}{(c_{i,0}c_{j,1} - c_{i,1}c_{j,0})} & -\frac{(c_{i,1})^2}{(c_{i,0}c_{j,1} - c_{i,1}c_{j,0})} \\ \frac{(c_{j,1})^2}{(c_{i,0}c_{j,1} - c_{i,1}c_{j,0})} & -\frac{c_{i,1}c_{j,1}}{(c_{i,0}c_{j,1} - c_{i,1}c_{j,0})} \end{pmatrix} L \quad \text{and} \quad \mathcal{R}^{(l)} = \begin{pmatrix} c_{i,0} \\ c_{j,0} \end{pmatrix}. \quad (21)$$

It can be shown (see Forni et al., 2015) that no other autoregressive representation of order one exists, but many other autoregressive representations of order two and higher can be obtained. This, however, is not the case for *square* systems, i.e., when $n = q$, where only an infinite VAR representation of (17)-(18) would be possible.⁷

In view of $\mathbf{x}_{nt} = \boldsymbol{\chi}_{nt} + \boldsymbol{\xi}_{nt}$, we obtain

$$\mathbf{z}_{nt} := \mathbf{A}_n(L)\mathbf{x}_{nt} = \mathcal{R}_n\mathbf{u}_t + \mathbf{A}_n(L)\boldsymbol{\xi}_{nt} =: \boldsymbol{\psi}_{nt} + \boldsymbol{\phi}_{nt}, \quad n \in \mathbb{N}, t \in \mathbb{Z}. \quad (22)$$

As shown in Forni et al. (2017, Proposition 4), it immediately follows from Assumptions (A)-(B) and (9) that the eigenvalues of the spectral density matrix of $\boldsymbol{\phi}_n$ are uniformly bounded for $\theta \in [-\pi, \pi]$ for all $n \in \mathbb{N}$. Indeed, let $\boldsymbol{\Sigma}_n^\phi(\theta) := \mathbf{A}_n(e^{-i\theta})\boldsymbol{\Sigma}_n^\xi(\theta)\mathbf{A}_n'(e^{i\theta})$, and denote by $\boldsymbol{\Lambda}_n^\phi(\theta)$ the $n \times n$ diagonal matrix of the eigenvalues $\lambda_{nj}^\phi(\theta)$ of $\boldsymbol{\Sigma}_n^\phi(\theta)$. 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, $\boldsymbol{\Sigma}_n^\phi(\theta) = \mathbf{P}_n^\phi(\theta)\boldsymbol{\Lambda}_n^\phi(\theta)\mathbf{P}_n^{\phi\dagger}(\theta)$ where $\mathbf{P}_n^{\phi\dagger}$ stands for the transposed complex-conjugate of \mathbf{P}_n^ϕ . We have

$$\sup_{\theta \in [-\pi, \pi]} \sup_{n \in \mathbb{N}} \lambda_{n1}^\phi(\theta) \leq \sup_{\theta \in [-\pi, \pi]} \sup_{n \in \mathbb{N}} \lambda_{n1}^\xi(\theta)\lambda_{n1}^A(\theta) \leq B^\xi D^\phi =: B^\phi, \quad \text{say}, \quad (23)$$

where $\lambda_{n1}^A(\theta)$ is the largest eigenvalue of $\mathbf{A}_n(e^{-i\theta})\mathbf{A}_n'(e^{i\theta})$, which is finite because of Assumptions (B-a) and (B-c). Therefore, $\boldsymbol{\phi}_n$ is still idiosyncratic in the sense of Section 2, and we call (22) the singular VAR representation of the GDFM.

4 Dual static representation

Let us show that (22) is a static factor model for \mathbf{z}_{nt} in the sense of Bai (2003). Hereafter, we consider the nT -dimensional subprocess $\{x_{it} \mid 1 \leq i \leq n, 1 \leq t \leq T\}$ of \mathbf{x} , with elements that, for any given $n \in \mathbb{N}$

almost everywhere, that is, except over a set of measure zero; in topology and algebraic geometry, a generic property is one that holds on a dense open set, that is, except over a nowhere dense set or, more generally, a *meagre* set. Here, we adopt the measure-theoretic point of view.

⁷Interestingly, when $c_{i,0}c_{j,1} - c_{i,1}c_{j,0} \neq 0$, the noise u_t is fundamental for $(\chi_{it}, \chi_{jt})'$ even if (17)-(18) both are non-invertible (see Forni et al., 2015), that is, even if u_{it} (u_{jt}) is not fundamental for χ_{it} (χ_{jt}).

and $T \in \mathbb{N}$, can be rearranged as

$$\mathbf{X}_{nT} := \begin{pmatrix} x_{11} & \cdots & x_{i1} & \cdots & x_{n1} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ x_{1t} & \cdots & x_{it} & \cdots & x_{nt} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ x_{1T} & \cdots & x_{iT} & \cdots & x_{nT} \end{pmatrix} =: \begin{pmatrix} \mathbf{x}'_{n1} \\ \vdots \\ \mathbf{x}'_{nt} \\ \vdots \\ \mathbf{x}'_{nT} \end{pmatrix} =: (\mathbf{x}_T^1 \cdots \mathbf{x}_T^i \cdots \mathbf{x}_T^n), \quad (24)$$

with t th row $\mathbf{x}_{nt} = (x_{1t} \cdots x_{nt})'$ (an n -dimensional random vector) and i th column $\mathbf{x}_T^i = (x_{i1} \cdots x_{iT})'$ (a T -dimensional random vector), respectively. Similarly define the idiosyncratic $T \times n$ matrix

$$\Phi_{nT} = (\phi_{n1} \cdots \phi_{nt} \cdots \phi_{nT})' = (\varphi_T^1 \cdots \varphi_T^i \cdots \varphi_T^n),$$

with t th row $\phi'_{nt} = (\phi_{1t} \cdots \phi_{nt})$ (an n -dimensional random vector) and i th column $\varphi_T^i = (\phi_{i1} \cdots \phi_{iT})'$ (a T -dimensional random vector). Define the $T \times q$ matrix of common shocks $\mathbf{U}_T := (\mathbf{u}_1 \cdots \mathbf{u}_t \cdots \mathbf{u}_T)'$, with t th row \mathbf{u}'_t (a q -dimensional random vector). With this notation, the static representation (22) of the GDFM, takes the form (in matrix representation)

$$\mathbf{Z}_{nT} := (\mathbf{A}_n(L)\mathbf{X}'_{nT})' = \mathbf{U}_T \mathbf{R}'_n + \Phi_{nT} =: \Psi_{nT} + \Phi_{nT}, \quad (25)$$

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

Furthermore, the matrix representation (25) can be written under (transposed) row-vector form as a cross-sectional projection which, denoting by \mathbf{e}_{Tt} the t th column of the $T \times T$ identity matrix \mathbf{I}_T , is given by

$$\mathbf{z}_{nt} = (\mathbf{A}_n(L)\mathbf{X}'_{nT}) \mathbf{e}_{Tt} = \mathbf{R}_n \mathbf{u}_t + \phi_{nt}, \quad n \in \mathbb{N}, t \in \mathbb{Z} \quad (26)$$

(that is, the static factor model (22)). Or, equivalently, by denoting as \mathbf{R}'_i the i th q -dimensional row of \mathbf{R}_n and by \mathbf{e}_{ni} the i th column of the $n \times n$ identity matrix \mathbf{I}_n , the matrix representation (25) can be written under column-vector form as a temporal projection given by

$$\mathbf{z}_T^i = (\mathbf{A}_n(L)\mathbf{X}'_{nT})' \mathbf{e}_{ni} = \mathbf{U}_T \mathbf{R}_i + \varphi_T^i, \quad i \in \mathbb{N}, T \in \mathbb{N}. \quad (27)$$

The two forms (26) and (27) constitute dual static factor model representations, with the time- and cross-sectional-dimensions (the indices i and t) exchanging roles—that is, \mathbf{X}'_{nT} replacing \mathbf{X}_{nT} . The duality between (26) and (27), as we shall see, plays a fundamental role when studying the limiting properties of the estimators. Exploiting it, however, requires some structure on (27) and the \mathbf{R}_i s that parallels the assumed structure provided by Assumption (A-*a*) for (26) and the \mathbf{u}_t s, along with some bounds on the second-order moments of the idiosyncratic components ϕ_{it} .

Assumption (C).

- (a) For all n and T , the distribution of \mathbf{X}_{nT} , hence also the distributions of \mathbf{Z}_{nT} , Ψ_{nT} , and Φ_{nT} are invariant under column permutations, i.e., they are cross-sectionally exchangeable;
- (b) $\{\mathbf{R}_i := (R_{i1} \cdots R_{iq})' | i \in \mathbb{N}\}$ is a q -dimensional i.i.d. zero-mean stochastic process, with positive definite covariance $\mathbb{E}[\mathbf{R}_i \mathbf{R}'_i] =: \Sigma^R$; moreover, for all $j = 1, \dots, q$ and all $i \in \mathbb{N}$, $\mathbb{E}[|R_{ij}|^{4+\varepsilon}] \leq M_R$ for some $\varepsilon > 0$ and some finite constant $M_R > 0$ independent of i and j ;

- (c) $\{\mathbf{R}_i | i \in \mathbb{N}\}$ and $\{(\mathbf{u}_t, \boldsymbol{\eta}_t) | t \in \mathbb{Z}\}$ are mutually independent processes;
- (d-i) for all $i, j \in \mathbb{N}$, $\lim_{T \rightarrow \infty} \mathbb{E}[(\frac{1}{\sqrt{T}} \sum_{t=1}^T \{\phi_{it}\phi_{jt} - \mathbb{E}[\phi_{it}\phi_{jt}]\})^2] \leq M_\Gamma$ for some finite constant $M_\Gamma > 0$ independent of i and j ;
- (d-ii) for all $t, s \in \mathbb{Z}$, $\lim_{n \rightarrow \infty} \mathbb{E}[(\frac{1}{\sqrt{n}} \sum_{i=1}^n \{\phi_{it}\phi_{is} - \mathbb{E}[\phi_{it}\phi_{is}]\})^2] \leq M_G$ for some finite constant $M_G > 0$ independent of t and s .

Remark 8. Under Assumption (C), the duality between representations (26) and (27) is fully established: 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 \mathbf{R}_n in (26) and random vectors \mathbf{R}_i loaded by cross-sectional item i at time t via random loadings \mathbf{u}_T in (27). Both \mathbf{u}_t and \mathbf{R}_i are i.i.d. noise, the only difference being that $\{\mathbf{u}_t\}$ is simply i.i.d. while $\{\mathbf{R}_i\}$ also is *exchangeable* in the sense of Assumption (C-a), which takes into account the irrelevance of the cross-sectional ordering (while chronological ordering is of fundamental importance). Note also that the linear rate of divergence of exploding eigenvalues in Assumption (A-a-iii) is the only rate compatible with cross-sectional exchangeability. Appendix E clarifies the relationship between Assumption (C) and the notion of exchangeability.

Remark 9. Assumption (C-d-i) is the same as Assumption 8 in Forni et al. (2009) and it basically says that, for any given $n \in \mathbb{N}$, we can estimate consistently all entries of the covariance matrix of $\{\phi_{nt}\}$. This is quite natural and mild in view of stationarity and the existence of 4-th moments of the innovations in (6), see Remark 10 below. Moreover, if for any fixed $n \in \mathbb{N}$ we let $\mathbf{\Gamma}_n^\phi$ be the $n \times n$ symmetric matrix with entries $\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \mathbb{E}[\phi_{it}\phi_{jt}] =: \gamma_{ij}^\phi$ (due to stationarity), $i, j = 1, \dots, n$, then, Assumption (C-d-i) implies that, as $T \rightarrow \infty$,

$$\frac{1}{n} \left\| \frac{\mathbf{\Phi}'_{nT} \mathbf{\Phi}_{nT}}{T} - \mathbf{\Gamma}_n^\phi \right\| = O_P \left(\frac{1}{\sqrt{T}} \right). \quad (28)$$

Indeed,

$$\begin{aligned} \lim_{T \rightarrow \infty} T \mathbb{E} \left[\frac{1}{n^2} \left\| \frac{\mathbf{\Phi}'_{nT} \mathbf{\Phi}_{nT}}{T} - \mathbf{\Gamma}_n^\phi \right\|^2 \right] &\leq \lim_{T \rightarrow \infty} T \mathbb{E} \left[\frac{1}{n^2} \left\| \frac{\mathbf{\Phi}'_{nT} \mathbf{\Phi}_{nT}}{T} - \mathbf{\Gamma}_n^\phi \right\|_F^2 \right] \\ &= \frac{1}{n^2} \sum_{i,j=1}^n \lim_{T \rightarrow \infty} T \mathbb{E} \left[\left(\frac{1}{T} \sum_{t=1}^T \phi_{it}\phi_{jt} - \gamma_{ij}^\phi \right)^2 \right] \\ &\leq \max_{i,j=1,\dots,n} \lim_{T \rightarrow \infty} \mathbb{E} \left[\left(\frac{1}{\sqrt{T}} \sum_{t=1}^T \{ \phi_{it}\phi_{jt} - \mathbb{E}[\phi_{it}\phi_{jt}] \} \right)^2 \right] \leq M_\Gamma, \end{aligned}$$

and (28) follows from Chebychev's inequality. Similarly, if for any fixed $T \in \mathbb{N}$ we let \mathbf{G}_T^ϕ be the $T \times T$ symmetric matrix with entries $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \mathbb{E}[\phi_{it}\phi_{is}] =: g_{ts}^\phi$, $t, s = 1, \dots, T$, then, it follows from Assumption (C-d-ii) that, as $n \rightarrow \infty$,

$$\frac{1}{T} \left\| \frac{\mathbf{\Phi}_{nT} \mathbf{\Phi}'_{nT}}{n} - \mathbf{G}_T^\phi \right\| = O_P \left(\frac{1}{\sqrt{n}} \right). \quad (29)$$

The proof of (29) is analogous to the proof of (28), thus is omitted.

Remark 10. Necessary and sufficient conditions for Assumptions (C-*d-i*) and (C-*d-ii*) to hold are

$$\begin{aligned} \frac{1}{T} \sum_{t,s=1}^T |\text{cum}(\phi_{it}, \phi_{jt}, \phi_{is}, \phi_{js})| &\leq M, \\ \frac{1}{T} \sum_{t,s=1}^T |\text{Cov}(\phi_{it}, \phi_{is})\text{Cov}(\phi_{jt}, \phi_{js})| &\leq M, \quad \frac{1}{T} \sum_{t,s=1}^T |\text{Cov}(\phi_{it}, \phi_{js})\text{Cov}(\phi_{is}, \phi_{jt})| \leq M, \end{aligned} \quad (30)$$

and

$$\begin{aligned} \frac{1}{n} \sum_{i,j=1}^n |\text{cum}(\phi_{it}, \phi_{jt}, \phi_{is}, \phi_{js})| &\leq M, \\ \frac{1}{n} \sum_{i,j=1}^n |\text{Cov}(\phi_{it}, \phi_{jt})\text{Cov}(\phi_{is}, \phi_{js})| &\leq M, \quad \frac{1}{n} \sum_{i,j=1}^n |\text{Cov}(\phi_{it}, \phi_{js})\text{Cov}(\phi_{is}, \phi_{jt})| \leq M, \end{aligned} \quad (31)$$

respectively, for some finite constant $M > 0$ (possibly different across conditions), where $\text{cum}(a, b, c, d)$ is the mixed 4-th order cumulant.⁸ Hence, conditions (30) and (31) are equivalent to summability of 4-th order moments, and can be interpreted as mixing conditions in the time series and cross-sectional sense, respectively. For (30), see also Hannan (1970, pp. 209-211), while (31) is similar to Assumption C.5 in Bai (2003), where, however, it is assumed summability also of 8-th order moments.

In fact, given the VARMA representation of ϕ_{it} , which follows from (6) and (22), and given that $\mathbb{E}[\phi_{it}^4] \leq C_0$ for some positive real C_0 independent of i and t (see below), then, it is easy to see that, under Assumption (A-*b*), the process $\{\phi_{it}\}$ is ergodic for all i , which implies that also the process $\{\phi_{it}\phi_{jt}\}$ is ergodic for all i and j (White, 2001, Theorem 3.35). Thus Assumption (C-*d-i*) holds under our setting.⁹ Although Assumption (C-*d-i*) is in principle redundant, since it could be proved using standard time series results, the cross-sectional condition in Assumption (C-*d-ii*) must be assumed, therefore, for symmetry, we prefer to assume both conditions.

Finally, in order to show that $\mathbb{E}[\phi_{it}^4] \leq C_0$, notice that, since $\phi_{it} = \mathbf{e}'_{ni} \mathbf{A}_n(L) \boldsymbol{\xi}_{nt}$ and each row of $\mathbf{A}_n(L)$ is of finite order and has only $q + 1$ non-zero elements by Assumption (B-*a*), we just need to show that $\mathbb{E}[\xi_{it}^4] \leq C_1$ for some positive real C_1 independent of i and t . To see that this is indeed the case, just notice that from (6), Assumptions (A-*b-i*) and (A-*b-ii*), and using the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned} \mathbb{E}[\xi_{it}^4] &= \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} \sum_{k_3=0}^{\infty} \sum_{k_4=0}^{\infty} \sum_{j_1=1}^{\infty} \sum_{j_2=1}^{\infty} \sum_{j_3=1}^{\infty} \sum_{j_4=1}^{\infty} \beta_{ij_1, k_1} \beta_{ij_2, k_2} \beta_{ij_3, k_3} \beta_{ij_4, k_4} \mathbb{E}[\eta_{j_1, t-k_1} \eta_{j_2, t-k_2} \eta_{j_3, t-k_3} \eta_{j_4, t-k_4}] \\ &\leq \sum_{k=0}^{\infty} \sum_{j_1=1}^{\infty} \sum_{j_2=1}^{\infty} \sum_{j_3=1}^{\infty} \sum_{j_4=1}^{\infty} B_{ij_1} B_{ij_2} B_{ij_3} B_{ij_4} \rho^{4k} \mathbb{E}[|\eta_{j_1, t-k} \eta_{j_2, t-k} \eta_{j_3, t-k} \eta_{j_4, t-k}|] \\ &\leq \frac{B^4}{1 - \rho^4} \sup_{j_1, j_2, j_3, j_4 \in \mathbb{N}} (\mathbb{E}[\eta_{j_1, t}^4])^{1/4} (\mathbb{E}[\eta_{j_2, t}^4])^{1/4} (\mathbb{E}[\eta_{j_3, t}^4])^{1/4} (\mathbb{E}[\eta_{j_4, t}^4])^{1/4} \leq \frac{B^4 M_{\eta}^4}{1 - \rho^4} =: C_1, \text{ say,} \end{aligned}$$

⁸Recall that, since $\mathbb{E}[\phi_{it}] = 0$, then its mixed 4-th order cumulant is

$$\text{cum}(\phi_{it}, \phi_{jt}, \phi_{is}, \phi_{js}) = \mathbb{E}[\phi_{it}\phi_{jt}\phi_{is}\phi_{js}] - \mathbb{E}[\phi_{it}\phi_{jt}]\mathbb{E}[\phi_{is}\phi_{js}] - \mathbb{E}[\phi_{it}\phi_{is}]\mathbb{E}[\phi_{jt}\phi_{js}] - \mathbb{E}[\phi_{it}\phi_{js}]\mathbb{E}[\phi_{is}\phi_{jt}].$$

⁹This same result is also proved by Forni et al. (2017, Lemma 11) by direct substitution of (6) and (22) into the statement of Assumption (C-*d-i*).

and since B , ρ , and M_η are independent of i and t , this result holds uniformly.

Example. MA(1) idiosyncratic components. In (21), let $\mathbf{A}_n^{(l)}(L) = \mathbf{I}_2 - \mathbf{A}^{(l)}L$; for simplicity, denote by $a_{hk}^{(l)}$, $h, k = 1, 2$ the entries of $\mathbf{A}^{(l)}$. Then, the filtered idiosyncratic components are

$$\phi_{it} = \xi_{it} - a_{11}^{(l)}\xi_{i,t-1} - a_{12}^{(l)}\xi_{j_i,t-1}, \quad (32)$$

$$\phi_{jt} = \xi_{jt} - a_{21}^{(l)}\xi_{i,t-1} - a_{22}^{(l)}\xi_{j_i,t-1}. \quad (33)$$

where $i_l := 2l-1$ and $j_l := 2l$. Assuming independence of the ξ_{it} s both across t and i , with $\text{Var}(\xi_{it}) = \sigma_i^2$, one obtains that (32)-(33) is a VMA(1), implying

$$\begin{aligned} \text{Cov}(\phi_{it}, \phi_{is}) &= \begin{cases} \sigma_i^2 + \sigma_i^2(a_{11}^{(l)})^2 + \sigma_{j_i}^2(a_{12}^{(l)})^2, & t = s, \\ -\sigma_i^2 a_{11}^{(l)}, & s = t + 1, s = t - 1, \\ 0, & \text{otherwise.} \end{cases} \\ \text{Cov}(\phi_{jt}, \phi_{js}) &= \begin{cases} \sigma_{j_l}^2 + \sigma_{i_l}^2(a_{21}^{(l)})^2 + \sigma_{j_l}^2(a_{22}^{(l)})^2, & t = s, \\ -\sigma_{j_l}^2 a_{22}^{(l)}, & s = t + 1, s = t - 1, \\ 0, & \text{otherwise.} \end{cases} \\ \text{Cov}(\phi_{it}, \phi_{js}) &= \begin{cases} \sigma_i^2(a_{11}^{(l)}a_{21}^{(l)}) + \sigma_{j_i}^2(a_{12}^{(l)}a_{22}^{(l)}), & t = s, \\ -\sigma_i^2 a_{21}^{(l)}, & s = t + 1, \\ -\sigma_{j_i}^2 a_{12}^{(l)}, & s = t - 1, \\ 0, & \text{otherwise.} \end{cases} \\ \text{Cov}(\phi_{it}, \phi_{jms}) &= 0, \quad \text{for all } t, s \text{ and } m \neq l. \end{aligned}$$

The fourth order cumulants $\text{cum}(\phi_{it}, \phi_{jt}, \phi_{is}, \phi_{js})$ are always zero for units (i, j) belonging to distinct bivariate VARs, and not zero only when considering units within the same VAR for $\{i_l = j_l, t = s\}$, $\{i_l \neq j_l, t = s\}$, and $\{i_l = j_l, t = s \pm 1\}$.

Given that, by the cumulants theorem (see Brillinger, 2001, Theorem 2.3.2, Equation 2.3.7),

$$\begin{aligned} \mathbb{E}[(\phi_{it}\phi_{jt} - \gamma_{i_j}^\phi)(\phi_{is}\phi_{js} - \gamma_{i_j}^\phi)] &= \text{cum}(\phi_{it}, \phi_{jt}, \phi_{is}, \phi_{js}) + \text{Cov}(\phi_{it}, \phi_{is})\text{Cov}(\phi_{jt}, \phi_{js}) \\ &\quad + \text{Cov}(\phi_{it}, \phi_{js})\text{Cov}(\phi_{jt}, \phi_{is}), \end{aligned}$$

and defining $\kappa_4(i_l) := \text{cum}(\phi_{i_l t}, \phi_{i_l t}, \phi_{i_l t}, \phi_{i_l t})$, it follows that (28) holds in view of

$$\begin{aligned} &\sum_{t,s=1}^T \mathbb{E}[(\phi_{it}\phi_{jt} - \gamma_{i_j}^\phi)(\phi_{is}\phi_{js} - \gamma_{i_j}^\phi)] \\ &= T(\sigma_i^2 + \sigma_i^2(a_{11}^{(l)})^2 + \sigma_{j_i}^2(a_{12}^{(l)})^2)(\sigma_{j_l}^2 + \sigma_{i_l}^2(a_{21}^{(l)})^2 + \sigma_{j_l}^2(a_{22}^{(l)})^2) + 2(T-1)a_{11}^{(l)}a_{21}^{(l)} \\ &\quad + T\sigma_i^2\sigma_{j_l}^2(\sigma_i^2(a_{11}^{(l)}a_{22}^{(l)}) + \sigma_{j_l}^2(a_{12}^{(l)}a_{22}^{(l)}))^2 + 2(T-1)\sigma_i^2\sigma_{j_l}^2a_{21}^{(l)}a_{12}^{(l)} \\ &\quad + T\mathbb{I}_{i=j_l}(\kappa_4(i_l)(1 + (a_{11}^{(l)})^4 + (a_{12}^{(l)})^4) + \kappa_4(j_l)(1 + (a_{21}^{(l)})^4 + (a_{22}^{(l)})^4)) \\ &\quad + T\mathbb{I}_{i \neq j_l}(\kappa_4(i_l)(a_{11}^{(l)})^2(a_{12}^{(l)})^2 + \kappa_4(j_l)(a_{21}^{(l)})^2(a_{22}^{(l)})^2) \\ &\quad + 2T\mathbb{I}_{i=j_l}\kappa_4(i_l)((a_{11}^{(l)})^2 + (a_{21}^{(l)})^2) + T\mathbb{I}_{i \neq j_l}(\kappa_4(i_l)(a_{12}^{(l)})^2 + \kappa_4(j_l)(a_{21}^{(l)})^2) \leq TM \end{aligned}$$

($\mathbb{I}_{[\dots]}$ the indicator of $[\dots]$), which holds as long as $\sup_l(|a_{11}^{(l)}| + |a_{12}^{(l)}| + |a_{21}^{(l)}| + |a_{22}^{(l)}|) + \sup_l \sigma_i^2 + \sup_l \kappa_4(i_l)$

is bounded, which is always satisfied under our assumptions, since, as shown in Remark 10, the fourth order cumulants are bounded and the VAR coefficients are also bounded by Assumption (B-a), i.e., stationarity.

Similarly, given the fact that

$$\begin{aligned}\mathbb{E}[(\phi_{it}\phi_{is} - g_{ts}^\phi)(\phi_{jt}\phi_{js} - g_{ts}^\phi)] &= \text{cum}(\phi_{it}, \phi_{jt}, \phi_{is}, \phi_{js}) + \mathbb{C}\text{ov}(\phi_{it}, \phi_{jt})\mathbb{C}\text{ov}(\phi_{is}, \phi_{js}) \\ &\quad + \mathbb{C}\text{ov}(\phi_{it}, \phi_{js})\mathbb{C}\text{ov}(\phi_{jt}, \phi_{is}),\end{aligned}$$

it follows that (29) holds since the ϕ_{it} s have an MA(1) structure across units, i.e.,

$$\begin{aligned}\sum_{i,j=1}^n \text{cum}(\phi_{it}, \phi_{jt}, \phi_{is}, \phi_{js}) &= \sum_{l=1}^m \text{cum}(\phi_{it}, \phi_{jt}, \phi_{is}, \phi_{js}) \\ &\leq \mathbb{I}_{t=s} \sum_{l=1}^m (\kappa_4(i_l)(1 + (a_{11}^{(l)})^4 + (a_{12}^{(l)})^4) + \kappa_4(j_l)(1 + (a_{21}^{(l)})^4 + (a_{22}^{(l)})^4)) \\ &\quad + \mathbb{I}_{t=s} \sum_{l=1}^m (\kappa_4(i_l)(a_{11}^{(l)})^2(a_{12}^{(l)})^2 + \kappa_4(j_l)(a_{21}^{(l)})^2(a_{22}^{(l)})^2) \\ &\quad + (\mathbb{I}_{t=s-1} + \mathbb{I}_{s=t-1}) \sum_{l=1}^m (2\kappa_4(i_l)(a_{11}^{(l)})^2 + 2\kappa_4(j_l)(a_{21}^{(l)})^2) \\ &\quad + \mathbb{I}_{t=s-1} \sum_{l=1}^m \kappa_4(j_l)(a_{12}^{(l)})^2 + \mathbb{I}_{s=t-1} \sum_{l=1}^m \kappa_4(i_l)(a_{21}^{(l)})^2 \leq mM = n(M/2),\end{aligned}\quad (34)$$

$$\begin{aligned}\sum_{i,j=1}^n \mathbb{C}\text{ov}(\phi_{it}, \phi_{jt})\mathbb{C}\text{ov}(\phi_{is}, \phi_{js}) &= 2 \sum_{l=1}^m \mathbb{C}\text{ov}(\phi_{it}, \phi_{jt})\mathbb{C}\text{ov}(\phi_{is}, \phi_{js}) + \sum_{l=1}^m \text{Var}(\phi_{it})\text{Var}(\phi_{is}) \\ &= 2 \sum_{l=1}^m (\sigma_{i_l}^2(a_{11}^{(l)}a_{21}^{(l)} + \sigma_{j_l}^2(a_{12}^{(l)}a_{22}^{(l)}))^2 + \sum_{l=1}^m (\sigma_{i_l}^2 + \sigma_{i_l}^2(a_{11}^{(l)})^2 + \sigma_{j_l}^2(a_{12}^{(l)})^2)^2 \\ &\leq mM = n(M/2),\end{aligned}\quad (35)$$

and

$$\begin{aligned}\sum_{i,j=1}^n \mathbb{C}\text{ov}(\phi_{it}, \phi_{js})\mathbb{C}\text{ov}(\phi_{jt}, \phi_{is}) &= 2 \sum_{l=1}^m \mathbb{C}\text{ov}(\phi_{it}, \phi_{js})\mathbb{C}\text{ov}(\phi_{jt}, \phi_{is}) + \sum_{l=1}^m \mathbb{C}\text{ov}(\phi_{it}, \phi_{is})\mathbb{C}\text{ov}(\phi_{jt}, \phi_{js}) \\ &= 2 \sum_{l=1}^m (\sigma_{i_l}^2 a_{21}^{(l)} \sigma_{j_l}^2 a_{12}^{(l)}) + \sum_{l=1}^m (\sigma_{i_l}^2 a_{11}^{(l)} \sigma_{j_l}^2 a_{22}^{(l)}) \leq mM = n(M/2),\end{aligned}\quad (36)$$

where the bounds (34)-(36) hold whenever the elements of $\mathbf{A}^{(l)}(L)$ in (21) satisfy $\sum_{l=1}^m \|\mathbf{A}^{(l)}\|^s \leq mM$ for $s = 1, 2, 3, 4$, which follows from Assumption (B-a), i.e., by stationarity.

The two static representations (26)-(27) contain common components which are static across time and exchangeable across units, respectively. Therefore, for our asymptotic analysis, it suffices to study the behavior of their sample second moment across time and the cross-section, respectively, rather than their spectral density matrices. This is developed in the next section. At the same time, one needs to ensure that the idiosyncratic components in (26)-(27) are indeed idiosyncratic, i.e., are weakly dependent

both across time and units. This is taken care of in the subsequent section.

4.1 Static Common Components

From Assumption (A-*a-i*) and the Weak Law of Large Numbers, it immediately follows that

$$\frac{1}{T} \sum_{t=1}^T \mathbf{u}_t \mathbf{u}_t' = \frac{\mathbf{U}'_T \mathbf{U}_T}{T} \xrightarrow{P} \mathbf{\Gamma}^u \text{ as } T \rightarrow \infty, \quad (37)$$

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 Assumption (C-*b*), and the Weak Law of Large Numbers, there exists a finite $q \times q$ positive definite matrix $\mathbf{\Sigma}^R$ such that

$$\frac{1}{n} \sum_{i=1}^n \mathbf{R}_i \mathbf{R}_i' = \frac{\mathbf{R}'_n \mathbf{R}_n}{n} \xrightarrow{P} \mathbf{\Sigma}^R \text{ as } n \rightarrow \infty, \quad (38)$$

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 (37) and (38) are \sqrt{T} and \sqrt{n} , respectively (see Lemma 1 in the Appendix).

Remark 11. Assumption (C-*b*) and its consequence (38) pave the way for clarifying the issue of identification of the common shocks and their IRFs as discussed in Remark 7. In particular, full rank of $\mathbf{R}'_n \mathbf{R}_n/n$, for n sufficiently large, which requires a sufficient degree of heterogeneity of the elements of \mathbf{R}_n , implies that the space spanned by the q elements of \mathbf{u}_t is identified, or equivalently that the q elements of \mathbf{u}_t are identified up to an invertible linear transformation \mathbf{H} . It follows from (15) that the \mathbf{u}_t s are *fundamental* with respect to the χ_{nts} —that is, the space spanned by their present and past values coincides with the space spanned by the present and past values of χ_{nt} . In turn, this means that identification is reduced to the choice of a $q \times q$ invertible matrix \mathbf{H} such that economically motivated restrictions on the identified IRFs matrix $\mathbf{B}(L)\mathbf{H} = [\mathbf{A}_n(L)]^{-1} \mathbf{R}_n \mathbf{H}$ hold. This is achieved, for instance, by maximizing or minimizing an objective function involving $\mathbf{B}(L)\mathbf{H}$ or, alternatively, by imposing zero restrictions on its elements (see Forni et al., 2009 for a detailed discussion and examples). Notice that, under our assumptions, in particular ensuring a sufficient degree of heterogeneity of the elements of \mathbf{R}_n , the number of economic identification restrictions needed depends only on q and not on n .

Now, from (28), which follows from Assumption (C-*d-i*), (37), and Assumption (C-*c*),

$$\frac{\mathbf{Z}'_{nT} \mathbf{Z}_{nT}}{T} = \frac{\mathbf{R}_n \mathbf{U}'_T \mathbf{U}_T \mathbf{R}'_n}{T} + \frac{\mathbf{\Phi}'_{nT} \mathbf{\Phi}_{nT}}{T} + o_P(1) \xrightarrow{P} \mathbf{R}_n \mathbf{\Gamma}^u \mathbf{R}'_n + \mathbf{\Gamma}_n^\phi, \text{ as } T \rightarrow \infty. \quad (39)$$

Letting μ_{nj}^ψ denote the j th largest eigenvalue of $\mathbf{R}_n \mathbf{\Gamma}^u \mathbf{R}'_n$, because of (38) and since $\mathbf{\Gamma}^u$ is positive definite, for all $j = 1, \dots, q$, there exist two positive reals $\underline{\mu}_j^\psi$ and $\bar{\mu}_j^\psi$ such that

$$\underline{\mu}_j^\psi \leq \text{p-lim inf}_{n \rightarrow \infty} \frac{\mu_{nj}^\psi}{n} \leq \text{p-lim sup}_{n \rightarrow \infty} \frac{\mu_{nj}^\psi}{n} \leq \bar{\mu}_j^\psi. \quad (40)$$

This is similar to Assumption 6 in Forni et al. (2017), although here derived from our assumptions.

Likewise, from (29), which follows from Assumption (C-*d-ii*), (38), and Assumption (C-*c*),

$$\frac{\mathbf{Z}_{nT}\mathbf{Z}'_{nT}}{n} = \frac{\mathbf{U}_T\mathbf{R}'_n\mathbf{R}_n\mathbf{U}'_T}{n} + \frac{\mathbf{\Phi}_{nT}\mathbf{\Phi}'_{nT}}{n} + o_P(1) \rightarrow_P \mathbf{U}_T\mathbf{\Sigma}^R\mathbf{U}'_T + \mathbf{G}_T^\phi, \text{ as } n \rightarrow \infty. \quad (41)$$

Letting ν_{Tj}^ψ denote the j largest eigenvalue of $\mathbf{U}_T\mathbf{\Sigma}^R\mathbf{U}'_T$, because of (37) and since $\mathbf{\Sigma}^R$ is positive definite, for all $j = 1, \dots, q$, there exist two positive reals $\underline{\nu}_j^\psi$ and $\bar{\nu}_j^\psi$ such that

$$\underline{\nu}_j^\psi \leq \text{p-lim inf}_{T \rightarrow \infty} \frac{\nu_{Tj}^\psi}{T} \leq \text{p-lim sup}_{T \rightarrow \infty} \frac{\nu_{Tj}^\psi}{T} \leq \bar{\nu}_j^\psi. \quad (42)$$

In fact, by the Strong Law of Large Numbers, (37) and (38) hold also almost surely and weak convergence statements in Lemma 1 of the Appendix could be replaced by almost sure ones with convergence rates $O(T^{1/2-\epsilon})$ and $O(n^{1/2-\epsilon})$ for some arbitrarily small $\epsilon > 0$.¹⁰ As a consequence, the eigenvalue properties (40) and (42) could be shown to hold with probability one, as in the classical factor model literature.

Consistent estimation of eigenvectors, however, also requires the usual assumption of asymptotic separation of eigenvalues—a slight reinforcement of (40) and (42).

Assumption (D). For all $j = 2, \dots, q$, $\bar{\mu}_j^\psi < \underline{\mu}_{j-1}^\psi$ and $\bar{\nu}_j^\psi < \underline{\nu}_{j-1}^\psi$.

This assumption, together with (40) and (42), is the analog, for the covariance matrix of the static common component ψ_n of the assumption made on the eigenvalues of the spectral density matrix of the common component χ_n (Assumption (A-*a-iii*)).

4.2 Static Idiosyncratic Components

From (28) and (29), we have that $\mathbf{\Gamma}_n^\phi = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \mathbb{E}[\phi_{nt}\phi'_{nt}] = \mathbb{E}[\phi_{nt}\phi'_{nt}]$ because of stationarity, while $\mathbf{G}_T^\phi = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n \mathbb{E}[\varphi_T^i \varphi_T^{i'}]$. Therefore, for any $\mathbf{b}_n = (b_1 \cdots b_n)'$ such that $\mathbf{b}'_n \mathbf{b}_n = 1$,¹¹

$$\begin{aligned} \sup_{n \in \mathbb{N}} \mathbf{b}'_n \mathbf{\Gamma}_n^\phi \mathbf{b}_n &= \sup_{n \in \mathbb{N}} \mathbf{b}'_n \mathbb{E}[\phi_{nt}\phi'_{nt}] \mathbf{b}_n = \sup_{n \in \mathbb{N}} \sum_{i,j=1}^n b_i b_j \int_{-\pi}^{\pi} \sigma_{ij}^\phi(\theta) d\theta \\ &\leq \sup_{n \in \mathbb{N}} \sum_{i,j=1}^n |b_i b_j| \int_{-\pi}^{\pi} |\sigma_{ij}(\theta)| d\theta \leq \sup_{n \in \mathbb{N}} \sum_{i=1}^n |b_i|^2 2\pi B^\phi = 2\pi B^\phi, \end{aligned} \quad (43)$$

and, for any $\mathbf{c}_T = (c_1 \cdots c_T)'$ such that $\mathbf{c}'_T \mathbf{c}_T = 1$,¹²

¹⁰More precisely, it is possible to show that $\left\| \frac{\mathbf{U}'_T \mathbf{U}_T}{T} - \mathbf{\Gamma}^u \right\| = O_{\text{a.s.}} \left(\frac{\log \log T}{\sqrt{T}} \right)$ and $\left\| \frac{\mathbf{R}'_n \mathbf{R}_n}{n} - \mathbf{\Sigma}^R \right\| = O_{\text{a.s.}} \left(\frac{\log \log n}{\sqrt{n}} \right)$.

¹¹Because of (23), the diagonal entries of $\mathbf{\Sigma}_n^\phi(\theta)$ are such that

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

since eigenvectors are normalized. Notice that $\sigma_{ii}^\phi(\theta)$ is real and positive.

$$\begin{aligned}
\sup_{T \in \mathbb{N}} \mathbf{c}'_T \mathbf{G}_T^\phi \mathbf{c}_T &= \sup_{T \in \mathbb{N}} \mathbf{c}'_T \left\{ \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \mathbb{E}[\varphi_T^i \varphi_T^{i'}] \right\} \mathbf{c}_T = \sup_{T \in \mathbb{N}} \sum_{t,s=1}^T c_t c_s \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \left\{ \int_{-\pi}^{\pi} \sigma_{ii}^\phi(\theta) e^{\iota(t-s)\theta} d\theta \right\} \\
&\leq \sup_{T \in \mathbb{N}} \sum_{t,s=1}^T |c_t c_s| \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \left\{ \int_{-\pi}^{\pi} |\sigma_{ii}(\theta)| |e^{\iota(t-s)\theta}| d\theta \right\} \\
&\leq \sup_{T \in \mathbb{N}} \sum_{t,s=1}^T |c_t c_s| \sup_{i \in \mathbb{N}} \left\{ \int_{-\pi}^{\pi} |\sigma_{ii}(\theta)| |e^{\iota(t-s)\theta}| d\theta \right\} \leq \sup_{T \in \mathbb{N}} \sum_{t=1}^T |c_t|^2 2\pi B^\phi = 2\pi B^\phi. \tag{44}
\end{aligned}$$

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

$$\sup_{n \in \mathbb{N}} \|\mathbf{\Gamma}_n^\phi\| = \sup_{n \in \mathbb{N}} \max_{\substack{\mathbf{b}_n \\ \mathbf{b}'_n \mathbf{b}_n = 1}} \mathbf{b}'_n \mathbf{\Gamma}_n^\phi \mathbf{b}_n \leq 2\pi B^\phi \text{ and} \tag{45}$$

$$\sup_{T \in \mathbb{N}} \|\mathbf{G}_T^\phi\| = \sup_{T \in \mathbb{N}} \max_{\substack{\mathbf{c}'_T \\ \mathbf{c}'_T \mathbf{c}_T = 1}} \mathbf{c}'_T \mathbf{G}_T^\phi \mathbf{c}_T \leq 2\pi B^\phi, \tag{46}$$

respectively. Following a similar reasoning, it is straightforward to show that also Assumptions C2 and C3 of Bai (2003) hold.

5 Estimation

In order to estimate the common component, we need to estimate the common loading filters, i.e., the impulse response functions $\mathbf{B}_n(L) = [\mathbf{A}_n(L)]^{-1} \mathbf{R}_n$ and the common factors or shocks \mathbf{u}_t . That estimation proceeds in two steps: first we estimate $\mathbf{A}_n(L)$ and then, by considering the static representation (22) of the GDFM, we estimate \mathbf{u}_t and \mathbf{R}_n by performing 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 6 is devoted to their asymptotic properties.¹³

Throughout, we denote by \hat{q} a consistent estimator of the number q of factors; such an estimator can be obtained by means of the Hallin and Liška (2007) information criterion applied to the observed data matrix \mathbf{X}_{nT} or, alternatively, via the methods proposed by Onatski (2009) or Avarucci et al. (2022).

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

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

$$\hat{\Sigma}_n(\theta_h) := \frac{1}{2\pi} \sum_{k=-T+1}^{T-1} \mathcal{K} \left(\frac{k}{B_T} \right) e^{-\iota k \theta_h} \hat{\Gamma}_{n,k}, \quad \theta_h = \frac{\pi h}{B_T}, \quad |h| \leq B_T, \tag{47}$$

¹²Because of (23) and the Cauchy-Schwarz inequality, the off-diagonal entries of $\Sigma_n^\phi(\theta)$ satisfy

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

¹³We denote estimated quantities as $\hat{\Sigma}_n(\theta)$, $\hat{\Gamma}_n$, $\hat{\mathbf{A}}_n(L)$, $\hat{\mathbf{R}}_n$, etc. The use of “hats” implicitly implies dependence both on n and T , not to be confused with the use of the index n and/or T which refers to the dimension of the object.

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 (assuming to work with centered data, i.e., such that $T^{-1} \sum_{t=1}^T \mathbf{x}_{nt} = \mathbf{0}$) and $\mathcal{K}(\cdot)$ is a suitable kernel with bandwidth B_T . The choice of the kernel and its bandwidth B_T are standard problems and in our empirical analysis we use the classical choice of a Bartlett kernel with bandwidth $B_T = \lfloor T^{1/3} \rfloor$, a choice which is compatible with our theoretical results (see Assumption (K) in Section 6.2). The asymptotic properties of $\widehat{\mathbf{\Sigma}}_n(\theta_h)$, defined in (47), are given in Proposition 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 \widehat{q} largest eigenvalues of $\widehat{\mathbf{\Sigma}}_n(\theta_h)$ into the $(n \times \widehat{q})$ matrix $\widehat{\mathbf{P}}_n(\theta_h)$, and the corresponding eigenvalues into the $(\widehat{q} \times \widehat{q})$ diagonal matrix $\widehat{\mathbf{\Lambda}}_n(\theta_h)$. Our estimator of the spectral density matrix of the common component is $\widehat{\mathbf{\Sigma}}_n^\chi(\theta_h) := \widehat{\mathbf{P}}_n(\theta_h) \widehat{\mathbf{\Lambda}}_n(\theta_h) \widehat{\mathbf{P}}_n^\dagger(\theta_h)$, 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:

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

Consider the m diagonal $(\widehat{q} + 1) \times (\widehat{q} + 1)$ blocks $\widehat{\mathbf{\Gamma}}_k^{\chi(s)}$ of the $\widehat{\mathbf{\Gamma}}_{n,k}^\chi$'s. For each block, we estimate, via the Yule-Walker method, the coefficients of a $(\widehat{q} + 1)$ -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 (B).¹⁴ By combining the m estimators for the m diagonal blocks $\mathbf{A}^{(1)}(L), \dots, \mathbf{A}^{(m)}(L)$, we obtain an estimator $\widehat{\mathbf{A}}_n(L)$ of the VAR filter $\mathbf{A}_n(L)$ as defined in (13); the asymptotic properties of that estimator are given in Proposition 2.

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

Remark 12. 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 (C-a), 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 13. Although we assumed, for simplicity, that $n = m(\widehat{q} + 1)$ for some integer m , this might not be the case in practice. When n is not an integer multiple of $(\widehat{q} + 1)$, we can consider $\lfloor n/(\widehat{q} + 1) \rfloor - 1$ blocks of size $(\widehat{q} + 1)$ and a last one of size $(\widehat{q} + 1) + n - \lfloor n/(\widehat{q} + 1) \rfloor (\widehat{q} + 1)$, which is larger than $(\widehat{q} + 1)$ but smaller than $2(\widehat{q} + 1)$. Since the arguments from Forni et al. (2017) used in the next section apply to any partition into blocks of size $(\widehat{q} + 1)$ or larger, nothing changes for the subsequent asymptotic theory.

Remark 14. 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 $p_s(\widehat{q} + 1) \times p_s(\widehat{q} + 1)$ Toeplitz matrix.

¹⁴For example, in the VAR(1) case, i.e., $p_s = 1$, we have $\widehat{\mathbf{A}}^{(s)}(L) = \mathbf{I}_{\widehat{q}+1} - \widehat{\mathbf{A}}^{(s)} L$ with $\widehat{\mathbf{A}}^{(s)} := \widehat{\mathbf{\Gamma}}_1^{\chi(s)} (\widehat{\mathbf{\Gamma}}_0^{\chi(s)})^{-1}$.

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.

5.2 Estimation of \mathbf{U}_T and \mathcal{R}_n

We now describe how to estimate the the common component (the product of \mathbf{U}_T and \mathcal{R}_n) in the static representation (25) of the GDFM. It is well known (see Bai, 2003) that estimation of \mathbf{U}_T and \mathcal{R}_n involves (i) the evaluation of the (suitably normalized) eigenvectors of the sample covariance matrix of $\widehat{\mathbf{Z}}'_{nT} := \widehat{\mathbf{A}}'_n(L)\mathbf{X}'_{nT}$, and (ii) the construction of a linear projection of the $\widehat{\mathbf{Z}}_{nT}$ s onto a subset of the sample eigenvectors obtained in step (i), corresponding to the estimated number \widehat{q} of factors. By the duality outlined above, this can be done in two different ways, depending on whether the $n \times n$ sample covariance matrix $\widehat{\mathbf{Z}}'_{nT}\widehat{\mathbf{Z}}_{nT}/T$ or the $T \times T$ covariance matrix $\widehat{\mathbf{Z}}'_{nT}\widehat{\mathbf{Z}}_{nT}/n$ is considered. In particular, in the former case of the $n \times n$ sample covariance matrix, the estimator of \mathbf{U}_T , which we denote as $\widehat{\mathbf{U}}_T$, is obtained as a linear projection of the $\widehat{\mathbf{Z}}_{nT}$ onto the n -dimensional sample eigenvectors that are used to build an estimator of \mathcal{R}_n , denoted as $\check{\mathcal{R}}_n$. Our main insight is to exploit the linearity rooted in the linear projection operator to derive the limiting distribution of the estimator for \mathbf{U}_T , relying on the consistency of the sample eigenvectors estimator established in Forni et al. (2017). The same reasoning, now based on the $T \times T$ sample covariance matrix, permits to deriving the limiting distribution of a (different) estimator of \mathcal{R}_n , denoted as $\widehat{\mathcal{R}}_n$, which has the desired linear projection form, and is a function of the T -dimensional sample eigenvectors of the covariance matrix used to build a (different) estimator of \mathbf{U}_T , which we denote as $\check{\mathbf{U}}_T$. This procedure is warranted by the duality that characterizes the static model (25) as discussed above, leading to the representations (26) and (27).

Therefore, one has two sets of estimators for \mathbf{U}_T and \mathcal{R}_n . When inference is to be performed on either \mathbf{U}_T or \mathcal{R}_n , we focus on the asymptotic properties of the linear projection estimators $\widehat{\mathbf{U}}_T$ and $\widehat{\mathcal{R}}_n$ given in Theorems 1 and 2, respectively. However, for estimating the static common component Ψ_{nT} defined in (25), we will combine the two estimators $\widehat{\mathbf{U}}_T\check{\mathcal{R}}'_n$ and $\check{\mathbf{U}}_T\widehat{\mathcal{R}}'_n$, thus enjoying an efficiency gain (see the next section).¹⁵ To this end, we will derive also the asymptotic properties of the eigenvectors estimators $\check{\mathbf{U}}_T$ and $\check{\mathcal{R}}_n$ (see Appendix C).

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

$$\widehat{\mathbf{\Gamma}}_n^z := \frac{1}{T} \sum_{t=1}^T \widehat{\mathbf{z}}_{nt}\widehat{\mathbf{z}}'_{nt} = \frac{\widehat{\mathbf{Z}}'_{nT}\widehat{\mathbf{Z}}_{nT}}{T}. \quad (48)$$

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

$$\check{\mathcal{R}}_n = (\check{\mathbf{R}}_1 \cdots \check{\mathbf{R}}_n)' := \widehat{\mathbf{P}}_n^z \left(\widehat{\mathbf{\Lambda}}_n^z \right)^{1/2}. \quad (49)$$

Next consider the submatrix of $\check{\mathcal{R}}_n$ consisting of a selection of $\bar{n} \leq n$ rows. The asymptotic conditions \bar{n} should satisfy are discussed in Section 6.3, while the specific choice of \bar{n} in finite samples is considered

¹⁵Note that $\widehat{\mathbf{U}}_T\widehat{\mathcal{R}}'_n$ is not convenient as an estimator of Ψ_{nT} because $\widehat{\mathbf{U}}_T$ and $\widehat{\mathcal{R}}_n$ are, as expected, affected by two different (and unknown) rotation matrices. For the same reason, we cannot consider the estimator $\check{\mathbf{U}}_T\check{\mathcal{R}}'_n$.

in Section 7. Without loss of generality, assume that the first \bar{n} rows are selected, and define

$$\check{\mathbf{R}}_{\bar{n}} = (\check{\mathbf{R}}_1 \cdots \check{\mathbf{R}}_{\bar{n}})' := \widehat{\mathbf{P}}_{\bar{n}}^z \left(\widehat{\mathbf{\Lambda}}_n^z \right)^{1/2} \quad (50)$$

where $\widehat{\mathbf{P}}_{\bar{n}}^z$ is the $\bar{n} \times \hat{q}$ submatrix of $\widehat{\mathbf{P}}_n^z$'s first \bar{n} rows. Note that each entry of $\check{\mathbf{R}}_{\bar{n}}$ still is a 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 $\widehat{\mathbf{Z}}_{\bar{n}T} = (\widehat{\mathbf{z}}_T^1 \cdots \widehat{\mathbf{z}}_T^{\bar{n}})$ be the $T \times \bar{n}$ matrix of $\widehat{\mathbf{Z}}_{nT}$'s first \bar{n} columns. We estimate \mathbf{U}_T as the cross-sectional linear projection $\widehat{\mathbf{U}}_T$ of the $\widehat{\mathbf{z}}_T^i$ s onto $\check{\mathbf{R}}_{\bar{n}}$: namely,

$$\begin{aligned} \widehat{\mathbf{U}}_T &= (\widehat{\mathbf{u}}_1 \cdots \widehat{\mathbf{u}}_t \cdots \widehat{\mathbf{u}}_T)' := \widehat{\mathbf{Z}}_{\bar{n}T} \check{\mathbf{R}}_{\bar{n}} \left(\check{\mathbf{R}}_{\bar{n}}' \check{\mathbf{R}}_{\bar{n}} \right)^{-1} \\ &= \widehat{\mathbf{Z}}_{\bar{n}T} \widehat{\mathbf{P}}_{\bar{n}}^z \left(\widehat{\mathbf{\Lambda}}_n^z \right)^{1/2} \left(\left(\widehat{\mathbf{\Lambda}}_n^z \right)^{1/2} \widehat{\mathbf{P}}_{\bar{n}}^z \widehat{\mathbf{P}}_{\bar{n}}^z \left(\widehat{\mathbf{\Lambda}}_n^z \right)^{1/2} \right)^{-1} = \widehat{\mathbf{Z}}_{\bar{n}T} \widehat{\mathbf{P}}_{\bar{n}}^z \left(\widehat{\mathbf{\Lambda}}_n^z \right)^{-1/2}. \end{aligned} \quad (51)$$

This is the estimator we are proposing for \mathbf{U}_T , for which we provide the asymptotic properties in detail below.

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

$$\widehat{\mathbf{G}}_T^z := \frac{1}{n} \sum_{i=1}^n \widehat{\mathbf{z}}_T^i \widehat{\mathbf{z}}_T^{i'} = \frac{\widehat{\mathbf{Z}}_{nT} \widehat{\mathbf{Z}}_{nT}'}{n}. \quad (52)$$

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

$$\check{\mathbf{U}}_T = (\check{\mathbf{u}}_1 \cdots \check{\mathbf{u}}_T)' := \widehat{\mathbf{\Pi}}_T^z \left(\widehat{\mathbf{L}}_T^z \right)^{1/2}. \quad (53)$$

Next consider the submatrix of $\check{\mathbf{U}}_T$ consisting of a selection of $\bar{T} \leq T$ rows. The asymptotic conditions \bar{T} should satisfy are discussed in Section 6.4, while the specific choice of \bar{T} in finite samples is considered in Section 7. Without loss of generality, assume that the first \bar{T} rows are selected, and define

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

where $\widehat{\mathbf{\Pi}}_{\bar{T}}^z$ is the $\bar{T} \times \hat{q}$ submatrix of $\widehat{\mathbf{\Pi}}_T^z$'s first \bar{T} rows. Note that each entry of $\check{\mathbf{U}}_{\bar{T}}$ continues to be function of n and T only; in particular the matrix of eigenvalues $\widehat{\mathbf{L}}_{\bar{T}}^z$ 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 \mathbf{R}_n as the time-series linear projection $\widehat{\mathbf{R}}_n$ of the $\widehat{\mathbf{z}}_{nt}$'s onto $\check{\mathbf{U}}_{\bar{T}}$: namely,

$$\begin{aligned} \widehat{\mathbf{R}}_n &= \left(\widehat{\mathbf{R}}_1 \cdots \widehat{\mathbf{R}}_i \cdots \widehat{\mathbf{R}}_n \right)' := \widehat{\mathbf{Z}}_{n\bar{T}}' \check{\mathbf{U}}_{\bar{T}} \left(\check{\mathbf{U}}_{\bar{T}}' \check{\mathbf{U}}_{\bar{T}} \right)^{-1} \\ &= \widehat{\mathbf{Z}}_{n\bar{T}}' \widehat{\mathbf{\Pi}}_{\bar{T}}^z \left(\widehat{\mathbf{L}}_{\bar{T}}^z \right)^{1/2} \left(\left(\widehat{\mathbf{L}}_{\bar{T}}^z \right)^{1/2} \widehat{\mathbf{\Pi}}_{\bar{T}}^z \widehat{\mathbf{\Pi}}_{\bar{T}}^z \left(\widehat{\mathbf{L}}_{\bar{T}}^z \right)^{1/2} \right)^{-1} \\ &= \widehat{\mathbf{Z}}_{n\bar{T}}' \widehat{\mathbf{\Pi}}_{\bar{T}}^z \left(\widehat{\mathbf{L}}_{\bar{T}}^z \right)^{-1/2}. \end{aligned} \quad (54)$$

This is the estimator we are proposing for \mathbf{R}_n , the asymptotic properties of which are developed below.

Summing up, we have two sets of estimators for \mathbf{u}_T , namely $\hat{\mathbf{u}}_T$ and $\check{\mathbf{u}}_T$, and two sets of estimators for \mathcal{R}_n , namely $\hat{\mathcal{R}}_n$ and $\check{\mathcal{R}}_n$. As mentioned above, for the purpose of inference on \mathbf{u}_T and \mathcal{R}_n , we will consider $\hat{\mathbf{u}}_T$ and $\hat{\mathcal{R}}_n$ only. Instead, to conduct inference on Ψ_{nT} , we will also consider $\check{\mathcal{R}}_n$ and $\check{\mathbf{u}}_T$.

5.3 Estimation of ψ_{it} and χ_{it}

To estimate the static common component ψ_{it} , we make full use of the four estimators (two for \mathbf{u}_T and two for \mathcal{R}_n) defined in the previous section. We either have $\hat{\psi}_{it} := \hat{\mathbf{R}}_i' \hat{\mathbf{u}}_t$ and/or $\check{\psi}_{it} := \check{\mathbf{R}}_i' \hat{\mathbf{u}}_t$. However, as discussed in Section 6.5 below, efficient estimators are convex linear combinations of the form

$$\hat{\psi}_{it} := \omega_{nT} \check{\mathbf{R}}_i' \hat{\mathbf{u}}_t + (1 - \omega_{nT}) \hat{\mathbf{R}}_i' \check{\mathbf{u}}_t, \quad i = 1, \dots, n, t = 1, \dots, T, \quad (55)$$

where the weight ω_{nT} is such that $\omega_{nT} = 1/2$ if $n = T$, $\omega_{nT} \uparrow 1$ if $n/T \downarrow 0$, and $\omega_{nT} \downarrow 0$ if $T/n \downarrow 0$, to take advantage of the dimension of the panel. The asymptotic properties of $\hat{\psi}_{it}$ defined in (55) are given in Theorem 3.

Finally, recalling the definitions $\mathbf{C}_n(L) := [\mathbf{A}_n(L)]^{-1}$ and $\mathcal{I}_s := \{\ell \mid \ell = (s-1)(\hat{q}+1)+1, \dots, s(\hat{q}+1)\}$, the set of integers indexing the series belonging to block s , $s = 1, \dots, m$, (see (16)), our estimator of χ_{it} is

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

where K is a finite integer, $\hat{c}_{i,j_s,k}$ is the (i, j_s) th entry of the k th coefficient matrix of $\hat{\mathbf{C}}_n(L) := [\hat{\mathbf{A}}_n(L)]^{-1}$, and $\hat{\psi}_{j_s,t-k}$ is the estimator of the static common component defined in (55). Notice that, in (56), we sum over a finite number of lags K only 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 uniformly negligible. This is the same standard problem that arises when reporting IRFs after estimating a VAR. In our empirical analysis, we set $K = 20$. The asymptotic properties of $\hat{\chi}_{it}$ defined in (56) are given in Theorem 4.

6 Asymptotic properties

6.1 The number of factors

Any of the estimators of q available in the literature, generically denoted as \hat{q} , converges in probability to q as $n, T \rightarrow \infty$ (see Bai and Ng, 2002; Hallin and Liška, 2007; Onatski, 2009; Avarucci et al., 2022). Since q is an integer, this means that, for any $\epsilon > 0$, there exist $n^*(\epsilon)$ and $T^*(\epsilon)$ such that $P(\hat{q} = q) > 1 - \epsilon$ for all $n > n^*(\epsilon)$ and $T > T^*(\epsilon)$. It follows that, for any $v \in \mathbb{R}$, as $n, T \rightarrow \infty$,

$$\begin{aligned} P(\hat{\chi}_{it} \leq v) &= P(\hat{\chi}_{it} \leq v; \hat{q} = q) + P(\hat{\chi}_{it} \leq v; \hat{q} \neq q) = P(\hat{\chi}_{it} \leq v; \hat{q} = q) + o(1) \\ &= P(\hat{\chi}_{it} \leq v | \hat{q} = q) P(\hat{q} = q) + o(1) = P(\hat{\chi}_{it} \leq v | \hat{q} = q) + o(1) \end{aligned}$$

(see also Bai, 2003, footnote 5). So the asymptotic distribution of $\hat{\chi}_{it}$ does not depend on q being substituted with \hat{q} ; the same holds true for all asymptotic statements in the rest of this section. Hence, without loss of generality, we hereafter can assume that q is known.

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

The first step in our estimation procedure is the computation of a lag-window estimator (47) of the spectral density matrix $\Sigma_n(\theta)$. This requires a kernel $\mathcal{K}(\cdot)$ and a bandwidth B_T on which we make the following standard assumptions.

Assumption (K).

- (a) The kernel $\mathcal{K}(u)$ is even, bounded, with support $[-1, 1]$, and
- (i) $|\mathcal{K}(u) - 1| = O(|u|^\kappa)$, as $u \rightarrow 0$, for some positive real κ ;
 - (ii) $\int_{-1}^1 \mathcal{K}^2(u) du < \infty$;
 - (iii) $\sum_{j \in \mathbb{Z}} \sup_{|s-j| \leq 1} |\mathcal{K}(jw) - \mathcal{K}(sw)| = O(1)$, as $w \rightarrow 0$;
- (b) the bandwidth B_T is such that $c_1 T^\delta \leq B_T \leq c_2 T^\delta$ for some $0 < \delta < 1$ and positive reals c_1 and c_2 .

Let $\sigma_{ij}^\chi(\theta)$ and $\hat{\sigma}_{ij}(\theta)$, $i, j = 1, \dots, n$, denote the (i, j) th entries of $\Sigma^\chi(\theta)$ and $\hat{\Sigma}(\theta)$, 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 $p > 4$, and B_T and κ satisfy Assumption (K). Then, under Assumptions (A) and (K), for any $\epsilon > 0$, there exist $\eta(\epsilon)$, $T^*(\epsilon)$, and $n^*(\epsilon)$, all independent of i and j , such that

$$\mathbb{P}\left(\max_{|h| \leq B_T} \frac{|\hat{\sigma}_{ij}(\theta_h) - \sigma_{ij}(\theta_h)|}{\max\left(\eta_{T;\kappa,p}, \frac{1}{\sqrt{n}}\right)} \geq \eta(\epsilon)\right) \leq \epsilon \quad (57)$$

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

Remark 15. The rate $\eta_{T;\kappa,p}$ in (57) depends on (i) the kernel smoothness κ , (ii) the bandwidth B_T which, by Assumption (K-b), is such that $B_T \asymp T^\delta$, and (iii) the minimum number p of finite moments we allow to exist which, for the result in Proposition 1 to hold, must be such that $p > 4$, in agreement with Assumptions (A-a-i) and (A-b-i).

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. Now, the first term is larger than the third if $\delta \geq \frac{1}{2\kappa+1}$: 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 $\delta \geq \frac{1}{3}$ or $\delta \geq \frac{1}{5}$ to get rid of the bias. Moreover, the first term in $\eta_{T;\kappa,p}$ is always larger than the second one if $\delta \leq 1 - \frac{4}{p}$.

Remark 16. To obtain the classical rate $\eta_{T;\kappa,p} = \sqrt{\frac{B_T \log T}{T}}$, we must have $p > 5$ if we choose $\kappa = 2$, while we need $p > 6$ if we choose $\kappa = 1$. Whereas, if $4 < p \leq 5$, setting $\delta \geq \frac{1}{2\kappa+1}$ yields the slightly worse rate $\eta_{T;\kappa,p} = \frac{T^{2/p} B_T (\log T)^{2+2/p}}{T}$ which is such that $\sqrt{\frac{B_T \log T}{T}} < \eta_{T;\kappa,p} < \frac{B_T (\log T)^{5/2}}{\sqrt{T}}$. In fact, recent results by Barigozzi and Farnè (2021) show that it is possible to obtain Proposition 1 with the classical $\sqrt{\frac{B_T \log T}{T}}$ convergence rate, regardless of the minimum number p of moments we assume to exist, as long as at least $p > 4$. These results, however, are based on assumptions that slightly differ from the usual GDFM setup considered in this paper. Therefore, we prefer to stick to the convergence rate presented in Proposition 1.

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 $s = 1, \dots, m$. Then, Forni et al. (2017, Proposition 9) prove the following.

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

6.3 Asymptotics for $\widehat{\mathbf{U}}_T$

Considering the spectral decomposition

$$\mathcal{R}_n \mathbf{\Gamma}^u \mathcal{R}_n' = \mathbf{P}_n^\psi \mathbf{\Lambda}_n^\psi \mathbf{P}_n^{\psi'}, \quad (58)$$

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

Assumption (E). *Let \bar{n} be such that $\frac{1}{\bar{n}} + \frac{\bar{n}}{n} \rightarrow 0$ as $n \rightarrow \infty$. Then, for any $t \in \mathbb{Z}$,*

$$\sqrt{\frac{n}{\bar{n}}} \mathbf{P}_{\bar{n}}^{\psi'} \boldsymbol{\phi}_{\bar{n}t} \rightarrow_d \mathcal{N}(\mathbf{0}_q, \mathcal{P}_t^u) \quad \text{as } n \rightarrow \infty, \quad (59)$$

where $\mathcal{P}_t^u := \lim_{n \rightarrow \infty} \frac{n}{\bar{n}} \mathbb{E}[\mathbf{P}_{\bar{n}}^{\psi'} \boldsymbol{\phi}_{\bar{n}t} \boldsymbol{\phi}_{\bar{n}t}' \mathbf{P}_{\bar{n}}^\psi]$ is positive definite.

Remark 17. Note that \mathcal{P}_t^u is not $\lim_{n \rightarrow \infty} \frac{n}{\bar{n}} \mathbf{P}_{\bar{n}}^{\psi'} \mathbb{E}[\boldsymbol{\phi}_{\bar{n}t} \boldsymbol{\phi}_{\bar{n}t}'] \mathbf{P}_{\bar{n}}^\psi$ since the eigenvectors in $\mathbf{P}_{\bar{n}}^\psi$ 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.

Consistency and asymptotic normality of $\widehat{\mathbf{u}}_t$ are proved in the next theorem.

Theorem 1. *There exists a $q \times q$ diagonal matrix $\widehat{\mathbf{W}}^z$ with diagonal entries ± 1 depending on n, T , and the observations, such that, for any $t = 1, \dots, T$ and any \bar{n} satisfying $\frac{1}{\bar{n}} + \frac{\bar{n}}{n} \rightarrow 0$ as $n, T \rightarrow \infty$,*

(i) *under Assumptions (A) through (K),*

$$\left\| \widehat{\mathbf{u}}_t - \widehat{\mathbf{W}}^z (\mathbf{\Gamma}^u)^{-1/2} \mathbf{u}_t \right\| = O_{\mathbb{P}} \left(\max \left(\frac{1}{\sqrt{\bar{n}}}, \zeta_{nT} \right) \right) \quad \text{as } n, T \rightarrow \infty;$$

(ii) *under Assumptions (A) through (E), with \bar{n} satisfying*

$$\frac{1}{\bar{n}} + \sqrt{\bar{n}} \zeta_{nT} \rightarrow 0 \quad \text{as } n, T \rightarrow \infty, \quad (60)$$

$\widehat{\mathbf{W}}^z - \mathbf{W}^u \rightarrow_{\mathbb{P}} \mathbf{0}_{q \times q}$ for some \mathbf{W}^u and, letting $\mathcal{M}^u := \text{plim}_{n \rightarrow \infty} \frac{n}{\bar{n}} \mathbf{P}_{\bar{n}}^{\psi'} \mathbf{P}_{\bar{n}}^\psi$ and $\mathcal{L}^u := \text{plim}_{n \rightarrow \infty} \frac{\mathbf{\Lambda}_n^\psi}{n}$,

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

where \mathcal{P}_t^u is defined in Assumption (E).

Remark 18. For condition (60) to hold we can assume \bar{n} to be of the form $\bar{n} = \zeta_{nT}^{-2} L^{-1}(\zeta_{nT}^{-1})$ for some slowly varying at infinity function $L(\cdot)$, i.e., such that $L(\zeta_{nT}^{-1}) \rightarrow \infty$ but $L(a\zeta_{nT}^{-1})/L(\zeta_{nT}^{-1}) \rightarrow 1$ for any real a , as $n, T \rightarrow \infty$. This implies that the rate in part (ii) is only marginally slower than the consistency rate ζ_{nT}^{-1} obtained in Forni et al. (2017). In fact, inspection of the proof of part (i) reveals that consistency holds with a faster rate and we could relax (60) to $\frac{\bar{n}}{\sqrt{n}} \zeta_{nT} \rightarrow 0$. However, since we need (60) anyway in Section 6.5 for deriving the asymptotic properties of the common component, we stick to it also in Theorem 1.

Remark 19. Note that \bar{n} then depends on both n and T . Let us assume, for simplicity, that $p > 6$. Then, in light of Remark 16, we have $\zeta_{nT} = \max\left(\sqrt{\frac{B_T \log T}{T}}, \frac{1}{\sqrt{n}}\right)$. So, in view of the previous remark, we can achieve a convergence rate almost equal to $\sqrt{\bar{n}}$ if $T/(B_T n) \downarrow 0$, which is the rate obtained by Bai (2003) in the static factor model. If $(nB_T)/T \downarrow 0$, we can achieve a convergence rate almost equal to $\sqrt{T/B_T}$, which is slower than the rates $\sqrt{\bar{n}}$ or T (depending on whether $\sqrt{\bar{n}}/T \downarrow 0$ or $T/\sqrt{\bar{n}} \downarrow 0$) obtained in Bai (2003).

Remark 20. A consistent estimator of the asymptotic covariance matrix of $\sqrt{\bar{n}}(\hat{\mathbf{u}}_t - \widehat{\mathbf{W}}^z (\boldsymbol{\Gamma}^u)^{-1/2} \mathbf{u}_t)$ is

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

where $\widehat{\boldsymbol{\Lambda}}_n^z$ is the $q \times q$ diagonal matrix of largest eigenvalues of $\widehat{\boldsymbol{\Gamma}}_n^z$ defined in (48), $\widehat{\mathbf{P}}_n^z$ is $n \times q$ matrix of the corresponding normalized eigenvectors, and $\widehat{\mathcal{P}}_t^u$ is a consistent estimator of \mathcal{P}_t^u . This requires specific assumptions on the form of the cross-sectional dependence of $\{\phi_{it}\}$. For instance, when the ϕ_{it} s are cross-sectionally independent, then the approach of Section 5(a) in Bai (2003) can be adapted, providing¹⁶

$$\widehat{\mathcal{P}}_t^u = \frac{n}{\bar{n}^2} \sum_{i=1}^{\bar{n}} \widehat{\mathbf{p}}_i^z \widehat{\mathbf{p}}_i^{z'} \left\{ \frac{1}{T} \sum_{t=1}^T \widehat{\phi}_{it}^2 \right\} \quad (61)$$

where $\widehat{\phi}_{it} = \widehat{z}_{it} - \check{\mathbf{R}}_i' \widehat{\mathbf{u}}_t$ and $\widehat{\mathbf{p}}_i^{z'}$ is the i th row of $\widehat{\mathbf{P}}_n^z$. If we want to address cross-sectional dependence we can instead follow Bai and Ng (2006, Section 3), and consider, for example, the Cross-Sectional HAC estimator defined therein.

6.4 Asymptotics for $\widehat{\mathcal{R}}_n$

Thanks to the duality between (26) and (27), the asymptotics for $\widehat{\mathcal{R}}_n$ follow along the same lines as for $\widehat{\mathcal{U}}_T$. Consider the spectral decomposition

$$\mathbf{u}_T \boldsymbol{\Sigma}^R \mathbf{u}_T' = \boldsymbol{\Pi}_T^\psi \mathbf{L}_T^\psi \boldsymbol{\Pi}_T^{\psi'}, \quad (62)$$

where \mathbf{L}_T^ψ is the $q \times q$ diagonal matrix of $\mathbf{u}_T \boldsymbol{\Sigma}^R \mathbf{u}_T'$'s eigenvalues and $\boldsymbol{\Pi}_T^\psi$ the $T \times q$ with columns the corresponding orthonormal eigenvectors. Similar to (59), we make the following assumption.

¹⁶Although $\widehat{\mathcal{P}}_t^u$ 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.

Assumption (F). Let \bar{T} be such that $\frac{1}{T} + \frac{\bar{T}}{T} \rightarrow 0$ as $T \rightarrow \infty$. Then, for any $i \in \mathbb{N}$,

$$\sqrt{\frac{\bar{T}}{T}} \mathbf{\Pi}_{\bar{T}}^{\psi'} \boldsymbol{\varphi}_{\bar{T}}^i \rightarrow_d \mathcal{N}(\mathbf{0}_q, \mathcal{P}^R) \quad \text{as } T \rightarrow \infty, \quad (63)$$

where $\mathcal{P}_i^R := \lim_{T \rightarrow \infty} \frac{\bar{T}}{T} \mathbb{E}[\mathbf{\Pi}_{\bar{T}}^{\psi'} \boldsymbol{\varphi}_{\bar{T}}^i \boldsymbol{\varphi}_{\bar{T}}^{i'} \mathbf{\Pi}_{\bar{T}}^{\psi}]$ is positive definite.

Remark 21. Here again, notice that \mathcal{P}_i^R is not $\lim_{T \rightarrow \infty} (T/\bar{T}) \mathbf{\Pi}_{\bar{T}}^{\psi'} \mathbb{E}[\boldsymbol{\varphi}_{\bar{T}}^i \boldsymbol{\varphi}_{\bar{T}}^{i'}] \mathbf{\Pi}_{\bar{T}}^{\psi}$ since the eigenvectors in $\mathbf{\Pi}_{\bar{T}}^{\psi}$ are random; so we must assume its existence. If eigenvectors were not random, its existence would follow from Lemma 18 in the Appendix, for all $T \in \mathbb{N}$. Moreover, \mathcal{P}_i^R is positive definite since it is a Toeplitz matrix containing all the autocovariances of the i th idiosyncratic component. A similar assumption is made also in Bai (2003, Assumption F4); it is satisfied, for example, by all α -mixing processes.

The next theorem, which can be proved along the same lines as Theorem 1, then establishes the consistency and asymptotic normality of $\widehat{\mathbf{R}}_i$.

Theorem 2. There exists a $q \times q$ diagonal matrix $\widehat{\mathbf{W}}^z$ with diagonal entries ± 1 depending on n , T , and the observations, such that, for any $i = 1, \dots, n$ and any \bar{T} satisfying $\frac{1}{T} + \frac{\bar{T}}{T} \rightarrow 0$, as $n, T \rightarrow \infty$,

(i) under Assumptions (A) through (K),

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

(ii) under Assumptions (A) through (K) and (F), with \bar{T} such that

$$\frac{1}{\bar{T}} + \sqrt{\bar{T}} \zeta_{nT} \rightarrow 0, \quad \text{as } n, T \rightarrow \infty, \quad (64)$$

$\widehat{\mathbf{W}}^z - \mathbf{W}^R \rightarrow_{\mathbb{P}} \mathbf{0}_{q \times q}$ for some \mathbf{W}^R and, letting $\mathcal{M}_R := \text{plim}_{T \rightarrow \infty} \frac{T}{\bar{T}} \mathbf{\Pi}_{\bar{T}}^{\psi'} \mathbf{\Pi}_{\bar{T}}^{\psi}$, and $\mathcal{L}_R := \text{plim}_{T \rightarrow \infty} \frac{L^{\psi}}{T}$,

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

where \mathcal{P}_i^R is defined in Assumption (F).

Remark 22. For condition (64) to hold, we can assume \bar{T} to be of the form $\bar{T} = \zeta_{nT}^{-2} L^{-1}(\zeta_{nT}^{-1})$ for some slowly varying at infinity function $L(\cdot)$, i.e., such that $L(\zeta_{nT}^{-1}) \rightarrow \infty$ but $L(a\zeta_{nT}^{-1})/L(\zeta_{nT}^{-1}) \rightarrow 1$ for any real a , as $n, T \rightarrow \infty$. This implies, in part (ii), a marginally slower rate than ζ_{nT}^{-1} , which is the consistency rate obtained in Forni et al. (2017). In fact, inspection of the proof of part (i) reveals that consistency holds with a faster rate, hence that we could relax (64) to $\frac{\bar{T}}{\sqrt{\bar{T}}} \zeta_{nT} \rightarrow 0$. However, since we need (64) anyway for deriving the properties of the common component in Section 6.5, we stick with it also in Theorem 2.

Remark 23. Note that \bar{T} here depends on both n and T . Let us assume, for simplicity, that $p > 6$. In light of Remark 16, we then have $\zeta_{nT} = \max \left(\sqrt{\frac{B_T \log T}{T}}, \frac{1}{\sqrt{n}} \right)$. So, if $T/(B_T n) \downarrow 0$, we can achieve a convergence rate almost equal to \sqrt{n} , while if $(nB_T)/T \downarrow 0$, we can achieve a convergence rate almost

equal to $\sqrt{T/B_T}$. Both rates are slower than the rates \sqrt{T} or n (depending on whether $\sqrt{T}/n \downarrow 0$ or $n/\sqrt{T} \downarrow 0$) in Bai (2003). This is because we need to estimate a spectral density before running PCA. In particular, if $(nB_T)/T \downarrow 0$, the best rates we can achieve are $T^{2/5}$ if we choose a quadratic kernel, i.e. $\kappa = 2$, with optimal bandwidth $B_T \asymp T^{1/5}$, or $T^{1/3}$ if we choose a Bartlett kernel, i.e. $\kappa = 1$, with optimal bandwidth $B_T \asymp T^{1/3}$.

Remark 24. A consistent estimator of the asymptotic covariance matrix of $\sqrt{T}(\widehat{\mathbf{R}}_i - \widehat{\mathbf{W}}_T^z(\boldsymbol{\Sigma}^R)^{-1/2}\mathbf{R}_i)$ is

$$\left(\frac{\widehat{\mathbf{L}}_T^z}{T}\right)^{-1/2} \left(\frac{T}{T}\widehat{\boldsymbol{\Pi}}_T^{z'}\widehat{\boldsymbol{\Pi}}_T^z\right)^{-1} \widehat{\mathcal{P}}_i^R \left(\frac{T}{T}\widehat{\boldsymbol{\Pi}}_T^{z'}\widehat{\boldsymbol{\Pi}}_T^z\right)^{-1} \left(\frac{\widehat{\mathbf{L}}_T^z}{T}\right)^{-1/2},$$

where $\widehat{\mathbf{L}}_T^z$ is the $q \times q$ diagonal matrix of largest eigenvalues of $\widehat{\mathbf{G}}_T^z$ defined in (52), $\widehat{\boldsymbol{\Pi}}_T^z$ the $n \times q$ matrix of the corresponding normalized eigenvectors, and $\widehat{\mathcal{P}}_i^R$ is a consistent estimator of \mathcal{P}_i^R . If we assume that $\{\phi_{it}\}$ is not autocorrelated, we can use

$$\widehat{\mathcal{P}}_i^R := \frac{T}{T^2} \sum_{t=1}^{\bar{T}} \widehat{\pi}_t^z \widehat{\pi}_t^{z'} \widehat{\phi}_{it}^2, \quad (65)$$

where $\widehat{\phi}_{it} = \widehat{z}_{it} - \check{\mathbf{u}}_t' \widehat{\mathbf{R}}_i$ and $\widehat{\pi}_t^{z'}$ is the t th row of $\widehat{\boldsymbol{\Pi}}_T^z$. To address idiosyncratic autocorrelation, a natural choice is the usual HAC estimator used also in Bai (2003, Section 5(b)).

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

Using the estimators developed in Section 5.2 for the loadings \mathbf{R}_i and the common shocks \mathbf{u}_t , 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} . However, due to the presence of the product of the rotation matrices $\widehat{\mathbf{W}}^z(\boldsymbol{\Gamma}^u)^{-1/2}$ and $\widehat{\mathbf{W}}^z(\boldsymbol{\Sigma}^R)^{-1/2}$, identification with $\widehat{\mathbf{R}}_i' \widehat{\mathbf{u}}_t$, is not warranted. Indeed, the rotation $\widehat{\mathbf{W}}^z(\boldsymbol{\Gamma}^u)^{-1/2}$ arises when estimating \mathbf{u}_t , while the rotation $\widehat{\mathbf{W}}^z(\boldsymbol{\Sigma}^R)^{-1/2}$ when estimating \mathbf{R}_i , and they are not necessarily identical. In contrast, as shown in Appendix D, it holds that

$$\left\| \check{\mathbf{R}}_i' - \mathbf{R}_i' (\boldsymbol{\Gamma}^u)^{1/2} \widehat{\mathbf{W}}^z \right\| = O_P(\zeta_{nT}) \quad \text{and} \quad \left\| \check{\mathbf{u}}_t' - \mathbf{u}_t' (\boldsymbol{\Sigma}^R)^{1/2} \widehat{\mathbf{W}}^z \right\| = O_P(\zeta_{nT}),$$

which, together with Theorems 1 and 2, imply that $\check{\mathbf{R}}_i' \widehat{\mathbf{u}}_t$ and $\widehat{\mathbf{R}}_i' \check{\mathbf{u}}_t$ are rotation-free as they only involve either $(\widehat{\mathbf{W}}^z)^2 = \mathbf{I}_q$ or $(\widehat{\mathbf{W}}^z)^2 = \mathbf{I}_q$. This is why the estimator $\widehat{\psi}_{it}$ we are proposing for ψ_{it} are of the form (55). 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 $\bar{n} = \bar{T}$, they avoid the technical difficulty of combining estimators with different rates of convergence (see Bai, 2003, proof of Theorem 3).

In order to establish the asymptotic properties of $\widehat{\psi}_{it}$, we need to slightly strengthen Assumptions (E) and (F) as follows.

Assumption (G). Set $\bar{n} = \bar{T} =: \bar{h}$ such that $\frac{1}{\bar{h}} + \frac{\bar{h}}{\min(n, T)} \rightarrow 0$ as $n, T \rightarrow \infty$. Then, for any $t \in \mathbb{Z}$ and any $i \in \mathbb{N}$,

$$\sqrt{\frac{\bar{n}}{\bar{h}}} \mathbf{P}_{\bar{h}}^{\psi'} \phi_{\bar{h}t} + \sqrt{\frac{\bar{T}}{\bar{h}}} \boldsymbol{\Pi}_{\bar{h}}^{\psi'} \boldsymbol{\varphi}_{\bar{h}}^i \rightarrow_d \mathcal{N}(\mathbf{0}_q, \mathcal{P}_t^u + \mathcal{P}_i^R + \boldsymbol{\Omega}_{it} + \boldsymbol{\Omega}'_{it}) \quad \text{as } n, T \rightarrow \infty,$$

where \mathcal{P}_i^u and \mathcal{P}_i^R are defined in Assumptions (E) and (F), respectively,

$$\mathbf{\Omega}_{it} := \lim_{n,T \rightarrow \infty} \left(\frac{\sqrt{nT}}{\bar{h}} \right) \mathbb{E} \left[\mathbf{P}_{\bar{n}}^{\psi'} \phi_{\bar{n}t} \varphi_{\bar{T}}^{i'} \mathbf{\Pi}_{\bar{T}}^{\psi} \right],$$

and the $q \times q$ asymptotic covariance matrix $(\mathcal{P}_t^u + \mathcal{P}_i^R + \mathbf{\Omega}_{it} + \mathbf{\Omega}_{it}')$ is positive definite.

The next theorem establishes the consistency and asymptotic normality of $\hat{\psi}_{it}$.

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} such that $\frac{1}{\bar{h}} + \frac{\bar{h}}{\min(n, T)} \rightarrow 0$ as $n, T \rightarrow \infty$,

(i) under Assumptions (A) through (K),

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

(ii) if also Assumption (G) holds, and if \bar{h} is such that

$$\frac{1}{\bar{h}} + \sqrt{\bar{h}} \zeta_{nT} \rightarrow 0, \quad \text{as } n, T \rightarrow \infty, \quad (66)$$

and $\widehat{\mathbf{W}}^z \rightarrow_{\mathbb{P}} \mathbf{W}^u$, $\widehat{\mathbf{W}}^z \rightarrow_{\mathbb{P}} \mathbf{W}^R$, then

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

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

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

with \mathcal{P}_t^u , $\boldsymbol{\mathcal{M}}^u$, and $\boldsymbol{\mathcal{L}}^u$ as defined in Theorem 1, \mathcal{P}_i^R , $\boldsymbol{\mathcal{M}}^R$, and $\boldsymbol{\mathcal{L}}^R$ as defined in Theorem 2, and $\mathbf{\Omega}_{it}$ as defined in Assumption (G).

Remark 25. The challenge that we need to resolve is that \mathbf{Z}_{nT} is not observed but rather estimated by $\widehat{\mathbf{Z}}'_{nT} := \widehat{\mathbf{A}}'_n(L) \mathbf{X}'_{nT}$, for otherwise one could apply Bai (2003) directly to derive the limiting properties of the static common component $\boldsymbol{\Psi}_{n,T}$. This would necessarily involve a certain loss in the rate of convergence of the estimator with respect to the $\min(\sqrt{n}, \sqrt{T})$ rate obtained by Bai (2003), and thus a larger sampling variability (for given n and T) of our estimator.

For instance, assuming $p > 6$, and considering a very smooth kernel ($\kappa = \infty$), so that we can pick B_T as constant independent of T , we would have $\zeta_{n,T} = \max \left(\frac{1}{\sqrt{T}}, \frac{1}{\sqrt{n}} \right)$.¹⁷ Then, in light of Remarks 18 and (22), we can choose $\bar{h} = \min(n, T) L^{-1}(\zeta_{n,T}^{-1})$ for some slowly varying at infinity function $L(\cdot)$, thus

¹⁷Notice that the $\log T$ factor in $\zeta_{n,T}$ comes from the estimation of the spectral density using B_T autocovariances and the fact that $B_T = O(T^\delta)$ in Assumption (K). Thus, when B_T is treated as a constant, this $\log T$ will disappear from the expression for $\zeta_{n,T}$.

ensuring $\frac{1}{h} + \frac{\bar{h}}{\min(n, T)} \rightarrow 0$ as $n, T \rightarrow \infty$. This means that, up to the slowly varying function $L(\cdot)$, we can achieve the same convergence rate as the one given in Bai (2003) for the static factor model case.

Remark 26. Notice that, in agreement with (38), we always can write (see the proof of Theorem 1)

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

which implies $\mathbf{R}'_i = \mathbf{p}_i^{\psi'} (\boldsymbol{\Lambda}_n^\psi)^{1/2} (\boldsymbol{\Gamma}^u)^{-1/2}$ (here $\mathbf{p}_i^{\psi'}$ is the i th row of \mathbf{P}_n^ψ). Moreover, $\check{\mathbf{R}}'_i = \widehat{\mathbf{p}}_i^{z'} (\widehat{\boldsymbol{\Lambda}}_n^z)^{1/2}$ by definition, where $\widehat{\mathbf{p}}_i^{z'} (\widehat{\boldsymbol{\Lambda}}_n^z)^{1/2}$ (see the proof of Theorem 3) is a consistent estimator of $\mathbf{R}'_i (\boldsymbol{\Gamma}^u)^{1/2} \mathbf{W}^u$. Therefore, a natural estimator of V_{it}^u is

$$\begin{aligned} \widehat{V}_{it}^u &:= \check{\mathbf{R}}'_i \left(\frac{\widehat{\boldsymbol{\Lambda}}_n^z}{n} \right)^{-1/2} \left(\frac{n}{\bar{n}} \widehat{\mathbf{P}}_n^{z'} \widehat{\mathbf{P}}_n^z \right)^{-1} \widehat{\mathcal{P}}_t^u \left(\frac{n}{\bar{n}} \widehat{\mathbf{P}}_n^{z'} \widehat{\mathbf{P}}_n^z \right)^{-1} \left(\frac{\widehat{\boldsymbol{\Lambda}}_n^z}{n} \right)^{-1/2} \check{\mathbf{R}}_i \\ &= n \widehat{\mathbf{p}}_i^{z'} \left(\frac{n}{\bar{n}} \widehat{\mathbf{P}}_n^{z'} \widehat{\mathbf{P}}_n^z \right)^{-1} \widehat{\mathcal{P}}_t^u \left(\frac{n}{\bar{n}} \widehat{\mathbf{P}}_n^{z'} \widehat{\mathbf{P}}_n^z \right)^{-1} \widehat{\mathbf{p}}_i^z, \end{aligned} \quad (67)$$

where $\widehat{\mathcal{P}}_t^u$ is an estimator of \mathcal{P}_t^u as, for example, the one defined in (61). The estimator in (67) is rotation-free as it neither depends on the unknown matrix $\boldsymbol{\Gamma}^u$ nor on the sign matrix \mathbf{W}^u . Similarly, we always can write

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

which implies $\mathbf{u}'_t = \boldsymbol{\pi}_t^{\psi'} (\mathbf{L}_T^\psi)^{1/2} (\boldsymbol{\Sigma}^R)^{-1/2}$ (here $\boldsymbol{\pi}_t^{\psi'}$ is the t th row of $\boldsymbol{\Pi}_T^\psi$). Moreover, by definition, we have that $\check{\mathbf{u}}'_t = \widehat{\boldsymbol{\pi}}_t^{z'} (\widehat{\mathbf{L}}_T^z)^{1/2}$ where (see the proof of Theorem 3) $\widehat{\boldsymbol{\pi}}_t^{z'} (\widehat{\mathbf{L}}_T^z)^{1/2}$ is a consistent estimator of $\mathbf{u}'_t (\boldsymbol{\Sigma}^R)^{1/2} \mathbf{W}^R$. Therefore, a natural estimator of V_{it}^R is

$$\begin{aligned} \widehat{V}_{it}^R &:= \check{\mathbf{u}}'_t \left(\frac{\widehat{\mathbf{L}}_T^z}{T} \right)^{-1/2} \left(\frac{T}{\bar{T}} \widehat{\boldsymbol{\Pi}}_T^{z'} \widehat{\boldsymbol{\Pi}}_T^z \right)^{-1} \widehat{\mathcal{P}}_i^R \left(\frac{T}{\bar{T}} \widehat{\boldsymbol{\Pi}}_T^{z'} \widehat{\boldsymbol{\Pi}}_T^z \right)^{-1} \left(\frac{\widehat{\mathbf{L}}_T^z}{T} \right)^{-1/2} \check{\mathbf{u}}_t \\ &= T \widehat{\boldsymbol{\pi}}_t^{z'} \left(\frac{T}{\bar{T}} \widehat{\boldsymbol{\Pi}}_T^{z'} \widehat{\boldsymbol{\Pi}}_T^z \right)^{-1} \widehat{\mathcal{P}}_i^R \left(\frac{T}{\bar{T}} \widehat{\boldsymbol{\Pi}}_T^{z'} \widehat{\boldsymbol{\Pi}}_T^z \right)^{-1} \widehat{\boldsymbol{\pi}}_t^z, \end{aligned} \quad (68)$$

where $\widehat{\mathcal{P}}_i^R$ is an estimator of \mathcal{P}_i^R as, for example, the one defined in (65). The estimator in (68) is also rotation-free since it neither depends on the unknown matrix $\boldsymbol{\Sigma}^R$ nor on the sign matrix \mathbf{W}^R . Finally, an estimator of C_{it} can be computed along the same lines.

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

For the estimator of the common component χ_{it} , defined in (56), we have the following result.

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} such that $\frac{1}{h} + \frac{\bar{h}}{\min(n, T)} \rightarrow 0$, as $n, T \rightarrow \infty$,

(i) under Assumptions (A) through (K),

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

(ii) if also Assumption (G) holds, if $\widehat{\mathbf{W}}^z \rightarrow_{\mathbb{P}} \mathbf{W}^u$, $\widehat{\mathbf{W}}^z \rightarrow_{\mathbb{P}} \mathbf{W}^R$ and \bar{h} is such that

$$\frac{1}{\bar{h}} + \sqrt{\bar{h}} \zeta_{nT} \rightarrow 0, \quad \text{as } n, T \rightarrow \infty, \quad (69)$$

then

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

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

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

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,

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

$$\begin{aligned} \boldsymbol{\nu}_{t \dots t-K}^u &:= \left\{ \mathbf{I}_{K+1} \otimes \left[\mathbf{W}^u (\boldsymbol{\mathcal{L}}^u)^{-1/2} (\boldsymbol{\mathcal{M}}^u)^{-1} \right] \right\} \boldsymbol{\mathcal{P}}_{t \dots t-K}^u \left\{ \mathbf{I}_{K+1} \otimes \left[(\boldsymbol{\mathcal{M}}^u)^{-1} (\boldsymbol{\mathcal{L}}^u)^{-1/2} \mathbf{W}^u \right] \right\}, \\ \boldsymbol{\nu}_{i_1 \dots i_{q+1}}^R &:= \left\{ \mathbf{I}_{q+1} \otimes \left[\mathbf{W}^R (\boldsymbol{\mathcal{L}}^R)^{-1/2} (\boldsymbol{\mathcal{M}}^R)^{-1} \right] \right\} \boldsymbol{\mathcal{P}}_{i_1 \dots i_{q+1}}^R \left\{ \mathbf{I}_{q+1} \otimes \left[(\boldsymbol{\mathcal{M}}^R)^{-1} (\boldsymbol{\mathcal{L}}^R)^{-1/2} \mathbf{W}^R \right] \right\}, \\ \boldsymbol{\Omega}_{i_1 \dots i_{q+1}} \left\{ \mathbf{I}_{K+1} \otimes \left[\mathbf{W}^u (\boldsymbol{\mathcal{L}}^u)^{-1/2} (\boldsymbol{\mathcal{M}}^u)^{-1} \right] \right\} &:= \boldsymbol{\Omega}_{i_1 \dots i_{q+1}} \left\{ \mathbf{I}_{q+1} \otimes \left[(\boldsymbol{\mathcal{M}}^R)^{-1} (\boldsymbol{\mathcal{L}}^R)^{-1/2} \mathbf{W}^R \right] \right\}, \end{aligned}$$

with \mathcal{M}^u and \mathcal{L}^u as defined in Theorem 1, \mathcal{M}^R and \mathcal{L}^R as defined in Theorem 2, and

$$\begin{aligned}\mathcal{P}_{t\dots t-K}^u &:= \lim_{n \rightarrow \infty} \frac{n}{\bar{h}} \mathbb{E} \left[\left\{ \mathbf{I}_{K+1} \otimes \mathbf{P}_{\bar{h}}^{\psi'} \right\} \begin{pmatrix} \phi_{\bar{h}t} \\ \vdots \\ \phi_{\bar{h}t-K} \end{pmatrix} \begin{pmatrix} \phi_{\bar{h}t} \\ \vdots \\ \phi_{\bar{h}t-K} \end{pmatrix}' \left\{ \mathbf{I}_{K+1} \otimes \mathbf{P}_{\bar{h}}^{\psi'} \right\}' \right], \\ \mathcal{P}_{i_1 \dots i_{q+1}}^R &:= \lim_{T \rightarrow \infty} \frac{T}{\bar{h}} \mathbb{E} \left[\left\{ \mathbf{I}_{q+1} \otimes \mathbf{\Pi}_{\bar{h}}^{\psi'} \right\} \begin{pmatrix} \varphi_{\bar{h}}^{i_1} \\ \vdots \\ \varphi_{\bar{h}}^{i_{q+1}} \end{pmatrix} \begin{pmatrix} \varphi_{\bar{h}}^{i_1} \\ \vdots \\ \varphi_{\bar{h}}^{i_{q+1}} \end{pmatrix}' \left\{ \mathbf{I}_{q+1} \otimes \mathbf{\Pi}_{\bar{h}}^{\psi'} \right\}' \right], \\ \Omega_{i_1 \dots i_{q+1}} &:= \lim_{n, T \rightarrow \infty} \frac{\sqrt{nT}}{\bar{h}} \mathbb{E} \left[\left\{ \mathbf{I}_{K+1} \otimes \mathbf{P}_{\bar{h}}^{\psi'} \right\} \begin{pmatrix} \phi_{\bar{h}t} \\ \vdots \\ \phi_{\bar{h}t-K} \end{pmatrix} \begin{pmatrix} \varphi_{\bar{h}}^{i_1} \\ \vdots \\ \varphi_{\bar{h}}^{i_{q+1}} \end{pmatrix}' \left\{ \mathbf{I}_{q+1} \otimes \mathbf{\Pi}_{\bar{h}}^{\psi'} \right\}' \right].\end{aligned}$$

Remark 27. To appreciate the formulas given in Theorem 4, let us consider a simple illustrative 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

$$\sqrt{\bar{h}}(\widehat{\chi}_{1t} - \chi_{1t}) = \sqrt{\bar{h}} \sum_{k=0}^1 \sum_{j_s=1}^2 \left\{ c_{1,j_s,k} \mathbf{R}'_{j_s} (\widehat{\mathbf{u}}_{t-k} - \mathbf{u}_{t-k}) \right\} + o_P(1), \quad (70)$$

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

$$\begin{aligned}W_{it}^u &= \lim_{n, T \rightarrow \infty} \bar{h} \left(c_{1,1,0}^2 R_1^2 \text{Var}(\widehat{u}_t - u_t) + c_{1,2,0}^2 R_2^2 \text{Var}(\widehat{u}_t - u_t) \right. \\ &\quad + c_{1,1,1}^2 R_1^2 \text{Var}(\widehat{u}_{t-1} - u_{t-1}) + c_{1,2,1}^2 R_2^2 \text{Var}(\widehat{u}_{t-1} - u_{t-1}) \\ &\quad + 2c_{1,1,0}c_{1,2,0}R_1R_2 \text{Cov}(\widehat{u}_t - u_t) + 2c_{1,1,1}c_{1,2,1}R_1R_2 \text{Cov}(\widehat{u}_{t-1} - u_{t-1}) \\ &\quad + 2c_{1,1,0}c_{1,1,1}R_1^2 \text{Cov}((\widehat{u}_t - u_t), (\widehat{u}_{t-1} - u_{t-1})) \\ &\quad + 2c_{1,2,0}c_{1,2,1}R_2^2 \text{Cov}((\widehat{u}_t - u_t), (\widehat{u}_{t-1} - u_{t-1})) \\ &\quad + 2c_{1,1,0}c_{1,2,1}R_1R_2 \text{Cov}((\widehat{u}_t - u_t), (\widehat{u}_{t-1} - u_{t-1})) \\ &\quad \left. + 2c_{1,2,0}c_{1,1,1}R_1R_2 \text{Cov}((\widehat{u}_t - u_t), (\widehat{u}_{t-1} - u_{t-1})) \right). \quad (71)\end{aligned}$$

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

$$\sqrt{\bar{h}}(\widehat{\chi}_{1t} - \chi_{1t}) = \sqrt{\bar{h}} \sum_{k=0}^1 \sum_{j_s=1}^2 \left\{ c_{1,j_s,k} \mathbf{u}'_{t-k} (\widehat{\mathbf{R}}_{j_s} - \mathbf{R}_{j_s}) \right\} + o_P(1), \quad (72)$$

which has asymptotic variance

$$\begin{aligned}
W_{it}^R &= \lim_{n, T \rightarrow \infty} \bar{h} \left(c_{1,1,0}^2 u_t^2 \text{Var}(\widehat{R}_1 - R_1) + c_{1,2,0}^2 u_t^2 \text{Var}(\widehat{R}_2 - R_2) \right. \\
&\quad + c_{1,1,1}^2 u_{t-1}^2 \text{Var}(\widehat{R}_1 - R_1) + c_{1,2,1}^2 u_{t-1}^2 \text{Var}(\widehat{R}_2 - R_2) \\
&\quad + 2c_{1,1,0}c_{1,1,1}u_t u_{t-1} \text{Var}(\widehat{R}_1 - R_1) + 2c_{1,2,0}c_{1,2,1}u_t u_{t-1} \text{Var}(\widehat{R}_2 - R_2) \\
&\quad + 2c_{1,1,0}c_{1,2,0}u_t^2 \text{Cov}((\widehat{R}_1 - R_1), (\widehat{R}_2 - R_2)) \\
&\quad + 2c_{1,1,1}c_{1,2,1}u_{t-1}^2 \text{Cov}((\widehat{R}_1 - R_1), (\widehat{R}_2 - R_2)) \\
&\quad + 2c_{1,1,0}c_{1,2,1}u_t u_{t-1} \text{Cov}((\widehat{R}_1 - R_1), (\widehat{R}_2 - R_2)) \\
&\quad \left. + 2c_{1,2,0}c_{1,1,1}u_t u_{t-1} \text{Cov}((\widehat{R}_1 - R_1), (\widehat{R}_2 - R_2)) \right). \tag{73}
\end{aligned}$$

The variances in (71) and (73) 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 (70) and those in (72), which contribute to the term G_{it} in the expression of the asymptotic variance.

7 Monte Carlo Simulations

The main goal of this section is to check whether the asymptotic distributions derived in Theorem 4 are empirically confirmed. We set $q = \{1, 2\}$ and we consider the data-generating process (a slightly modified version of the one used by Forni et al., 2017)

$$x_{it} = \sum_{j=1}^q \frac{a_{ij}}{(1 - \alpha_{ij}L)} u_{jt} + \xi_{it}. \tag{74}$$

We generate u_{jt} and ξ_{it} either as i.i.d. standard normal variables, or as Student-t with 5 degrees of freedom, a_{ij} as i.i.d. normal random variables with mean and variance both equal to one, and α_{ij} as i.i.d. random variables uniformly distributed over $[0.1, 0.8]$. Finally, each idiosyncratic component ξ_{it} is rescaled so that the share of variance of x_{it} it is accounting for is $\theta/(1 + \theta)$, with $\theta = 0.5$ for all i .

We simulate panels of size $n = T \in \{120, 240, 480\}$ and 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 $\chi_{it}^{(b)}$ as in (56) with $K = 20$ lags for the MA representation (notice that according to (74) the largest AR coefficient we allow for is 0.8, and, since $0.8^{20} \simeq 0.01$, the truncation error is likely to be negligible). To estimate the spectral density we use a Bartlett kernel with bandwidth $B_T = \lfloor T^{1/3} \rfloor$. The size of the blocks of the singular VAR representation is set at $q + 1$ and for each of them we fit a VAR(1) model. Since in all considered cases n is an integer multiple of $(q + 1)$, the number of VARs to be fitted is always $m = n/(q + 1)$. Because we simulate panels with $n = T$, 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 $\omega_{nT} = 1$. We also set $\bar{n} = n$ and $\bar{T} = T$: indeed, in light of Remarks 18 and 22, for any fixed n and T , we can always set \bar{n} and \bar{T} arbitrarily close to n and T , respectively. In fact, \bar{n} and \bar{T} differ from n and T by an ‘‘arbitrarily slow’’ slowly varying function, which is in practice arbitrarily close to, hence indistinguishable from, one. Finally, the asymptotic variance of $\widehat{\chi}_{it}^{(b)}$ is computed as described in Theorem 4 when using the estimators in (67) and (68).

Table 1: ESTIMATION OF THE NUMBER OF FACTORS WITH THE HALLIN-LIŠKA CRITERION

q	T	n	$u_{jt}, \xi_{it} \sim \mathcal{N}(0, 1)$	$u_{jt}, \xi_{it} \sim t_5$
			% of times $\hat{q} = q$	% of times $\hat{q} = q$
1	120	120	97.1	98.2
1	240	240	100	100
1	480	480	100	100
2	120	120	98.2	98.7
2	240	240	100	100
2	480	480	100	100

In the sequel, we treat q as known because although the effect of the estimation of q for the estimation of the common component is in principle an important issue, in practice, this issue poses no problem, at least when the Hallin and Liška (2007) method is adopted. Indeed, the results in Table 1 show that we almost always recover the true value of q when we apply the Hallin and Liška (2007) criterion to our simulated data. A similar approach is adopted in other simulation studies about factor models, see, e.g., Bai (2003) and Forni et al. (2017).

Table 2 reports the average Standardized Mean Squared Error (S-MSE) and the average Standardized Mean Absolute Error (S-MAE)

$$\text{S-MSE} := \frac{1}{B} \sum_{b=1}^B \frac{\sum_{i=1}^n \sum_{t=1}^T (\hat{\chi}_{it}^{(b)} - \chi_{it}^{(b)})^2}{\sum_{i=1}^n \sum_{t=1}^T (\chi_{it}^{(b)})^2}, \quad \text{S-MAE} := \frac{1}{B} \sum_{b=1}^B \frac{\sum_{i=1}^n \sum_{t=1}^T |\hat{\chi}_{it}^{(b)} - \chi_{it}^{(b)}|}{\sum_{i=1}^n \sum_{t=1}^T |\chi_{it}^{(b)}|} \quad (75)$$

of the estimator of χ_{it} . The results clearly show that the estimator in (56) works very well. As n and T increase, the S-MSE and S-MAE decrease monotonically, to the point that, for $n = T = 480$, the S-MSE (S-MAE) is more than 70% (65%) lower than for $n = T = 120$. This holds even when $q = 2$ and, in particular, when the shocks are generated from a heavy-tailed distribution.

Having looked at the properties of the estimator of the common components, we look into the estimator of the common shocks \mathcal{U}_T . Since the common shocks are not fully identified, we report the values of the multivariate R^2 coefficient

$$\text{TR}_{\mathcal{U}}^{(b)} := \frac{\text{tr} \left((\mathcal{U}_T^{(b)'} \hat{\mathcal{U}}_T^{(b)}) (\hat{\mathcal{U}}_T^{(b)'} \hat{\mathcal{U}}_T^{(b)})^{-1} (\hat{\mathcal{U}}_T^{(b)'} \mathcal{U}_T^{(b)}) \right)}{\text{tr} \left(\mathcal{U}_T^{(b)'} \mathcal{U}_T^{(b)} \right)}.$$

These trace statistics are always positive and smaller than one, and they tend to one when the space spanned by the true and estimated quantities are closer, i.e., when the empirical canonical correlations between the two tend to one. The results in Table 3 clearly demonstrate that our estimator does a good job at estimating the common shocks.

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

$$Z_{it}^{(b)} := \left(\frac{1}{4} \widehat{W}_{it}^u + \frac{1}{4} \widehat{W}_{it}^R \right)^{-1/2} \left(\hat{\chi}_{it}^{(b)} - \chi_{it}^{(b)} \right) \quad (76)$$

which, according to Theorem 4, is asymptotically standard normal. Figure 1 shows histograms of $\{Z_{it}^{(b)} : i = 1, \dots, n, t = 1, \dots, T, b = 1, \dots, B\}$. These histograms indicate that, while struggling

Table 2: STANDARDIZED MEAN SQUARED/ABSOLUTE ERRORS

COMMON COMPONENTS						
q	T	n	$u_{jt}, \xi_{it} \sim \mathcal{N}(0, 1)$		$u_{jt}, \xi_{it} \sim t_5$	
			S-MSE	S-MAE	S-MSE	S-MAE
1	120	120	0.29	1.13	0.29	1.23
1	240	240	0.04	0.21	0.19	0.75
1	480	480	0.02	0.15	0.08	0.35
2	120	120	0.28	0.93	0.28	0.95
2	240	240	0.17	0.57	0.16	0.56
2	480	480	0.08	0.32	0.08	0.34

Table 3: MULTIVARIATE R^2
COMMON SHOCKS

q	T	n	$u_{jt}, \xi_{it} \sim \mathcal{N}(0, 1)$	$u_{jt}, \xi_{it} \sim t_5$
1	120	120	0.96	0.95
1	240	240	0.98	0.97
1	480	480	0.99	0.98
2	120	120	0.92	0.91
2	240	240	0.95	0.95
2	480	480	0.98	0.97

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—of course, the empirical distribution of $Z_{it}^{(b)}$ worsens a little bit when there are multiple factors, or when the shocks are heavy-tailed, but they are still reasonably close to a normal distribution. As shown in Proposition 1, the fatter than normal tails are, the slower is the convergence in estimating the spectral density and therefore in estimating $\mathbf{A}_n(L)$, and in this case the first step of estimation has a larger impact on the asymptotic distributions.

8 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.¹⁸

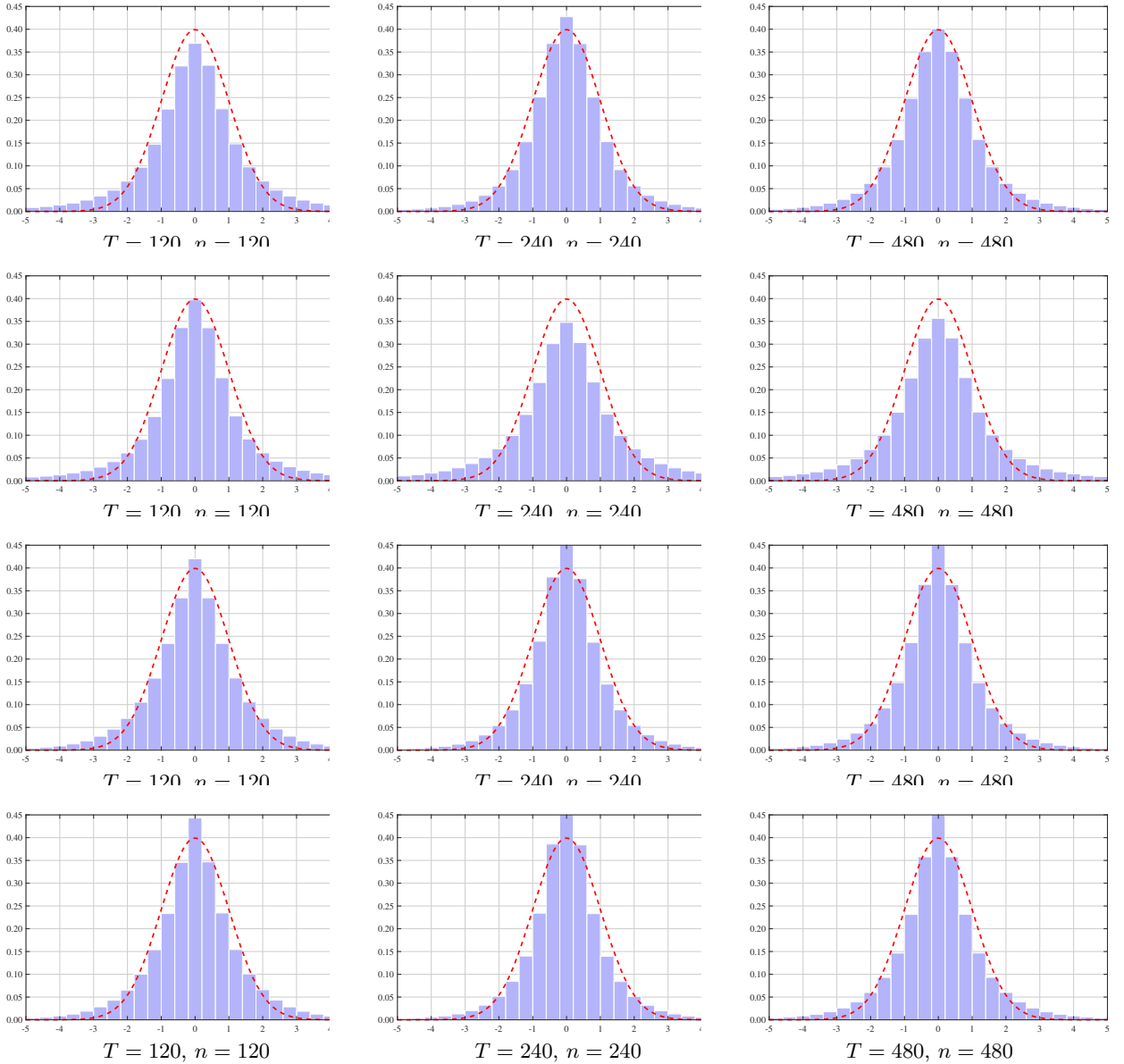
Nowadays, the notion of core inflation in the U.S. is mainly associated with inflation excluding food and energy. The rationale for this indicator is that both food and energy prices are very volatile and often driven by idiosyncratic shocks (such as weather for food or OPEC decisions for energy). Thus, 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). Therefore, the literature has proposed alternative ways of measuring core inflation, such as trimmed means and factor model-based estimates.¹⁹

The idea of considering (low-dimensional) factor models to estimate core inflation dates back to

¹⁸Altissimo et al. (2009) estimate a dynamic factor model, on disaggregated inflation data that represents an over-simplified case of our setting, as it is assuming that the common components follow AR(1) processes with i.i.d. idiosyncratic components. This simplification allows to use a different estimation method. Unlike us, they estimate their model on euro area data.

¹⁹The rationale for the use of trimmed means as core inflation indexes is that a trimmed mean is a robust estimator of 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.

Figure 1: HISTOGRAMS OF THE SIMULATED $Z_{it}^{(b)}$ 'S IN (76), FOR VARIOUS VALUES OF n AND T



Bryan and Cecchetti (1993), while Cristadoro et al. (2005) and ? more recently have used high-dimensional dynamic factor models, similar to the GDFM, with the same objective.²⁰ The rationale for considering factor models in 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 ($T = 300$). 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

²⁰Other 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.

Table 4: PERCENTAGE OF EXPLAINED VARIANCE

	1	2	3	4	5	6	7	8	9	10
q	8.5	6.2	5.5	4.9	4.5	4.2	3.8	3.6	3.3	3.1

This table reports the percentage of total variance explained by the q largest eigenvalues of the spectral density matrix of the data.

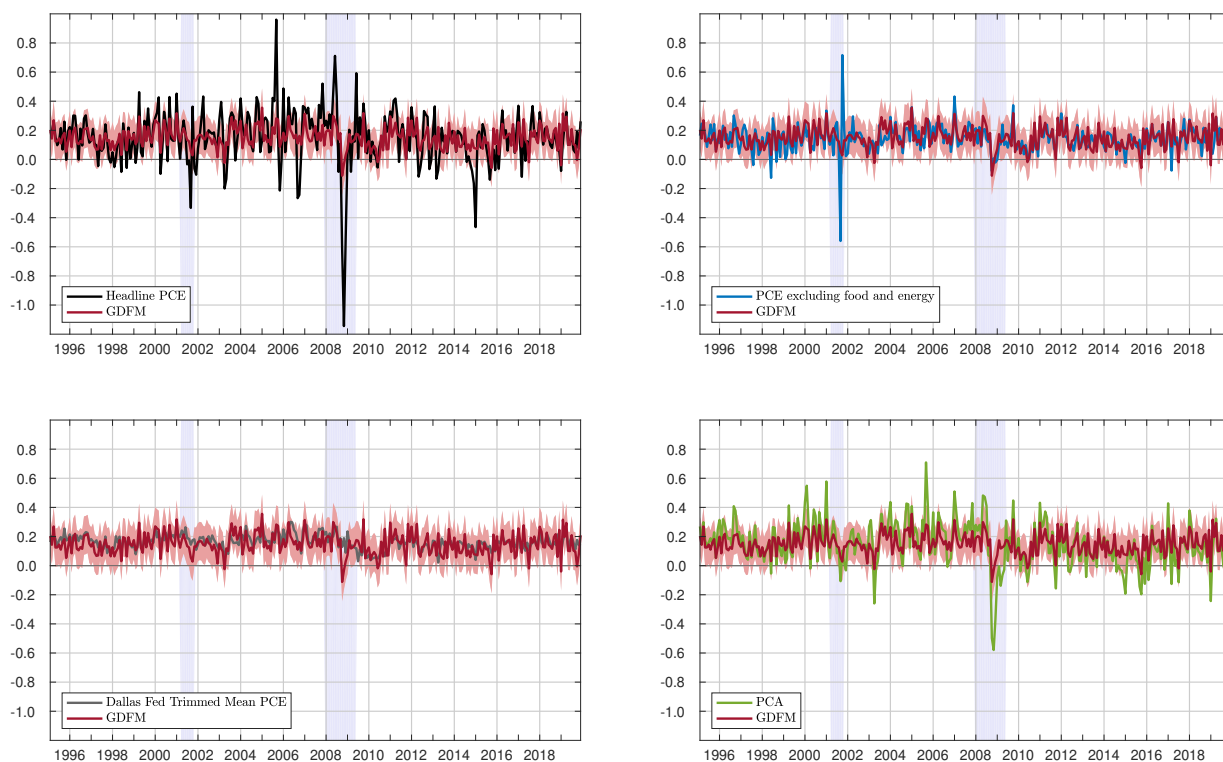
food and energy, and 146 disaggregated PCE prices. These 146 disaggregated PCE prices represent a particular disaggregation of PCE prices in which each disaggregated price index is constructed from a distinct data source. Indeed, most disaggregated PCE prices are measured using a corresponding index 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).

To determine the number q of factors, we look at the behavior of the eigenvalues of the spectral density matrix. From Table 4, we see that the first eigenvalue seems to separate from the rest, thus suggesting the presence of only one factor. This intuition is confirmed by the information criterion proposed by Hallin and Liška (2007), which exploits the behavior of the eigenvalues of the spectral density matrix of the data averaged across all frequencies. It is worth noting that the largest eigenvalue accounts only for a relatively small fraction of the variance. This is a known stylized fact of disaggregated monthly PCE prices, which are very volatile in an idiosyncratic way, and thus one common factor cannot capture all this high frequency. For example, it is well known that food prices are driven to a great extent by idiosyncratic factors, such as weather (e.g., draughts and hurricanes) or disease (e.g., avian flu). Likewise, nonenergy goods prices are also known to be idiosyncratic because they are primarily imported and thus not related to the U.S. business cycle (see Luciani, 2020, for a detailed discussion of commonality and idiosyncrasy in disaggregated PCE prices). However, if one were to filter out the “ultra-high” frequency fluctuations (i.e., those with a period shorter than six months), the common factor would account for a much larger share of the variance.

The specification used in this section features $q = 1$ common factor, the spectral density is estimated using a Bartlett kernel with bandwidth $B_T = \lfloor T^{1/3} \rfloor = 6$, the $m = n/(q + 1) = 74$ two-dimensional singular VARs are estimated using one lag, as determined via standard BIC, and the MA coefficients in the expression for the estimated dynamic common component are truncated at lag $K = 20$ (notice that the average lag-1 autocorrelation across the n series is 0.6, and, since $0.6^{20} \simeq 4 \cdot 10^{-5}$, the truncation error is clearly negligible). Finally, because $n \ll T$, to estimate \mathbf{U}_T and \mathbf{R}_n , and therefore the common components and the asymptotic variances, we set $\omega_{nT} = 1$. In light of this, we just need to set \bar{n} which, as in Section 7 we choose to be equal to n .

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). 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

Figure 2: “CORE” PCE PRICE MONTH-OVER-MONTH INFLATION



In all charts, the red line is based on our estimate, the shaded area is the asymptotic \pm one standard deviation confidence band.

rates, and then the estimated common component is computed by converting the month-over-month estimate into an year-over-year estimate.²¹

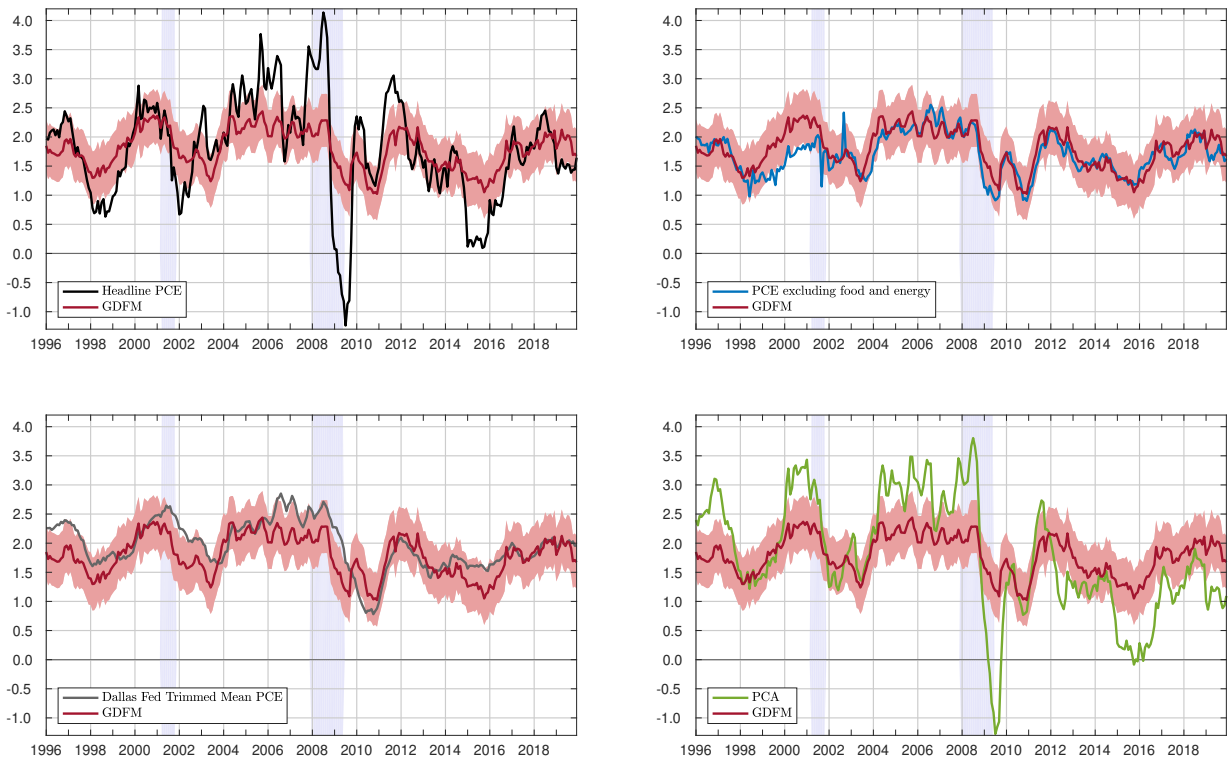
From simple visual inspection of the upper-left charts in Figures 2 and 3, we immediately see that our measure of core inflation is doing what it is supposed to do: tracking the trend of headline PCE price inflation while reducing the variance. Moreover, the confidence band seems to be quite well calibrated, 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 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.²² Moreover, as shown in Figure 4, our estimate of core inflation captures primarily fluctuations

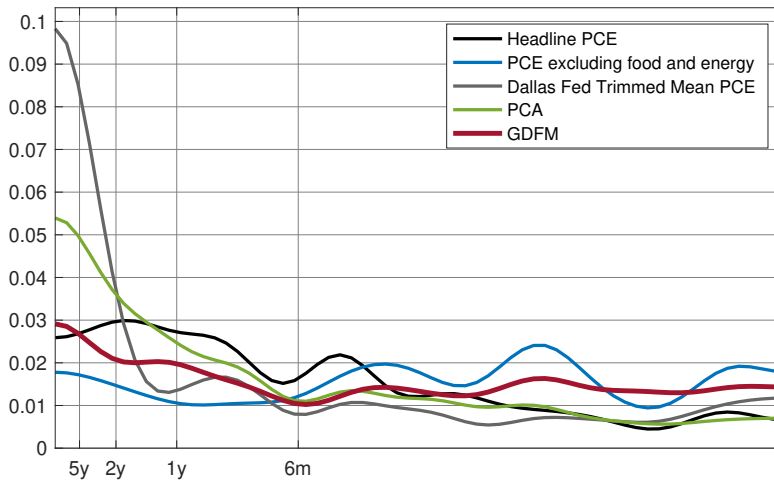
²¹As for the asymptotic variance, we took a shortcut for year-over-year estimates. Indeed we compute the variance for year-over-year estimates as $12\times$ the asymptotic variance over the month-over-month estimates. However, in doing so we are neglecting the autocorrelations, hence we can say that the confidence bands shown in Figure 3 are an approximation, which, most likely, are slightly tighter than they should be.

²²The 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 U.S.) brought back unlimited data plans.

Figure 3: “CORE” PCE PRICE YEAR-OVER-YEAR INFLATION



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



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.

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, 2021).

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.²³ Our measure and 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 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.

9 Conclusion

In the past decades, factor models have emerged as the most efficient tool for analyzing and predicting high-dimensional time series (high-dimensional panel data). The literature has proposed several factor models, 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 pave the way for inferential applications of the GDFM of great interest to macro and applied economists, such as asymptotic confidence intervals in prediction, impulse responses, and the construction of economic indicators. We illustrate the use of our methodology by constructing a new “core” inflation indicator for the U.S. economy. The GDFM-based indicator appears to provide more stable results than the current methods—it also outperforms its DFM-based counterpart estimated via Principal Components, which appears to be much more volatile.

References

- Altissimo, F., R. Cristadoro, M. Forni, M. Lippi, and G. Veronese (2010). New Eurocoin: Tracking economic growth in real time. *The Review of Economics and Statistics* 92(4), 1024–1034.
- Altissimo, F., B. Mojon, and P. Zaffaroni (2009). Can aggregation explain the persistence of inflation? *Journal of Monetary Economics* 56, 231–241.
- 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.
- Hallin, M., Mathias, Ch., Pirotte, H., and D. Veredas (2011). Market liquidity as dynamic factors. *Journal of Econometrics* 163, 42–50.
- Trucíos, C., Mazzeu, J.H.G., Hallin, M., Hotta, L.K., Valls Pereira, P.L., and M. Zevallos (2022). Forecasting

²³The Dallas Fed Trimmed Mean PCE price inflation estimates core inflation by taking the weighted trimmed mean of 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.

- conditional covariance matrices in high-dimensional time series: a general dynamic factor approach. *Journal of Business & Economic Statistics* 41, 40–52.
- Avarucci, M., M. Cavicchioli, M. Forni, and P. Zaffaroni (2022). The main business cycle shock(s): frequency-band estimation of the number of dynamic factors. CEPR Discussion Paper No. DP17281.
- Bai, J. (2003). Inferential theory for factor models of large dimensions. *Econometrica* 71, 135–171.
- Bai, J. and K. Li (2016). Maximum likelihood estimation and inference for approximate factor models of high dimension. *The Review of Economics and Statistics* 98, 298–309.
- Bai, J. and S. Ng (2002). Determining the number of factors in approximate factor models. *Econometrica* 70, 191–221.
- Bai, J. and S. Ng (2006). Confidence intervals for diffusion index forecasts and inference for factor augmented regressions. *Econometrica* 74, 1133–1150.
- Barigozzi, M. and M. Farnè (2022). An algebraic estimator for large spectral density matrices. *Journal of the American Statistical Association*, to appear.
- Barigozzi, M. and M. Hallin (2017). Generalized dynamic factor models and volatilities: estimation and forecasting. *Journal of Econometrics* 201, 307–321.
- Barigozzi, M. and M. Hallin (2020). Generalized dynamic factor models and volatilities: consistency, rates, and prediction intervals. *Journal of Econometrics* 216, 4–34.
- 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 (2019). Quasi maximum likelihood estimation and inference of large approximate dynamic factor models via the em algorithm. arXiv:1910.03821.v4.
- Barigozzi, M. and M. Luciani (2021). Measuring the output gap using large datasets. *The Review of Economics and Statistics*, 1–45.
- 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.
- Chamberlain, G. and M. Rothschild (1983). Arbitrage, factor structure, and mean-variance analysis on large asset markets. *Econometrica* 51, 1281–1304.
- 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.
- Doz, C., D. Giannone, and L. Reichlin (2012). A quasi maximum likelihood approach for large approximate dynamic factor models. *The Review of Economics and Statistics* 94(4), 1014–1024.
- 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., D. Giannone, M. Lippi, and L. Reichlin (2009). Opening the black box: structural factor models with large cross sections. *Econometric Theory* 25, 1319–1347.
- 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.
- Giannone, D. and T. Matheson (2007). A new core inflation indicator for new zealand. *International Journal of Central Banking* 3(4), 145–180.
- Giannone, D., L. Reichlin, and D. Small (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics* 55(4), 665–676.
- 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.
- Hallin, M. and R. Liška (2011). Dynamic factors in the presence of blocks. *Journal of Econometrics* 163, 29–41.
- Hannan, E. J. (1970). *Multiple Time Series*. John Wiley & Sons.
- 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. (2014). Forecasting with approximate dynamic factor models: the role of non-pervasive shocks. *International Journal of Forecasting* 30(1), 20–29.
- 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.
- Onatski, A. (2010). Determining the number of factors from empirical distribution of eigenvalues. *The Review of Economics and Statistics* 92, 1004–1016.
- Onatski, A. (2012). Asymptotics of the principal components estimator of large factor models with weakly influential factors. *Journal of Econometrics* 168, 244–258.
- Peña, D., E. Smucler, and V. J. Yohai (2021). Sparse estimation of dynamic principal components for forecasting high-dimensional time series. *International Journal of Forecasting* 37(4), 1498–1508.
- 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.
- Stock, J. H. and M. W. Watson (2016). Dynamic Factor Models, Factor-Augmented Vector Autoregressions, and Structural Vector Autoregressions in Macroeconomics. In J. B. Taylor and H. Uhlig (Eds.), *Handbook of Macroeconomics*, Volume 2, pp. 415–525. Elsevier.
- White, H. (2001). *Asymptotic Theory for Econometricians*. Academic press.
- 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.

APPENDIX

This Appendix collects the proofs of the main results. For simplicity, we throughout assume that Assumptions (A) through (G) 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 $n, T \rightarrow \infty$,*

$$(i) \quad \left\| \frac{\mathbf{U}'_T \mathbf{U}_T}{T} - \mathbf{\Gamma}^u \right\| = O_P \left(\frac{1}{\sqrt{T}} \right) \text{ as } T \rightarrow \infty; \quad (ii) \quad \left\| \frac{\mathbf{R}'_n \mathbf{R}_n}{n} - \mathbf{\Sigma}^R \right\| = O_P \left(\frac{1}{\sqrt{n}} \right) \text{ as } n \rightarrow \infty.$$

Proof. Part (i) follows from the i.i.d.-ness of \mathbf{u}_t in Assumption (A-a-i) and (37); part (ii) from the i.i.d.-ness of \mathbf{R}_i in Assumption (C-b) and (38). \square

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

$$\frac{1}{\sqrt{\bar{n}}} \|\widehat{\mathbf{z}}_{\bar{n}t} - \mathbf{z}_{\bar{n}t}\| = O_P(\zeta_{nT}) \quad \text{as } n, T \rightarrow \infty.$$

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

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

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 $n, T \rightarrow \infty$,*

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

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

Proof. From (28) which, as shown in the paper in Remark 9, follows from Assumption (C-d-i), equation

(45), and Lemma 1(i), we obtain that, as $n, T \rightarrow \infty$,

$$\begin{aligned} \frac{1}{n} \left\| \tilde{\Gamma}_n^z - \mathcal{R}_n \Gamma^u \mathcal{R}'_n \right\| &= \frac{1}{n} \left\| \mathcal{R}_n \frac{\mathcal{U}'_T \mathcal{U}_T}{T} \mathcal{R}'_n + \frac{\Phi'_{nT} \Phi_{nT}}{T} - \mathcal{R}_n \Gamma^u \mathcal{R}'_n \right\| \\ &\leq \frac{1}{n} \left\| \mathcal{R}_n \frac{\mathcal{U}'_T \mathcal{U}_T}{T} \mathcal{R}'_n - \mathcal{R}_n \Gamma^u \mathcal{R}'_n \right\|_F + \frac{1}{n} \left\| \frac{\Phi'_{nT} \Phi_{nT}}{T} \right\| \\ &= \frac{1}{n} \left\| \Gamma_n^\phi \right\| + O_P \left(\frac{1}{\sqrt{T}} \right) \leq \frac{2\pi B^\phi}{n} + O_P \left(\frac{1}{\sqrt{T}} \right) \\ &= O_P \left(\max \left(\frac{1}{\sqrt{T}}, \frac{1}{n} \right) \right), \end{aligned}$$

which implies

$$\frac{1}{n} \left\| \tilde{\Lambda}_n^\psi - \Lambda_n^\psi \right\| \leq \frac{1}{n} \left\| \tilde{\Gamma}_n^z - \mathcal{R}_n \Gamma^u \mathcal{R}'_n \right\| = O_P \left(\max \left(\frac{1}{\sqrt{T}}, \frac{1}{n} \right) \right),$$

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

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

where μ_{nj}^ψ are the eigenvalues of $\mathcal{R}_n \Gamma^u \mathcal{R}'_n$ (satisfying (40) and Assumption (D)), $\mu_{n0}^\psi := \infty$, and $\mu_{n,q+1}^\psi = 0$. Similarly, for $\bar{n} \leq n$,

$$\left\| \tilde{\mathbf{P}}_{\bar{n}}^z - \mathbf{P}_{\bar{n}}^\psi \widehat{\mathbf{W}}_1^z \right\| \leq \frac{2^{3/2} \sqrt{q} \left\| \tilde{\Gamma}_{\bar{n}}^z - \mathcal{R}_{\bar{n}} \Gamma^u \mathcal{R}'_{\bar{n}} \right\|}{\min(\mu_{n0}^\psi - \mu_{n1}^\psi, \mu_{nq}^\psi - \mu_{n,q+1}^\psi)} = O_P \left(\frac{\bar{n}}{n} \max \left(\frac{1}{\sqrt{T}}, \frac{1}{n} \right) \right), \quad (77)$$

which completes the proof. \square

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

(i) $\frac{1}{n} \left\| \widehat{\Lambda}_n^z - \tilde{\Lambda}_n^z \right\| = O_P(\zeta_{nT})$;

(ii) there exists a $q \times q$ diagonal matrix $\widehat{\mathbf{W}}_2^z$ with entries ± 1 such that, for any $\bar{n} \leq n$ satisfying $\frac{1}{\bar{n}} + \frac{\bar{n}}{n} \rightarrow 0$ and $\bar{n} \rightarrow \infty$ as $n \rightarrow \infty$, $\left\| \tilde{\mathbf{P}}_{\bar{n}}^z - \widehat{\mathbf{P}}_{\bar{n}}^z \widehat{\mathbf{W}}_2^z \right\| = O_P\left(\frac{\bar{n}}{n} \zeta_{nT}\right)$.

Proof. It immediately follows from Lemma 2 that $\frac{1}{n} \left\| \widehat{\Gamma}_n^z - \tilde{\Gamma}_n^z \right\| = O_P(\zeta_{nT})$, which implies

$$\frac{1}{n} \left\| \widehat{\Lambda}_n^z - \tilde{\Lambda}_n^z \right\| \leq \frac{1}{n} \left\| \widehat{\Gamma}_n^z - \tilde{\Gamma}_n^z \right\| = O_P(\zeta_{nT}),$$

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

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

Thus, from (40), with probability tending to one as $n, T \rightarrow \infty$,

$$\tilde{\mu}_{nj}^z \geq \mu_{nj}^\psi - c \max \left(\frac{1}{\sqrt{T}}, \frac{1}{n} \right) \geq n \underline{\mu}_j^\psi - c, \quad j = 1, \dots, q,$$

and $\tilde{\mu}_{nj}^z \leq c$, $j = q+1, \dots, n$. Therefore, for $n \geq 4c/\underline{\mu}_j^\psi$, with probability tending to one as $n, T \rightarrow \infty$,

it holds that

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

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

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

where $\tilde{\mu}_{nj}^z$ is the j th eigenvalue of $\tilde{\mathbf{\Gamma}}_n^z$ and $\tilde{\mu}_{n0}^z := \infty$. It follows that, for any $\bar{n} \leq n$,

$$\left\| \widehat{\mathbf{P}}_{\bar{n}}^z - \tilde{\mathbf{P}}_{\bar{n}}^z \widehat{\mathbf{W}}_2^z \right\| = O_P\left(\frac{\bar{n}}{n} \zeta_{nT}\right). \quad \square$$

Lemma 5. As $n, T \rightarrow \infty$,

$$(i) \frac{1}{n} \|\widehat{\mathbf{\Lambda}}_n^z - \mathbf{\Lambda}_n^\psi\| = O_P(\zeta_{nT});$$

(ii) for any $\bar{n} \leq n$ such that $\frac{1}{\bar{n}} + \frac{\bar{n}}{n} \rightarrow 0$ as $n \rightarrow \infty$, $\|\widehat{\mathbf{P}}_{\bar{n}}^z - \mathbf{P}_{\bar{n}}^\psi \widehat{\mathbf{W}}^z\| = O_P\left(\frac{\bar{n}}{n} \zeta_{nT}\right)$, 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

$$\frac{1}{n} \|\widehat{\mathbf{\Lambda}}_n^z - \mathbf{\Lambda}_n^\psi\| \leq \frac{1}{n} \|\widehat{\mathbf{\Lambda}}_n^z - \tilde{\mathbf{\Lambda}}_n^z\| + \frac{1}{n} \|\tilde{\mathbf{\Lambda}}_n^z - \mathbf{\Lambda}_n^\psi\| = O_P(\zeta_{nT}) + O_P\left(\max\left(\frac{1}{\sqrt{T}}, \frac{1}{n}\right)\right).$$

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) we obtain, since $\|\widehat{\mathbf{W}}_2^z\| = 1$,

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

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

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

$$\frac{\mathbf{\Lambda}_n^\psi}{n} \rightarrow_P \mathbf{L}^u \quad \text{as } n \rightarrow \infty.$$

Proof. The Lemma is an immediate consequence of (40). \square

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

Proof. Part (i) follows from (40), part (ii) from Lemma 5(i) and part (i). \square

Lemma 8. Denoting by \mathbf{e}_{ni} the i th column of \mathbf{I}_n ,

$$\max_{i=1, \dots, n} \left\| \mathbf{e}_{ni}' \mathbf{P}_n^\psi \right\| = O_P\left(\frac{1}{\sqrt{n}}\right) \quad \text{as } n \rightarrow \infty.$$

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

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

Indeed, $\|\mathbf{e}'_{ni}\mathbf{R}_n\mathbf{\Gamma}^u\mathbf{R}'_n\| = O_P(\sqrt{n})$, $\|\mathbf{P}_n^\psi\| = 1$, and $\|(\mathbf{\Lambda}_n^\psi)^{-1}\| = O_P(n^{-1})$, because of Lemma 7(i). \square

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

$$(i) \quad \|\mathbf{P}_{\bar{n}}^\psi\| = O_P\left(\sqrt{\frac{\bar{n}}{n}}\right) \quad \text{and} \quad (ii) \quad \|\mathbf{P}_{\bar{n}}^{\psi'}\mathbf{P}_{\bar{n}}^\psi\| = O_P\left(\frac{\bar{n}}{n}\right).$$

Proof. It follows from Lemma 8 that

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

Moreover,

$$\|\mathbf{P}_{\bar{n}}^{\psi'}\mathbf{P}_{\bar{n}}^\psi\| \leq \|\mathbf{P}_{\bar{n}}^\psi\|^2 = O_P\left(\frac{\bar{n}}{n}\right). \quad \square$$

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

$$\sqrt{\frac{n}{\bar{n}}}\|\mathbf{P}_{\bar{n}}^{\psi'}\phi_{\bar{n}t}\| = O_P(1) \quad \text{as } n \rightarrow \infty.$$

Proof. Recall that $\frac{n}{\bar{n}}\|\mathbf{P}_{\bar{n}}^{\psi'}\mathbf{P}_{\bar{n}}^\psi\| = O_P(1)$, because of Lemma 9(ii). Therefore, for the k th column of $\mathbf{P}_{\bar{n}}^\psi$, denoted as $\mathbf{p}_{\bar{n}k}^\psi$, it holds that $\frac{n}{\bar{n}}\mathbf{p}_{\bar{n}k}^{\psi'}\mathbf{p}_{\bar{n}k}^\psi = O_P(1)$. Let $\tilde{\mathbf{p}}_{\bar{n}k}^\psi := \mathbf{p}_{\bar{n}k}^\psi / \sqrt{\mathbf{p}_{\bar{n}k}^{\psi'}\mathbf{p}_{\bar{n}k}^\psi}$, so that $\tilde{\mathbf{p}}_{\bar{n}k}^{\psi'}\tilde{\mathbf{p}}_{\bar{n}k}^\psi = 1$. Let \tilde{p}_{ik}^ψ denote the i th entry of $\tilde{\mathbf{p}}_{\bar{n}k}^\psi$ and let $\tilde{\mathbf{P}}_{\bar{n}}^\psi$ be the matrix with columns $\tilde{\mathbf{p}}_{\bar{n}1}^\psi, \dots, \tilde{\mathbf{p}}_{\bar{n}\bar{n}}^\psi$. Due to normalization of $\tilde{\mathbf{p}}_{\bar{n}k}^\psi$ and Lemma 8, there exists a finite positive real \bar{c} such that $\max_{i=1,\dots,n} \max_{j=1,\dots,q} |\tilde{p}_{ij}^\psi| \leq \bar{c}/\sqrt{\bar{n}}$ with probability one. Then, denoting by $\boldsymbol{\iota}_{\bar{n}}$ a \bar{n} -dimensional column vector of ones, for any $t \in \mathbb{Z}$,

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

in view of (45). Hence, it follows from Chebychev's inequality that $\|\tilde{\mathbf{P}}_{\bar{n}}^{\psi'}\phi_{\bar{n}t}\| = O_P(1)$ and, therefore, $\|\mathbf{P}_{\bar{n}}^{\psi'}\phi_{\bar{n}t}\|$ is $O_P\left(\sqrt{\frac{\bar{n}}{n}}\right)$. \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 (51)

$$\begin{aligned}
\widehat{\mathbf{u}}_t &= \left(\left(\widehat{\boldsymbol{\Lambda}}_n^z \right)^{1/2} \widehat{\mathbf{P}}_{\bar{n}}^{z'} \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'} \widehat{\mathbf{z}}_{\bar{n}t} = \left(\widehat{\boldsymbol{\Lambda}}_n^z \right)^{-1/2} \left(\widehat{\mathbf{P}}_{\bar{n}}^{z'} \widehat{\mathbf{P}}_{\bar{n}}^z \right)^{-1} \widehat{\mathbf{P}}_{\bar{n}}^{z'} \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'} \widehat{\mathbf{P}}_{\bar{n}}^z \right)^{-1} \widehat{\mathbf{P}}_{\bar{n}}^{z'} \widehat{\mathbf{z}}_{\bar{n}t} \\
&\quad + \left(\boldsymbol{\Lambda}_n^\psi \right)^{-1/2} \left(\left(\widehat{\mathbf{P}}_{\bar{n}}^{z'} \widehat{\mathbf{P}}_{\bar{n}}^z \right)^{-1} - \left(\widehat{\mathbf{W}}^z \mathbf{P}_{\bar{n}}^{\psi'} \mathbf{P}_{\bar{n}}^\psi \widehat{\mathbf{W}}^z \right)^{-1} \right) \widehat{\mathbf{P}}_{\bar{n}}^{z'} \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'} \mathbf{P}_{\bar{n}}^\psi \widehat{\mathbf{W}}^z \right)^{-1} \left(\widehat{\mathbf{P}}_{\bar{n}}^{z'} - \widehat{\mathbf{W}}^z \mathbf{P}_{\bar{n}}^{\psi'} \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'} \mathbf{P}_{\bar{n}}^\psi \right)^{-1} \mathbf{P}_{\bar{n}}^{\psi'} \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'} \mathbf{P}_{\bar{n}}^\psi \right)^{-1} \mathbf{P}_{\bar{n}}^{\psi'} \mathbf{z}_{\bar{n}t} \\
&= I + II + III + IV + V, \text{ say.}
\end{aligned} \tag{78}$$

For I , since

$$\begin{aligned}
\left(\left(\widehat{\boldsymbol{\Lambda}}_n^z \right)^{-1/2} - \left(\boldsymbol{\Lambda}_n^\psi \right)^{-1/2} \right) &= \left(\left(\widehat{\boldsymbol{\Lambda}}_n^z \right)^{-1} - \left(\boldsymbol{\Lambda}_n^\psi \right)^{-1} \right) \left(\left(\widehat{\boldsymbol{\Lambda}}_n^z \right)^{-1/2} + \left(\boldsymbol{\Lambda}_n^\psi \right)^{-1/2} \right)^{-1} \\
&= \left(\widehat{\boldsymbol{\Lambda}}_n^z \right)^{-1} \left(\boldsymbol{\Lambda}_n^\psi - \widehat{\boldsymbol{\Lambda}}_n^z \right) \left(\boldsymbol{\Lambda}_n^\psi \right)^{-1} \left(\left(\widehat{\boldsymbol{\Lambda}}_n^z \right)^{-1/2} + \left(\boldsymbol{\Lambda}_n^\psi \right)^{-1/2} \right)^{-1}
\end{aligned} \tag{79}$$

and because of (79) and Lemmas 5(i) and 7, the norm of I is bounded from above by

$$\begin{aligned}
\left\| \left(\widehat{\boldsymbol{\Lambda}}_n^z \right)^{-1} \right\| \left\| \boldsymbol{\Lambda}_n^\psi - \widehat{\boldsymbol{\Lambda}}_n^z \right\| \left\| \left(\boldsymbol{\Lambda}_n^\psi \right)^{-1} \right\| \left\| \left(\left(\widehat{\boldsymbol{\Lambda}}_n^z \right)^{-1/2} + \left(\boldsymbol{\Lambda}_n^\psi \right)^{-1/2} \right)^{-1} \right\| \left\| \left(\widehat{\mathbf{P}}_{\bar{n}}^{z'} \widehat{\mathbf{P}}_{\bar{n}}^z \right)^{-1} \widehat{\mathbf{P}}_{\bar{n}}^{z'} \right\| \left\| \widehat{\mathbf{z}}_{\bar{n}t} \right\| \\
= O_{\mathbb{P}} \left(\frac{1}{n^2} \sqrt{n} \zeta_{nT} \sqrt{n} \frac{\sqrt{n}}{\sqrt{n}} \sqrt{n} \right)
\end{aligned} \tag{80}$$

since $\|\widehat{\mathbf{z}}_{\bar{n}t}\| = O_{\mathbb{P}}(\sqrt{n})$ by Lemma 2, and $\|(\widehat{\mathbf{P}}_{\bar{n}}^{z'} \widehat{\mathbf{P}}_{\bar{n}}^z)^{-1} \widehat{\mathbf{P}}_{\bar{n}}^{z'}\| = O_{\mathbb{P}}(\sqrt{n/\bar{n}})$ by Lemma 9(i) and (ii). This yields $I = O_{\mathbb{P}}(\zeta_{nT}/\sqrt{n})$.

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

$$\left\| \mathbf{P}_{\bar{n}}^\psi \widehat{\mathbf{W}}^z - \widehat{\mathbf{P}}_{\bar{n}}^z \right\| = O_{\mathbb{P}} \left(\frac{\bar{n}}{n} \zeta_{nT} \right). \tag{81}$$

Then,

$$\begin{aligned}
&\left(\left(\widehat{\mathbf{P}}_{\bar{n}}^{z'} \widehat{\mathbf{P}}_{\bar{n}}^z \right)^{-1} - \left(\widehat{\mathbf{W}}^z \mathbf{P}_{\bar{n}}^{\psi'} \mathbf{P}_{\bar{n}}^\psi \widehat{\mathbf{W}}^z \right)^{-1} \right) \\
&= \left(\widehat{\mathbf{P}}_{\bar{n}}^{z'} \widehat{\mathbf{P}}_{\bar{n}}^z \right)^{-1} \left(\widehat{\mathbf{W}}^z \mathbf{P}_{\bar{n}}^{\psi'} \mathbf{P}_{\bar{n}}^\psi \widehat{\mathbf{W}}^z - \widehat{\mathbf{P}}_{\bar{n}}^{z'} \widehat{\mathbf{P}}_{\bar{n}}^z \right) \left(\widehat{\mathbf{W}}^z \mathbf{P}_{\bar{n}}^{\psi'} \mathbf{P}_{\bar{n}}^\psi \widehat{\mathbf{W}}^z \right)^{-1} \\
&= \left(\widehat{\mathbf{P}}_{\bar{n}}^{z'} \widehat{\mathbf{P}}_{\bar{n}}^z \right)^{-1} \left(\widehat{\mathbf{W}}^z \mathbf{P}_{\bar{n}}^{\psi'} \left(\mathbf{P}_{\bar{n}}^\psi \widehat{\mathbf{W}}^z - \widehat{\mathbf{P}}_{\bar{n}}^z \right) + \left(\widehat{\mathbf{W}}^z \mathbf{P}_{\bar{n}}^{\psi'} - \widehat{\mathbf{P}}_{\bar{n}}^{z'} \right) \widehat{\mathbf{P}}_{\bar{n}}^z \right) \left(\widehat{\mathbf{W}}^z \mathbf{P}_{\bar{n}}^{\psi'} \mathbf{P}_{\bar{n}}^\psi \widehat{\mathbf{W}}^z \right)^{-1}
\end{aligned}$$

and, because of (81) and Lemma 9,

$$\left\| \left(\left(\widehat{\mathbf{P}}_{\bar{n}}^{z'} \widehat{\mathbf{P}}_{\bar{n}}^z \right)^{-1} - \left(\widehat{\mathbf{W}}^z \mathbf{P}_{\bar{n}}^{\psi'} \mathbf{P}_{\bar{n}}^\psi \widehat{\mathbf{W}}^z \right)^{-1} \right) \right\| = O_{\mathbb{P}} \left(\sqrt{\frac{\bar{n}}{n}} \frac{\bar{n}}{n} \zeta_{nT} \frac{n}{\bar{n}} \right) = O_{\mathbb{P}} \left(\sqrt{\frac{\bar{n}}{n}} \zeta_{nT} \right). \tag{82}$$

Because of (82), Lemmas 2, 7(i), and 9(i), the norm of II is bounded from above by

$$\begin{aligned} & \left\| \left(\mathbf{\Lambda}_n^\psi \right)^{-1/2} \right\| \left\| \left(\left(\widehat{\mathbf{P}}_n^{z'} \widehat{\mathbf{P}}_n^z \right)^{-1} - \left(\widehat{\mathbf{W}}^z \mathbf{P}_n^{\psi'} \mathbf{P}_n^\psi \widehat{\mathbf{W}}^z \right)^{-1} \right) \right\| \left\| \widehat{\mathbf{P}}_n^{z'} \right\| \left\| \widehat{\mathbf{z}}_{nt} \right\| \\ & = O_P \left(\frac{1}{\sqrt{\bar{n}}} \sqrt{\frac{\bar{n}}{n}} \zeta_{nT} \sqrt{\frac{\bar{n}}{n}} \sqrt{\bar{n}} \right) = O_P \left(\sqrt{\frac{\bar{n}}{n}} \zeta_{nT} \right), \end{aligned} \quad (83)$$

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

By (81) and Lemmas 2, 7(i), and 9(i), one immediately gets $III = O_P \left(\sqrt{\frac{\bar{n}}{n}} \zeta_{nT} \right)$ and $IV = O_P \left(\frac{\zeta_{nT}}{\sqrt{n}} \right)$.

Finally, consider term V . Recall that, from Assumption (S-d1), (45), (58), and Lemma 1(i), for any $n \in \mathbb{N}$, as $T \rightarrow \infty$,

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

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

$$\widehat{\mathbf{u}}_t - \widehat{\mathbf{W}}^z \left(\mathbf{\Gamma}^u \right)^{-1/2} \mathbf{u}_t = I + II + III + IV + \widehat{\mathbf{W}}^z \left(\mathbf{\Lambda}_n^\psi \right)^{-1/2} \left(\mathbf{P}_{\bar{n}}^{\psi'} \mathbf{P}_{\bar{n}}^\psi \right)^{-1} \mathbf{P}_{\bar{n}}^{\psi'} \phi_{\bar{n}t}. \quad (84)$$

Recalling that $\|\widehat{\mathbf{W}}^z\| = 1$, it follows from (84) that, in view of Lemmas 7(i), 9(ii), and 10,

$$\begin{aligned} \left\| \widehat{\mathbf{u}}_t - \widehat{\mathbf{W}}^z \left(\mathbf{\Gamma}^u \right)^{-1/2} \mathbf{u}_t \right\| & \leq \left\| \left(\mathbf{\Lambda}_n^\psi \right)^{-1/2} \right\| \left\| \left(\mathbf{P}_{\bar{n}}^{\psi'} \mathbf{P}_{\bar{n}}^\psi \right)^{-1} \right\| \left\| \mathbf{P}_{\bar{n}}^{\psi'} \phi_{\bar{n}t} \right\| + O_P \left(\sqrt{\frac{\bar{n}}{n}} \zeta_{nT} \right) \\ & = O_P \left(\frac{1}{\sqrt{\bar{n}}} \frac{n}{\bar{n}} \sqrt{\frac{\bar{n}}{n}} \right) + O_P \left(\sqrt{\frac{\bar{n}}{n}} \zeta_{nT} \right) = O_P \left(\max \left(\frac{1}{\sqrt{\bar{n}}}, \sqrt{\frac{\bar{n}}{n}} \zeta_{nT} \right) \right). \end{aligned}$$

This proves consistency.

Now, by (40), there exists a $q \times q$ positive definite diagonal matrix \mathcal{L}_u such that $\mathbf{\Lambda}_n^\psi/n \rightarrow_P \mathcal{L}^u$ as $n \rightarrow \infty$. Similarly, by Lemma 9(ii), there exists a $q \times q$ positive definite matrix \mathcal{M}_u such that, as $n \rightarrow \infty$, $\frac{n}{\bar{n}} \mathbf{P}_{\bar{n}}^{\psi'} \mathbf{P}_{\bar{n}}^\psi \rightarrow_P \mathcal{M}^u$. Therefore, by Assumption (E), as $n, T \rightarrow \infty$,

$$\begin{aligned} \sqrt{\bar{n}} \left(\widehat{\mathbf{u}}_t - \widehat{\mathbf{W}}^z \left(\mathbf{\Gamma}^u \right)^{-1/2} \mathbf{u}_t \right) & = \sqrt{\bar{n}} \widehat{\mathbf{W}}^z \left(\mathbf{\Lambda}_n^\psi \right)^{-1/2} \left(\mathbf{P}_{\bar{n}}^{\psi'} \mathbf{P}_{\bar{n}}^\psi \right)^{-1} \mathbf{P}_{\bar{n}}^{\psi'} \phi_{\bar{n}t} + o_P(1) \\ & = \widehat{\mathbf{W}}^z \left(\frac{\mathbf{\Lambda}_n^\psi}{n} \right)^{-1/2} \left(\frac{n}{\bar{n}} \mathbf{P}_{\bar{n}}^{\psi'} \mathbf{P}_{\bar{n}}^\psi \right)^{-1} \sqrt{\frac{\bar{n}}{n}} \left(\mathbf{P}_{\bar{n}}^{\psi'} \phi_{\bar{n}t} \right) + o_P(1) \\ & \rightarrow_d \mathcal{N} \left(\mathbf{0}_q, \mathbf{W}^u \left(\mathcal{L}^u \right)^{-1/2} \left(\mathcal{M}^u \right)^{-1} \mathcal{P}_t^u \left(\mathcal{M}^u \right)^{-1} \left(\mathcal{L}^u \right)^{-1/2} \mathbf{W}^u \right) \end{aligned}$$

since $\sqrt{\bar{n}} \zeta_{nT} \rightarrow 0$, because of (60). \square

B Proof of Theorem 2

B.1 Preliminary lemmas

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

$$(i) \frac{1}{T} \left\| \widehat{\mathbf{L}}_T^z - \mathbf{L}_T^\psi \right\| = O_P \left(\max \left(\frac{1}{\sqrt{n}}, \frac{1}{T} \right) \right);$$

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

Proof. The claim follows along the same lines as for Lemma 3 but using Assumption (C-d-ii), (29), (46), and Lemma 1(ii) instead of Assumption (C-d-i), (28), (45), and Lemma 1(i). \square

Lemma 12. Collect the q largest eigenvalues of $\widehat{\mathbf{G}}_T^z := \widehat{\mathbf{Z}}_{nT} \widehat{\mathbf{Z}}'_{nT}/n$ in the $q \times q$ diagonal matrix $\widehat{\mathbf{L}}_T^z$ and the corresponding normalized eigenvectors in $\widehat{\mathbf{\Pi}}_T^z$. As $n, T \rightarrow \infty$,

$$(i) \frac{1}{T} \left\| \widehat{\mathbf{L}}_T^z - \widetilde{\mathbf{L}}_T^z \right\| = O_P(\zeta_{nT});$$

(ii) there exists a $q \times q$ diagonal matrix $\widehat{\mathbf{W}}_1^z$ with entries ± 1 such that, for any $\bar{T} \leq T$ satisfying $\frac{1}{\bar{T}} + \frac{\bar{T}}{T} \rightarrow 0$ and $\bar{T} \rightarrow \infty$ as $T \rightarrow \infty$, $\left\| \widehat{\mathbf{\Pi}}_{\bar{T}}^z - \widehat{\mathbf{\Pi}}_{\bar{T}}^z \widehat{\mathbf{W}}_2^z \right\| = O_P \left(\frac{\bar{T}}{T} \zeta_{nT} \right).$

Proof. The claim follows along the same lines as for Lemma 4 but using Lemma 12 and (42). \square

Lemma 13. As $n, T \rightarrow \infty$,

$$(i) \frac{1}{T} \left\| \widehat{\mathbf{L}}_T^z - \mathbf{L}_T^\psi \right\| = O_P(\zeta_{nT});$$

(ii) for any $\bar{T} \leq T$ such that $\frac{1}{\bar{T}} + \frac{\bar{T}}{T} \rightarrow 0$ as $T \rightarrow \infty$, $\left\| \widehat{\mathbf{\Pi}}_{\bar{T}}^z - \mathbf{\Pi}_{\bar{T}}^\psi \widehat{\mathbf{W}}^z \right\| = O_P \left(\frac{\bar{T}}{T} \zeta_{nT} \right)$, 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 11 and $\widehat{\mathbf{W}}_2^z$ in Lemma 12.

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

Lemma 14. There exists a positive definite $q \times q$ diagonal matrix \mathcal{L}_R such that $\mathbf{L}_T^\psi/T \xrightarrow{P} \mathcal{L}_R$ as $T \rightarrow \infty$.

Proof. This Lemma is an immediate consequence of (42). \square

$$\text{Lemma 15. } (i) \left\| \left(\frac{\mathbf{L}_T^\psi}{T} \right)^{-1} \right\| = O_P(1) \text{ as } T \rightarrow \infty; \quad (ii) \left\| \left(\frac{\widehat{\mathbf{L}}_T^z}{T} \right)^{-1} \right\| = O_P(1) \text{ as } n, T \rightarrow \infty.$$

Proof. Part (i) follows from (42), part (ii) from Lemma 13(i) and part (i). \square

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

$$\max_{t=1, \dots, T} \left\| \mathbf{e}'_{Tt} \mathbf{\Pi}_T^\psi \right\| = O_P \left(\frac{1}{\sqrt{T}} \right) \text{ as } T \rightarrow \infty.$$

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

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

$$(i) \left\| \mathbf{\Pi}_{\bar{T}}^\psi \right\| = O_P \left(\sqrt{\frac{\bar{T}}{T}} \right); \quad (ii) \left\| \mathbf{\Pi}_{\bar{T}}^{\psi'} \mathbf{\Pi}_{\bar{T}}^\psi \right\| = O_P \left(\frac{\bar{T}}{T} \right).$$

Proof. Same as Lemma 9 but using Lemma 16. \square

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

$$\sqrt{\frac{T}{\bar{T}}} \left\| \mathbf{\Pi}_{\bar{T}}^{\psi'} \boldsymbol{\varphi}_{\bar{T}}^i \right\| = O_P(1).$$

Proof. Recall that, in view of Lemma 17(ii), $(T/\bar{T}) \left\| \mathbf{\Pi}_{\bar{T}}^{\psi'} \mathbf{\Pi}_{\bar{T}}^\psi \right\| = O_P(1)$. Therefore, for the k th column of $\mathbf{\Pi}_{\bar{T}}^\psi$, denoted as $\boldsymbol{\pi}_{\bar{T}k}^\psi$, it holds that $(T/\bar{T}) \boldsymbol{\pi}_{\bar{T}k}^{\psi'} \boldsymbol{\pi}_{\bar{T}k}^\psi = O_P(1)$. Let $\widetilde{\boldsymbol{\pi}}_{\bar{T}k}^\psi := \boldsymbol{\pi}_{\bar{T}k}^\psi / \sqrt{\boldsymbol{\pi}_{\bar{T}k}^{\psi'} \boldsymbol{\pi}_{\bar{T}k}^\psi}$,

so that $\tilde{\boldsymbol{\pi}}_{\bar{T}k}^{\psi'} \tilde{\boldsymbol{\pi}}_{\bar{T}k}^{\psi} = 1$. Let $\tilde{\pi}_{ik}^{\psi}$ be the i th entry of $\tilde{\boldsymbol{\pi}}_{\bar{T}k}^{\psi}$ and denote by $\tilde{\boldsymbol{\Pi}}_{\bar{T}}^{\psi}$ the $\bar{T} \times \bar{T}$ matrix with columns $\tilde{\boldsymbol{\pi}}_{\bar{T}1}^{\psi}, \dots, \tilde{\boldsymbol{\pi}}_{\bar{T}\bar{T}}^{\psi}$. Due to normalization of $\tilde{\boldsymbol{\pi}}_{\bar{T}k}^{\psi}$ and Lemma 16, there exists a finite positive real \bar{c} such that

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

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

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

because of (46) and since $\mathbf{G}_{\bar{T}}^{\phi} = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \mathbb{E} [\boldsymbol{\varphi}_{\bar{T}}^i \boldsymbol{\varphi}_{\bar{T}}^{i'}]$. From Chebychev's inequality, $\|\tilde{\boldsymbol{\Pi}}_{\bar{T}}^{\psi'} \boldsymbol{\varphi}_{\bar{T}}^i\| = O_{\mathbb{P}}(1)$ and, therefore, $\|\boldsymbol{\Pi}_{\bar{T}}^{\psi'} \boldsymbol{\varphi}_{\bar{T}}^i\| = O_{\mathbb{P}}\left(\sqrt{\frac{\bar{T}}{T}}\right)$. \square

B.2 Proof of Theorem 2

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

C Proof of Theorem 3

First, from the proof of Theorem 1, $\mathbf{R}'_i = \mathbf{p}_i^{\psi'} (\boldsymbol{\Lambda}_n^{\psi})^{1/2} (\boldsymbol{\Gamma}^u)^{-1/2}$ for any $i = 1, \dots, n$. Therefore, from the definition of $\check{\mathbf{R}}'_i$ in (49),

$$\begin{aligned} \check{\mathbf{R}}'_i - \mathbf{R}'_i (\boldsymbol{\Gamma}^u)^{1/2} \widehat{\mathbf{W}}^z &= \widehat{\mathbf{p}}_i^{z'} \left(\widehat{\boldsymbol{\Lambda}}_n^z \right)^{1/2} - \mathbf{p}_i^{\psi'} \left(\boldsymbol{\Lambda}_n^{\psi} \right)^{1/2} \widehat{\mathbf{W}}^z = \widehat{\mathbf{p}}_i^{z'} \left(\widehat{\boldsymbol{\Lambda}}_n^z \right)^{1/2} - \mathbf{p}_i^{\psi'} \widehat{\mathbf{W}}^z \left(\boldsymbol{\Lambda}_n^{\psi} \right)^{1/2} \\ &= \mathbf{p}_i^{\psi'} \widehat{\mathbf{W}}^z \left(\widehat{\boldsymbol{\Lambda}}_n^z - \boldsymbol{\Lambda}_n^{\psi} \right)^{1/2} + \left(\widehat{\mathbf{p}}_i^{z'} - \mathbf{p}_i^{\psi'} \widehat{\mathbf{W}}^z \right) \left(\boldsymbol{\Lambda}_n^{\psi} \right)^{1/2} + \left(\widehat{\mathbf{p}}_i^{z'} - \mathbf{p}_i^{\psi'} \widehat{\mathbf{W}}^z \right) \left(\widehat{\boldsymbol{\Lambda}}_n^z - \boldsymbol{\Lambda}_n^{\psi} \right)^{1/2} \\ &= I + II + III, \text{ say.} \end{aligned} \tag{85}$$

Term I is $O_{\mathbb{P}}(\zeta_{nT})$ because of Lemmas 5(i) and 8, term II is $O_{\mathbb{P}}\left(\sqrt{\frac{\bar{n}}{n}} \zeta_{nT}\right)$ because of Lemmas 5(ii) and 8 (see also the arguments in Lemma 6 of Forni et al., 2017), and term III is $o_{\mathbb{P}}(\zeta_{nT})$. From (85), we get

$$\left\| \check{\mathbf{R}}'_i - \mathbf{R}'_i (\boldsymbol{\Gamma}^u)^{1/2} \widehat{\mathbf{W}}^z \right\| = O_{\mathbb{P}}(\zeta_{nT}) \tag{86}$$

which, combined with Theorem 1(i), yields

$$\left\| \check{\mathbf{R}}'_i \widehat{\mathbf{u}}_t - \mathbf{R}'_i \mathbf{u}_t \right\| = O_{\mathbb{P}} \left(\max \left(\frac{1}{\sqrt{\bar{n}}}, \zeta_{nT} \right) \right) = O_{\mathbb{P}} \left(\max \left(\frac{1}{\sqrt{h}}, \zeta_{nT} \right) \right). \tag{87}$$

Following a reasoning similar to (85), since, from the proof of Theorem 2, for any $t = 1, \dots, T$ we

have $\mathbf{u}'_t = \boldsymbol{\pi}_t^{\psi'} (\mathbf{L}_T^\psi)^{1/2} (\boldsymbol{\Sigma}^R)^{-1/2}$, the definition of $\check{\mathbf{u}}'_t$ in (53) and Lemmas 13 and 16 imply that

$$\left\| \check{\mathbf{u}}'_t - \mathbf{u}'_t (\boldsymbol{\Sigma}^R)^{1/2} \widehat{\mathbf{W}}^z \right\| = O_{\mathbb{P}}(\zeta_{nT}) \quad (88)$$

which, combined with Theorem 2(i), yields

$$\left\| \check{\mathbf{u}}'_t \widehat{\mathbf{R}}_i - \mathbf{u}'_t \mathbf{R}_i \right\| = O_{\mathbb{P}} \left(\max \left(\frac{1}{\sqrt{\bar{n}}}, \zeta_{nT} \right) \right) = O_{\mathbb{P}} \left(\max \left(\frac{1}{\sqrt{\bar{h}}}, \zeta_{nT} \right) \right). \quad (89)$$

Part (i) of the theorem follows from (87) and (89).

Now, from the proof of Theorems 1 and 2 and using (86) and (88),

$$\begin{aligned} \check{\mathbf{R}}'_i \widehat{\mathbf{u}}_t &= \mathbf{R}'_i \mathbf{u}_t + \mathbf{R}'_i (\boldsymbol{\Gamma}^u)^{1/2} \widehat{\mathbf{W}}^z \left(\widehat{\mathbf{u}}_t - \widehat{\mathbf{W}}^z (\boldsymbol{\Gamma}^u)^{-1/2} \mathbf{u}_t \right) + \left(\check{\mathbf{R}}'_i - \mathbf{R}'_i (\boldsymbol{\Gamma}^u)^{1/2} \widehat{\mathbf{W}}^z \right) \widehat{\mathbf{u}}_t \\ &= \mathbf{R}'_i \mathbf{u}_t + \mathbf{R}'_i (\boldsymbol{\Gamma}^u)^{1/2} \widehat{\mathbf{W}}^z \left(\boldsymbol{\Lambda}_n^\psi \right)^{-1/2} \left(\mathbf{P}_{\bar{n}}^{\psi'} \mathbf{P}_{\bar{n}}^\psi \right)^{-1} \mathbf{P}_{\bar{n}}^{\psi'} \boldsymbol{\phi}_{\bar{n}t} + \left(\check{\mathbf{R}}'_i - \mathbf{R}'_i (\boldsymbol{\Gamma}^u)^{1/2} \widehat{\mathbf{W}}^z \right) \widehat{\mathbf{u}}_t + O_{\mathbb{P}}(\zeta_{nT}) \\ &= \mathbf{R}'_i \mathbf{u}_t + \mathbf{R}'_i (\boldsymbol{\Gamma}^u)^{1/2} \widehat{\mathbf{W}}^z \left(\boldsymbol{\Lambda}_n^\psi \right)^{-1/2} \left(\mathbf{P}_{\bar{n}}^{\psi'} \mathbf{P}_{\bar{n}}^\psi \right)^{-1} \mathbf{P}_{\bar{n}}^{\psi'} \boldsymbol{\phi}_{\bar{n}t} + O_{\mathbb{P}}(\zeta_{nT}), \end{aligned} \quad (90)$$

and

$$\begin{aligned} \check{\mathbf{u}}'_t \widehat{\mathbf{R}}_i &= \mathbf{u}'_t \mathbf{R}_i + \mathbf{u}'_t (\boldsymbol{\Sigma}^R)^{1/2} \widehat{\mathbf{W}}^z \left(\widehat{\mathbf{R}}_i - \widehat{\mathbf{W}}^z (\boldsymbol{\Sigma}^R)^{1/2} \mathbf{R}_i \right) + \left(\check{\mathbf{u}}'_t - \mathbf{u}'_t (\boldsymbol{\Sigma}^R)^{1/2} \widehat{\mathbf{W}}^z \right) \widehat{\mathbf{R}}_i \\ &= \mathbf{u}'_t \mathbf{R}_i + \mathbf{u}'_t (\boldsymbol{\Sigma}^R)^{1/2} \widehat{\mathbf{W}}^z \left(\mathbf{L}_T^\psi \right)^{-1/2} \left(\boldsymbol{\Pi}_{\bar{T}}^{\psi'} \boldsymbol{\Pi}_{\bar{T}}^\psi \right)^{-1} \boldsymbol{\Pi}_{\bar{T}}^{\psi'} \boldsymbol{\varphi}_{\bar{T}}^i + \left(\check{\mathbf{u}}'_t - \mathbf{u}'_t (\boldsymbol{\Sigma}^R)^{1/2} \widehat{\mathbf{W}}^z \right) \widehat{\mathbf{R}}_i + O_{\mathbb{P}}(\zeta_{nT}) \\ &= \mathbf{u}'_t \mathbf{R}_i + \mathbf{u}'_t (\boldsymbol{\Sigma}^R)^{1/2} \widehat{\mathbf{W}}^z \left(\mathbf{L}_T^\psi \right)^{-1/2} \left(\boldsymbol{\Pi}_{\bar{T}}^{\psi'} \boldsymbol{\Pi}_{\bar{T}}^\psi \right)^{-1} \boldsymbol{\Pi}_{\bar{T}}^{\psi'} \boldsymbol{\varphi}_{\bar{T}}^i + O_{\mathbb{P}}(\zeta_{nT}) \end{aligned} \quad (91)$$

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

From Theorem 1, (90), and because of (69), as $n, T \rightarrow \infty$,

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

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

Likewise, from Theorem 2, (91), and because of (69), as $n, T \rightarrow \infty$,

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

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

Moreover, because of Assumption (G), when $\bar{n} = \bar{T} = \bar{h}$,

$$\sqrt{\frac{\bar{n}}{n}} \mathbf{P}_{\bar{n}}^{\psi'} \boldsymbol{\phi}_{\bar{n}t} + \sqrt{\frac{\bar{T}}{T}} \boldsymbol{\Pi}_{\bar{T}}^{\psi'} \boldsymbol{\varphi}_{\bar{T}}^i \rightarrow_d \mathcal{N} \left(\mathbf{0}_q, \boldsymbol{\mathcal{P}}_t^u + \boldsymbol{\mathcal{P}}_i^R + \boldsymbol{\Omega}_{it} + \boldsymbol{\Omega}'_{it} \right) \quad \text{as } n, T \rightarrow \infty,$$

where

$$\Omega_{it} := \lim_{n,T \rightarrow \infty} \left(\frac{\sqrt{nT}}{\bar{h}} \right) \mathbb{E} \left[\mathbf{P}_{\bar{n}}^{\psi'} \phi_{\bar{n}t} \varphi_T^i \mathbf{\Pi}_T^{\psi} \right].$$

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

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

where $\omega = \lim_{n,T \rightarrow \infty} \omega_{nT}$ and

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

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 $t = 1, \dots, T$,

$$\begin{aligned} \widehat{\chi}_{it} - \chi_{it} &= \mathbf{e}_i' \left(\widehat{\mathbf{C}}_n(L) \widehat{\boldsymbol{\psi}}_{nt} - \mathbf{C}_n(L) \boldsymbol{\psi}_{nt} \right) \\ &= \mathbf{e}_i' \left(\widehat{\mathbf{C}}_n(L) - \mathbf{C}_n(L) \right) \boldsymbol{\psi}_{nt} + \mathbf{e}_i' \mathbf{C}_n(L) \left(\widehat{\boldsymbol{\psi}}_{nt} - \boldsymbol{\psi}_{nt} \right) + \mathbf{e}_i' \left(\widehat{\mathbf{C}}_n(L) - \mathbf{C}_n(L) \right) \left(\widehat{\boldsymbol{\psi}}_{nt} - \boldsymbol{\psi}_{nt} \right) \\ &= I + II + III, \quad \text{say,} \end{aligned} \quad (92)$$

where \mathbf{e}_i denotes the i th 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 $n, T \rightarrow \infty$,

$$\max_{\ell=1, \dots, p_s} \max_{j_s, h_s=1, \dots, (q+1)} (\widehat{a}_{j_s, h_s, \ell} - a_{j_s, h_s, \ell})^2 \leq \left\| \widehat{\mathbf{A}}^{[s]} - \mathbf{A}^{[s]} \right\|^2 = O_{\mathbb{P}}(\zeta_{n,T}^2), \quad (93)$$

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 s th $(q+1)$ -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 j th column of block s of \mathbf{A}_n^k and $\widehat{\mathbf{A}}_n^k$).

Assumption (B-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

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

Hence, we can select K such that

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

Then, the norm of I is such that

$$\begin{aligned}
\left| \mathbf{e}'_i \left(\widehat{\mathbf{C}}_n(L) - \mathbf{C}_n(L) \right) \boldsymbol{\psi}_{nt} \right| &\leq \sum_{k=0}^K \left(\sum_{j_s=1}^{q+1} (\widehat{c}_{i,j_s,k} - c_{i,j_s,k})^2 \psi_{j_s,t-k}^2 \right)^{1/2} + o_{\mathbb{P}}(\zeta_{nT}) \\
&\leq \sum_{k=0}^K \left(\sum_{j_s=1}^{q+1} \psi_{j_s,t-k}^4 \right)^{1/4} \left(\sum_{j_s=1}^{q+1} (\widehat{c}_{i,j_s,k} - c_{i,j_s,k})^4 \right)^{1/4} + o_{\mathbb{P}}(\zeta_{nT}) \\
&= O_{\mathbb{P}}(\zeta_{nT}),
\end{aligned} \tag{94}$$

because of (93) and the continuous mapping theorem.

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

$$\left| \sum_{j_s=1}^{q+1} \sum_{k=K+1}^{\infty} c_{i,j_s,k} \left(\widehat{\psi}_{j_s,t-k} - \psi_{j_s,t-k} \right) \right| = o_{\mathbb{P}}(\zeta_{nT}) \tag{95}$$

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

$$\begin{aligned}
\left| \mathbf{e}'_i \mathbf{C}_n(L) \left(\widehat{\boldsymbol{\psi}}_{nt} - \boldsymbol{\psi}_{nt} \right) \right| &\leq \sum_{k=0}^K \left(\sum_{j_s=1}^{q+1} c_{i,j_s,k}^2 \left(\widehat{\psi}_{j_s,t-k} - \psi_{j_s,t-k} \right)^2 \right)^{1/2} + o_{\mathbb{P}}(\zeta_{nT}) \\
&\leq \sum_{k=0}^K \left(\sum_{j_s=1}^{q+1} c_{i,j_s,k}^4 \right)^{1/4} \left(\sum_{j_s=1}^{q+1} \left(\widehat{\psi}_{j_s,t-k} - \psi_{j_s,t-k} \right)^4 \right)^{1/4} + o_{\mathbb{P}}(\zeta_{nT}) \\
&= O_{\mathbb{P}} \left(\max \left(\frac{1}{\sqrt{\bar{h}}}, \zeta_{nT} \right) \right) + o_{\mathbb{P}}(\zeta_{nT}).
\end{aligned} \tag{96}$$

Obviously $III = o_{\mathbb{P}}(\zeta_{nT})$. Therefore, substituting (94) and (96) into (92), we proved consistency.

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

$$\sqrt{\bar{n}} \left(\begin{pmatrix} \widehat{\mathbf{u}}_t \\ \vdots \\ \widehat{\mathbf{u}}_{t-k} \end{pmatrix} - \begin{pmatrix} \mathbf{u}_t \\ \vdots \\ \mathbf{u}_{t-k} \end{pmatrix} \right) \rightarrow_d \mathcal{N}(\mathbf{0}_{qk}, \boldsymbol{\mathcal{V}}_{t\dots t-k}^u), \tag{97}$$

where

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

and

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

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

$$\sqrt{\bar{T}} \left(\begin{pmatrix} \widehat{\mathbf{R}}_{i_1} \\ \vdots \\ \widehat{\mathbf{R}}_{i_\ell} \end{pmatrix} - \begin{pmatrix} \mathbf{R}_{i_1} \\ \vdots \\ \mathbf{R}_{i_\ell} \end{pmatrix} \right) \rightarrow_d \mathcal{N}(\mathbf{0}_{q\ell}, \boldsymbol{\mathcal{V}}_{i_1\dots i_\ell}^R), \tag{98}$$

where

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

and

$$\mathcal{P}_{i_1 \dots i_\ell}^R = \lim_{T \rightarrow \infty} \frac{T}{T} \mathbb{E} \left[\left\{ \mathbf{I}_\ell \otimes \mathbf{\Pi}_T^{\psi'} \right\} \begin{pmatrix} \varphi_T^{i_1} \\ \vdots \\ \varphi_T^{i_\ell} \end{pmatrix} \begin{pmatrix} \varphi_T^{i_1} \\ \vdots \\ \varphi_T^{i_\ell} \end{pmatrix}' \left\{ \mathbf{I}_\ell \otimes \mathbf{\Pi}_T^{\psi'} \right\}' \right].$$

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

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

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

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

where $\boldsymbol{\nu}_q$ is a q -dimensional vector of ones. For given $i = 1, \dots, n$, let $\mathbf{R}'_{i_{j_s}}$ be the row of \mathbf{R}_n corresponding to the j_s th series in block s , which is the block to which series i belongs. Then, from (92), (97), (98), and given K as defined in (95), for any $i = 1, \dots, n$ and $t = 1, \dots, T$, as $n, T \rightarrow \infty$,

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

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

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

with

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

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

E A comment on exchangeability

The duality of the static representations (26) and (27) can be motivated if we assume that 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: (Step I) the stochastic selection, via some unspecified distribution \mathbb{P} , of the distributional features of \mathbf{x} as a time-indexed stochastic process, followed by (Step II) 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 $T \rightarrow \infty$.

A rigorous description of Step I can be given as follows. Consider the space

$$\mathcal{F}_T := \{ \mathbb{P} : \mathbb{P} \text{ a probability measure over } (\mathbb{R}^T, \mathcal{B}^T) \}$$

of all probability measures over the space $(\mathbb{R}^T, \mathcal{B}^T)$ of realizations of length T of scalar-valued processes. Equip that space with the σ -field \mathcal{A}_T of all Borel (with respect to the topology of weak convergence) sets. As usual, denote by $(\Omega, \mathcal{A}_\Omega)$ some adequate probability space; distributions over $(\Omega, \mathcal{A}_\Omega)$ will be denoted as \mathbb{P} . Let $\underline{\mathbb{P}}^{(n)}$, with $\underline{\mathbb{P}}^{(n)}(\omega) = (\mathbb{P}_1^{(n)}(\omega), \dots, \mathbb{P}_n^{(n)}(\omega))$, denote a measurable map from $(\Omega, \mathcal{A}_\Omega)$ to $(\mathcal{F}_T^{\otimes n}, \mathcal{A}_T^{\otimes n})$, where $\mathcal{F}_T^{\otimes n}$ is the n -fold product space $\mathcal{F}_T \times \dots \times \mathcal{F}_T$ and $\mathcal{A}_T^{\otimes n}$ the corresponding product σ -field. The map $\underline{\mathbb{P}}^{(n)}$ is called *exchangeable* if

$$\mathbb{P} \left[(\mathbb{P}_{\pi_n(1)}^{(n)}(\omega), \dots, \mathbb{P}_{\pi_n(n)}^{(n)}(\omega)) \in A_T^{(n)} \right] = \mathbb{P} \left[(\mathbb{P}_1^{(n)}(\omega), \dots, \mathbb{P}_n^{(n)}(\omega)) \in A_T^{(n)} \right]$$

for any permutation π_n of $\{1, \dots, n\}$ and any $A_T^{(n)} \in \mathcal{A}_T^{\otimes n}$.

The distribution \mathbb{P} , however, plays the role of a nuisance. Statistical practice in such cases consists in conducting inference on the realization observed in Step II conditional on the (unobserved) result of Step I (see, e.g., Chapter 10 of Lehmann and Romano (2006)), so that \mathbb{P} needs no further description. Under such conditional approach, the distributional features of the stochastic process of which the

observed panel \mathbf{X}_{nT} is a finite realization are treated as unknown but fixed, which is precisely what the deterministic approach is doing. An important feature of Step I, however, is that its result should be a cross-sectionally *exchangeable* process: the distributions of any of the resulting $n \times T$ subprocesses thus 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 motivates our Assumption (C-*a*) and provides the intuitive justification for it.

This random cross-section approach is the one we are implicitly adopting in the paper; and it yields, \mathbb{P} -a.s. conditionally on Step I the same results as the deterministic approach based on Assumption (A) only. The main benefit of Assumption (C-*a*), thus, is to provide a justification for Assumptions (C-*b*) and (C-*d-ii*) which determine the special form of \mathbf{G}_T^ϕ and the linear divergence of exploding eigenvalues, which otherwise would be “brutally” imposed.