TI 2025-016/III
Tinbergen Institute Discussion Paper

New rank-based tests and estimators for Common Primitive Shocks

Federico Carlini[1]
Mirco Rubin[2]
Pierluigi Vallarino[3]

1 LUISS Business School
2 EDHEC Business School
3 Erasmus University Rotterdam, Tinbergen Institute

New rank-based tests and estimators for Common Primitive Shocks *

Federico Carlini† Mirco Rubin‡ Pierluigi Vallarino §

This version: February 28, 2025

Abstract

We propose a new rank-based test for the number of common primitive shocks, q, in large panel data. After estimating a VAR(1) model on r static factors extracted by principal component analysis, we estimate the number of common primitive shocks by testing the rank of the VAR residuals' covariance matrix. The new test is based on the asymptotic distribution of the sum of the smallest $r - q$ eigenvalues of the residuals' covariance matrix. We develop both plug-in and bootstrap versions of this eigenvalue-based test. The eigenvectors associated to the q largest eigenvalues allow us to construct an easy-to-implement estimator of the common primitive shocks. We illustrate our testing and estimation procedures with applications to panels of macroeconomic variables and individual stocks' volatilities.

Keywords: Common primitive shocks; Dynamic factor models; Rank tests; Distribution of eigenvalues; Principal component analysis.

*We thank Matteo Barigozzi, Peter Boswijk, Valentina Corradi, Patrick Gagliardini, Silvia Gonçalves, Manfred Deistler, Alex Horenstein, Frank Kleibergen, Oliver Linton, Marco Lippi, Hao Ma, Serena Ng, Aleksey Onatski, Richard Paap, Richard Smith, Lorenzo Trapani, and Paolo Zaffaroni for their useful comments. In addition, we thank participants at the 27^{th} International Panel Data Conference in Bertinoro (Italy), the Fifth Workshop on High-Dimensional Times Series in Macroeconomics and Finance, hosted at the Institute for Advanced Studies in Vienna, the 2022 EC^2 conference at ESSEC Paris, the 2022 Rome-Waseda Time Series Symposym in Rome, the 2023 Asian Meeting of the Econometric Society, the 2023 Florence-Paris workshop on Statistics of Random Processes and its Applications to Financial Econometrics, the 2024 European Summer Meeting of the Econometric Society at Erasmus University Rotterdam, the 2024 CIREQ-CMP Econometrics Conference in Honor of Eric Ghysels in Montreal, the 2024 FinML conference at USI Lugano, and seminar participants at CREST-ENSAE Paris, and the Universities of Bern, Cambridge, Rotterdam (Erasmus), Queen Mary in London, and Verona.

†LUISS Business School, Rome, Italy (fcarlini@luiss.it).

‡EDHEC Business School, Nice, France (mirco.rubin@edhec.edu), corresponding author.

§Econometrics Institute, Erasmus School of Economics, Rotterdam, The Netherlands (vallarino@ese.eur.nl).

1 Introduction

We introduce a rank-based test to determine the number of common primitive shocks in large panel data. That is, we study the problem of testing and estimating the smallest number q of common sources of randomness characterizing the N-dimensional vector of observable variables $y_t = [y_{1,t}, \ldots, y_{N,t}]'$, for $t = 1, ..., T$, in a linear dynamic factor model.

As a motivating example of a Data Generating Process (DGP) where a large dimensional dynamic system is driven by $q \ll N$ primitive shocks, let us consider *restricted* dynamic factor models of the kind:

$$
\begin{aligned}
y_t &= \Lambda_0 h_t + \cdots + \Lambda_s h_{t-s} + \varepsilon_t, && (1.1)\\
h_t &= \Phi_1 h_{t-1} + \cdots \Phi_p h_{t-p} + w_t, && (1.2)
\end{aligned}
$$

where y_t depends on contemporaneous h_t and lagged $(h_{t-1}, \ldots, h_{t-s})$ values of q latent common *dynamic factors* through the (N, q) matrices $\Lambda_0, \Lambda_1, \ldots, \Lambda_s$. Vector h_t evolves according to the stationary VAR(p) in equation (1.2), where we assume $p \leq s < \infty$ to simplify the exposition.[1] The N-dimensional vector ε_t comprises weakly correlated (over time and across entities) zero-mean innovations, while w_t is a non-degenerate q-dimensional vector of mutually orthogonal shocks. As p and s are finite, we can introduce a new r-dimensional vector of factors $f_t = [h_t', h_{t-1}', \ldots, h_{t-s}']'$, with $r = q(s + 1)$, a new matrix of loadings $\Lambda = [\Lambda_0, ..., \Lambda_s]$, and write the *static factor* representation of the model in (1.1)-(1.2) as:

$$
\begin{aligned}
y_t &= \Lambda f_t + \varepsilon_t, && (1.3)\\
f_t &= \Phi f_{t-1} + v_t, && (1.4)
\end{aligned}
$$

where f_t follows a stationary VAR(1) with auto-regressive matrix Φ based on matrices $\Phi_1, \ldots, \Phi_p$ in (1.2), and innovations $v_t = \begin{bmatrix} I_q & 0 & \ldots & 0 \end{bmatrix}' w_t$ whose (r, r) covariance matrix $V(v_t)$ has rank $q < r$, i.e. it is rank deficient so that its smallest $r - q$ eigenvalues are zero. Thus, the common component of the factor model in (1.3)-(1.4) is based on the r-dimensional random vector f_t whose stochasticity is entirely driven by the a smaller number q of common *primitive shocks* w_t. In the rest of the paper

[1] The case $p > s$ would require a heavier notation.

we focus on all the models which can be written as (1.3)-(1.4) such that $V(v_t)$ is reduced rank, with model (1.1)-(1.2) being a special case.

Dynamic factor models based on (1.1)-(1.2), and their generalizations, have been extensively used in macroeconometrics following the seminal works of Forni and Reichlin (1998) and Forni, Hallin, Lippi, and Reichlin (2000).[2] For restricted dynamic factor models which can be written as (1.3)-(1.4), estimation and inference on w_t, and on its dimension q, have been studied by Bai and Ng (2007), Amengual and Watson (2007), and Breitung and Pigorsch (2013), who start by estimating the r static factors f_t through Principal Component Analysis (PCA) on the observed panel.[3]

We contribute to the literature on large dynamic factor models by introducing a sequential testing procedure for the rank of $V(v_t)$, which yields a consistent estimator for the number of common shocks q. This new procedure relies on testing the rank of $V(v_t)$ after having estimated a VAR(1) model on the Principal Components (PCs) estimate $\hat{f}_t$ of the r static factors f_t. In particular, the test is based on the asymptotic distribution of the sum of the smallest $r-q$ eigenvalues of the sample covariance matrix $\hat{V}(v_t)$. These are all equal to zero in population, but are strictly positive in finite samples due to the estimation error of factors $\hat{f}_t$, which induces an estimation error in $\hat{V}(v_t)$ and its eigenvalues. We propose two implementations of the test: a first one based on plug-in estimators of the asymptotic bias and the variance, and a second one relying on a residual (wild) bootstrap of the standardized test statistics. We also develop an easy-to-implement estimator for the common shocks w_t based on the eigenvectors associated to the largest q eigenvalues of $\hat{V}(v_t)$. Notably, the derivation of the asymptotic distribution of the smallest $r-q$ eigenvalues of the positive semi-definite matrix (p.s.d.) $\hat{V}(v_t)$, discussed more in detail below, is the main theoretical contribution of our work.

Consistent estimation procedures for q based on Information Criteria (IC) have been derived by Amengual and Watson (2007), Bai and Ng (2007), and Breitung and Pigorsch (2013). Differently from these approaches, we construct a consistent estimator for q starting from an eigenvalue-based rank test procedure. This is similar to the approach of Onatski (2009), who tests and estimates the number of common primitive shocks (or, equivalently, of dynamic factors) for generalized dynamic

[2] Generalized dynamic factor models, introduced by Forni, Hallin, Lippi, and Reichlin (2000), extend the model (1.1) - (1.2) to account for an infinite number of lags in the r.h.s. of (1.1) and/or (1.2), e.g. when h_t follows an invertible MA(q) process. See Barigozzi and Hallin (2024) for a review of generalized dynamic factor models and their relation with restricted ones. Presently, our methodology is restricted to dynamic factor models with a finite number of lags s and p.

[3] Other estimation techniques based on the Expectation Maximization algorithm are surveyed by Barigozzi and Luciani (2019), who also establish the asymptotic properties of the resulting factor and loading estimators.

factor models by exploiting the asymptotic distribution of functions of the eigenvalues of the estimated spectral density matrix of the observables $y_{i,t}$.[4] Under the same setting, Hallin and Liska (2007) derive a consistent selection procedure for the number of common dynamic factors/primitive shocks. Kapetanios (2010) proposes an alternative to the test of Onatski (2009) based on the largest eigenvalues of the covariance matrix of the data. Importantly, our testing procedure allows for more general relative convergence rates of N and T compared those required for the asymptotic results in Onatski (2009) and Kapetanios (2010).

As discussed by Bai and Ng (2007) and Donald, Fortuna, and Pipiras (2010), testing the rank of a finite-dimensional p.s.d. matrix is a highly non-standard problem. While the literature has developed plenty of methods to test the rank of a matrix, e.g. Gill and Lewbel (1992), Cragg and Donald (1996), Robin and Smith (2000), Kleibergen and Paap (2006) and Donald, Fortuna, and Pipiras (2007), these approaches cannot deal with matrices that are symmetric and semi-definite. Indeed, Donald, Fortuna, and Pipiras (2007) showed that when the rank of a semi-definite matrix, say M_0, is estimated using a semi-definite matrix, say $\hat{M}$, the asymptotic variance-covariance matrix of the estimator, say $W_0 = V(vech\{\hat{M}\})$, is necessarily singular. Hence, the aforementioned rank tests for indefinite matrices cannot be applied as they assume that W_0 is full rank. To the best of our knowledge, we are the first to solve the problem of testing the rank of a finite dimensional positive semi-definite (p.s.d.) matrix in panel data where both N and T diverge.[5]

Our solution consists in a sequential testing procedure based on the asymptotic distribution of the sum of the smallest $r - q$ eigenvalues of the estimated covariance matrix $\hat{V}(v_t)$. This distribution is derived through an asymptotic expansion of the PC estimator $\hat{f}_t$ and by applying perturbation methods to expand the sum of the smallest $r - q$ estimated eigenvalues.[6] Remarkably, under the null hypothesis of $V(v_t)$ having rank q, the asymptotic distribution of the test statistic is Gaussian, has a non-standard convergence rate $N\sqrt{T}$, and features an asymptotic bias of order $1/N$. This bias is due to the measurement error in the eigenvalues of $\hat{V}(v_t)$ originated by the estimation error of the static factors. Starting from this asymptotic distribution, we develop a consistent sequential testing procedure to determine

[4] The testing procedure of Onatski (2009) requires the use of frequency-domain techniques, as the test is developed for generalized dynamic factor models. Our testing and estimation procedures do not require frequency-domain techniques.

[5] Fortin, Gagliardini, and Scaillet (2023a,b) develop rank-based testing procedures for the number of static factors when N diverges but T is finite, under a more restrictive set of assumptions.

[6] Perturbation arguments are also used by Andreou, Gagliardini, Ghysels, and Rubin (2019), AGGR henceforth, to derive the asymptotic distribution of the sum of the largest canonical correlations of two sets of factors estimated by PCA from two separate panels to test for the presence of common factors among them.

the rank of $V(v_t)$ which takes into account the usual issues related to multiple testing in the spirit of, e.g., Robin and Smith (2000). To improve finite sample properties of the test, we develop a residual-based wild bootstrap inspired by the works of Gonçalves and Perron (2014, 2020), and Gonçalves, Koh, and Perron (2024).

Empirical sizes and powers of the two implementations are studied in a Monte Carlo (MC) analysis. The asymptotic test is over-sized when N is small, e.g. $N < 200$, but it has unitary empirical even when controlling for the size distortion. The bootstrap scheme refines the actual size while preserving most of its power. Both implementations return estimators of q that improve upon extant approaches. Finally, we use the new methodology in two separate empirical applications to study the factor structure of US macroeconomic variables and of volatility measures on US stocks.

The rest of the paper is organized as follows. Section 2 details the modelling setting and the identification strategy for the common primitive shocks and their number q. Section 3 describes estimators of the static factors, the primitive shocks, and their number. The large sample theory of the test and estimators are presented in Section 4. Section 5 introduces the bootstrap implementation of the test and related sequential estimation procedure. Section 6 shows results of different MC studies, while Section 7 covers the empirical applications. Section 8 concludes. Appendix A provides the regularity conditions, while Appendix B includes the proofs of Propositions and Theorems. The Online Appendix (OA) provides the proofs of additional technical results (Section C), details of the bootstrap implementation of the tests based either on the smallest $r - q$ eigenvalues (Section D), a detailed discussion of the alternative estimators of q considered in the MC experiments (Section E), and additional MC analyses (Section F).[7]

[7] Regarding notation, we partition an r-dimensional vector x_t as $x_t = [x'_{Ht}, x'_{Lt}]'$, where x_{Ht} is the upper q-dimensional sub-vector, and x_{Ht} is the lower $(r-q)$-dimensional sub-vector. Moreover, we partition any (r, r) matrix A in four blocks as:

$$A = \begin{bmatrix} A_{HH} & A_{HL} \\ A_{LH} & A_{LL} \end{bmatrix},$$

where A_{HH} is the upper-left (q, q) block, A_{HL} the upper-right $(q, r-q)$ block, A_{LH} is the bottom-left $(r-q, q)$ block and A_{LL} the bottom-right $(r-q, r-q)$ block.

2 The model

In the rest of the paper, we consider the following factor model:

$$
\begin{aligned}
y_t &= \breve{\Lambda}\breve{f}_t + \varepsilon_t &(2.1)\\
\breve{f}_t &= \breve{\Phi}\breve{f}_{t-1} + \breve{v}_t, &(2.2)\\
\breve{v}_t &= G\eta_t, \qquad \text{with} \qquad \eta_t \sim iid(0, I_q) &(2.3)
\end{aligned}
$$

where $y_t = [y_{1t}, ..., y_{Nt}]'$ is a vector of observations for N individuals at time $t = 1, \dots, T$, $\breve{\Lambda} = [\breve{\lambda}_1, ..., \breve{\lambda}_N]'$ is the (N, r) matrix of factor loadings, $\breve{f}_t$ is the r-dimensional vector of latent static factors with $1 < r \ll \min(N, T)$, and $\varepsilon_t = [\varepsilon_{1t}, ..., \varepsilon_{Nt}]'$ is an N-dimensional vector of weakly correlated (over time and across entities) error terms. Factors $\breve{f}_t$ follow a stationary VAR(1) process with autoregressive matrix $\breve{\Phi}$. The r-dimensional innovations vector $\breve{v}_t$ is a linear combination of the q-dimensional vector of "primitive shocks" η_t, with $1 \leq q \leq r$. These primitive shocks are orthogonal and independent over time. The (r, q) full-column rank matrix G maps the primitive shocks η_t into $\breve{v}_t$, so that

$$
\breve{v}_t \sim iid(0, \breve{\Sigma}_v), \tag{2.4}
$$

for $\breve{\Sigma}_v := E(\breve{v}_t\breve{v}_t') = GI_qG' = GG'$, and we denote by σ_ℓ^2 its ℓ-th largest eigenvalue, where $\ell = 1, ..., r$.[8] When $q < r$ the (r, r) matrix $\breve{\Sigma}_v$ has reduced rank q, so that its smallest $r - q$ (largest r) eigenvalues are zero (positive), i.e. $\sigma_1^2 \geq \sigma_2^2 \geq ... \geq \sigma_q^2 > \sigma_{q+1}^2 = \sigma_{q+2}^2 = ... = \sigma_r^2 = 0$.

2.1 Identification of static factors and primitive shocks

Our identification strategy for the number q of primitive shocks is based on the number $r - q$ of zero eigenvalues of $\breve{\Sigma}_v$ in (2.4). Similarly, the actual common shocks are identified by using the eigenvectors of $\breve{\Sigma}_v$ associated to its largest q eigenvalues.

Static factors $\breve{f}_t$ are identifiable, up to a rotation (and change of sign), by performing PCA on the

[8] The assumption $V(\eta_t) = I_q$ is an identification condition. Different values of variance and (non-perfect) correlation among the innovations $\breve{v}_t$ can be obtained by appropriate values of the entries of matrix G.

panel of observables y_t, under the standard identification assumptions:

$$E(\breve{f}_t) = 0 \qquad \text{and} \qquad V(\breve{f}_t) = E(\breve{f}_t \breve{f}_t') = I_r. \tag{2.5}$$

The zero mean assumption is implied by not including an intercept in the r.h.s. of (2.2). We refer to (2.5) as Assumption A.2 ii) in the list of regularity conditions in Appendix A.

When $\breve{\Sigma}_v$ has rank $q < r$, there exists an equivalent way of expressing the DGP in (2.1) - (2.4) which allows identification of the common primitive shocks η_t while simplifying the derivation of the distribution of the test statistics for q. Let Σ_v be the (r, r) diagonal matrix of the sorted eigenvalues of $\breve{\Sigma}_v$, that is

$$\Sigma_v := diag(\sigma_1^2, ..., \sigma_q^2, 0, ..., 0), \tag{2.6}$$

and let $W_v = [W_{v,q} \ , \ W_{v,r-q}]$ be the (r, r) matrix containing the associated orthonormal eigenvectors, with $W_{v,q}$ being the (r, q) matrix of eigenvectors associated to the largest q eigenvalues, and $W_{v,r-q}$ being the $(r, r-q)$ matrix of eigenvectors associated to the $r - q$ zero eigenvalues. Then,

$$\breve{\Sigma}_v W_v = W_v \Sigma_v \ , \qquad \text{with} \qquad W_v' W_v = W_v W_v' = I_r. \tag{2.7}$$

Let us define the rotated factors and associated loadings as

$$f_t = \left[f_{H,t}', \ \ f_{L,t}'\right]' := W_v' \breve{f}_t, \ \ t = 1, ..., T \qquad \text{and} \qquad \Lambda = [\lambda_1, ..., \lambda_N]' := \breve{\Lambda} W_v \tag{2.8}$$

with $\lambda_i = W_v' \breve{\lambda}_i$ for $i = 1, ..., N$. Then, equation (2.1) can be written as:

$$y_t = \Lambda f_t + \varepsilon_t. \tag{2.9}$$

By defining $\Phi := W_v' \breve{\Phi} W_v$, $v_t := W_v' \breve{v}_t = [(W_{v,q}' G \eta_t)' \ \ (W_{v,r-q}' G \eta_t)']'$, and by premultiplying both sides of equation (2.2) by W_v' we obtain $W_v' \breve{f}_t = W_v' \breve{\Phi} W_v W_v' \breve{f}_{t-1} + W_v' \breve{v}_t$, which is an equivalent DGP for the rotated factors:

$$f_t \ = \ \Phi f_{t-1} + v_t, \qquad \text{with} \qquad v_t \sim iid(0, \Sigma_v) \tag{2.10}$$

and $\Sigma_v := V(v_t) = W_v'\breve{\Sigma}_v W_v$. From equations (2.4) and (2.6), the $(r-q)$-dimensional sub-vector v_{Lt} of v_t is degenerate with $E[v_{Lt}] = 0$ and $V(v_{Lt}) = \Sigma_{v,LL} = 0_{(r-q,r-q)}$, so that

$$v_t = \begin{bmatrix} v_{Ht} \\ v_{Lt} \end{bmatrix} = \begin{bmatrix} v_{Ht} \\ 0 \end{bmatrix}, \tag{2.11}$$

almost surely, for any t. Hence, equation (2.10) can be re-written as:

$$\begin{bmatrix} f_{Ht} \\ f_{Lt} \end{bmatrix} = \begin{bmatrix} \Phi_{HH} & \Phi_{HL} \\ \Phi_{LH} & \Phi_{LL} \end{bmatrix} \begin{bmatrix} f_{Ht-1} \\ f_{Lt-1} \end{bmatrix} + \begin{bmatrix} v_{Ht} \\ 0 \end{bmatrix}, \tag{2.12}$$

or, by recursive substitution,

$$\begin{bmatrix} f_{Ht} \\ f_{Lt} \end{bmatrix} = \begin{bmatrix} v_{H,t} \\ 0 \end{bmatrix} + \sum_{s=1}^{\infty} \begin{bmatrix} \Phi_{HH} & \Phi_{HL} \\ \Phi_{LH} & \Phi_{LL} \end{bmatrix}^s \begin{bmatrix} v_{H,t-s} \\ 0 \end{bmatrix}.$$

In this DGP, the non-degenerate vector $v_{Ht} = W_{v,q}'G\eta_t$ collects a one-to-one linear transformation of the q primitive shocks η_t given by the full rank (q,q) matrix $W_{v,q}'G$. On the other hand, the $r-q$ factors collected in f_{Lt} have degenerate innovations $v_{Lt} = W_{v,r-q}'G\eta_t = 0$ for any t. Thus, factors f_{Ht} are the only ones to bring most recent information of the q common shocks η_t into the common component Λf_t: for this reason we refer to them as *non-redundant* factors. The remaining factors f_{Lt} are a deterministic function of the common shocks up to and including time $t-1$, and can be seen a linear transformation of lagged factors f_{t-1} or, equivalently, as a function only of past primitive shocks $v_{t-1}, v_{t-2}, ...$ Hence, we refer to them as *redundant* factors in what follows.[9]

3 Estimators

Section 3.1 presents the OLS estimator of $\breve{\Sigma}_v$ (resp. Σ_v) when factors $\breve{f}_t$ (resp. f_t) are observed. In Section 3.2 we introduce estimators for the static factors, the common shocks, VAR(1) parameters,

[9]This discussion implies the existence of $r-q$ different linear combinations of f_t (resp. $\breve{f}_t$) which are perfectly correlated with other $r-q$ linear combinations of f_{t-1} (resp. $\breve{f}_{t-1}$) or, equivalently, that there exist $r-q$ unitary canonical correlations between f_t (resp. $\breve{f}_t$) and f_{t-1} (resp. $\breve{f}_{t-1}$). A similar argument was first used by Breitung and Pigorsch (2013).

and eigenvalues/eigenvectors of $\breve{\Sigma}_v$ when static factors are estimated by PCA. Section 3.3 deals with the sequential testing strategy characterizing our estimator of q.

3.1 Estimation of Σ_v when factors are observed

Let $\tilde{\breve{\Phi}} = (\sum_{t=1}^{T} \breve{f}_t \breve{f}'_{t-1})(\sum_{t=1}^{T} \breve{f}_{t-1} \breve{f}'_{t-1})^{-1}$ be the Ordinary Least Squares (OLS) estimator of $\breve{\Phi}$ when factors $\breve{f}_t$ are observable, and let $\tilde{\breve{v}}_t = \breve{f}_t - \tilde{\breve{\Phi}} \breve{f}_{t-1}$ be the VAR residuals estimated using $\tilde{\breve{\Phi}}$. In this case, the OLS estimator of $\breve{\Sigma}_v$ is:

$$\tilde{\breve{\Sigma}}_v = \frac{1}{T} \sum_{t=1}^{T} \tilde{\breve{v}}_t \tilde{\breve{v}}'_t. \tag{3.1}$$

Moreover, let

$$\tilde{\Phi} := \left(\sum_{t=1}^{T} f_t f'_{t-1} \right) \left(\sum_{t=1}^{T} f_{t-1} f'_{t-1} \right)^{-1} = \begin{bmatrix} \tilde{\Phi}_{HH} & \tilde{\Phi}_{HL} \\ \tilde{\Phi}_{LH} & \tilde{\Phi}_{LL} \end{bmatrix}, \tag{3.2}$$

be the OLS estimator of Φ when factors f_t are observable, and $\tilde{v}_t = f_t - \tilde{\Phi} f_{t-1}$ be the associated VAR innovations. The OLS estimator of Σ_v reads:

$$\tilde{\Sigma}_v = \frac{1}{T} \sum_{t=1}^{T} \tilde{v}_t \tilde{v}'_t. \tag{3.3}$$

Note that both estimators $\tilde{\breve{\Sigma}}_v$ and $\tilde{\Sigma}_v$ are infeasible when factors are unobserved.

3.2 Estimation when factors are unobserved

Let us assume that the true number of static factors r is known, but the true factors $\breve{f}_t$ are unobservable and q is unknown. To determine q, we start by estimating factors $\breve{f}_t$ by PCA. Let $\hat{\breve{F}} = [\hat{\breve{f}}_0, \hat{\breve{f}}_1, ..., \hat{\breve{f}}_T]'$, i.e. the estimator of $\breve{F} = [\breve{f}_0, \breve{f}_1, ..., \breve{f}_T]'$, be the $(T+1, r)$ matrix of estimated Principal Components (PCs) extracted from the $(T+1, N)$ panel $Y = [y_0, y_1, ..., y_T]'$. These PCs are associated to the largest r eigenvalues of matrix $\frac{1}{N(T+1)} YY'$, i.e. $\hat{\breve{F}}$ satisfies the usual PCA eigenvalue-eigenvector equation:

$$\frac{1}{N(T+1)} YY' \hat{\breve{F}} = \hat{\breve{F}} \hat{V},$$

where $\hat{V}$ is the (r,r) diagonal matrix containing the largest r eigenvalues of $YY'/(N(T+1))$ in descending order, and the columns of matrix $\mathring{\hat{F}}$ are the associated normalized eigenvectors such that $\frac{1}{T+1}\mathring{\hat{F}}'\mathring{\hat{F}} = \frac{1}{T+1}\sum_{t=0}^{T}\mathring{\hat{f}}_t\mathring{\hat{f}}_t' = I_r$.[10]

Let $\mathring{\hat{\Phi}} = (\sum_{t=1}^{T}\mathring{\hat{f}}_t\mathring{\hat{f}}_{t-1}')(\sum_{t=1}^{T}\mathring{\hat{f}}_{t-1}\mathring{\hat{f}}_{t-1}')^{-1}$ be the OLS estimator of $\breve{\Phi}$ when factors $\breve{f}_t$ are estimated by PCA, and $\mathring{\hat{v}}_t = \mathring{\hat{f}}_t - \mathring{\hat{\Phi}}\mathring{\hat{f}}_{t-1}$ be the VAR residuals estimated using $\mathring{\hat{\Phi}}$. Then, the OLS estimator of $\breve{\Sigma}_v$ reads:

$$\mathring{\hat{\Sigma}}_v = \frac{1}{T}\sum_{t=1}^{T}\mathring{\hat{v}}_t\mathring{\hat{v}}_t'.$$

Let $\hat{W}_v$ be the (r,r) matrix collecting the eigenvectors associated to the ordered eigenvalues $\hat{\sigma}_\ell^2$, with $\ell = 1,...,r$, of $\mathring{\hat{\Sigma}}_v$:

$$\mathring{\hat{\Sigma}}_v\hat{W}_v = \hat{W}_v\hat{\Sigma}_v\ , \tag{3.4}$$

where $\hat{\Sigma}_v := diag(\hat{\sigma}_1^2,...,\hat{\sigma}_r^2)$ and $\hat{W}_v'\hat{W}_v = \hat{W}_v\hat{W}_v' = I_r$. From the definition of rotated factors in (2.8), we can define the estimator of $f_t = W_v'\breve{f}_t$ as $\hat{f}_t := \hat{W}_v'\mathring{\hat{f}}_t$, and construct the estimator of $F = [f_0, f_1, ..., f_T]'$ as $\hat{F} := [\hat{f}_0, \hat{f}_1, ..., \hat{f}_T]' = \mathring{\hat{F}}\hat{W}_v$. Analogously, we define the estimator $\hat{v}_t := \hat{W}_v'\mathring{\hat{v}}_t$ of $v_t = W_v'\breve{v}_t$. By denoting as $\hat{W}_{v,q}$ (resp. $\hat{W}_{v,r-q}$) the first q (resp. last $r-q$) columns of $\hat{W}_v$, i.e. $\hat{W}_v = [\hat{W}_{v,q},\ \hat{W}_{v,r-q}]$, we can also define a natural estimator of $f_{H,t}$ and $v_{H,t}$.

DEFINITION 1. *The estimator of the non-redundant static factors $f_{H,t}$ is $\hat{f}_{H,t} = \hat{W}_{v,q}'\mathring{\hat{f}}_t$, and the estimator of the q primitive shocks $v_{H,t}$ is $\hat{v}_{H,t} = \hat{W}_{v,q}'\mathring{\hat{v}}_t$, for all $t = 1,...,T$.*

The matrix of loadings Λ in (2.9) is estimated by time-series regressions of y_{it} on $\hat{f}_t$, yielding the estimator $\hat{\Lambda} = [\hat{\lambda}_1, ..., \hat{\lambda}_N]'$:

$$\hat{\Lambda} \ = \ Y'\hat{F}(\hat{F}'\hat{F})^{-1} \ = \ \frac{1}{T+1}Y'\hat{F}, \tag{3.5}$$

where the second equality follows from $\hat{F}'\hat{F}/(T+1) = \mathring{\hat{F}}'W_vW_v'\mathring{\hat{F}}/(T+1) = \mathring{\hat{F}}'\mathring{\hat{F}}/(T+1) = I_r$.

Let $\hat{\Phi}$ be the OLS estimator of Φ for the VAR(1) in equation (2.10):

$$\hat{\Phi} = \left(\sum_{t=1}^{T}\hat{f}_t\hat{f}_{t-1}'\right)\left(\sum_{t=1}^{T}\hat{f}_{t-1}\hat{f}_{t-1}'\right)^{-1} = \begin{bmatrix} \hat{\Phi}_{HH} & \hat{\Phi}_{HL} \\ \hat{\Phi}_{LH} & \hat{\Phi}_{LL} \end{bmatrix}, \tag{3.6}$$

[10]Let $\hat{F}^*$ be the orthonormal eigenvectors of $\frac{1}{N(T+1)}YY'$, s.t. $\frac{1}{N(T+1)}YY'\hat{F}^* = \hat{F}^*\hat{V}$ and $\hat{F}^{*\prime}\hat{F}^* = I_r$, then the normalized factor estimator $\mathring{\hat{F}}$ is computed as $\mathring{\hat{F}} = \sqrt{T+1}\cdot\hat{F}^*$.

so that $\hat{v}_t = \hat{f}_t - \hat{\Phi}\hat{f}_{t-1}$ for all $t = 1, ..., T$. Then, equation (3.6) implies that:

$$\frac{1}{T}\sum_{t=1}^{T} \hat{v}_t \hat{v}_t' = \hat{\Sigma}_v, \tag{3.7}$$

where $(1/T)\sum_{t=1}^{T} \hat{v}_t\hat{v}_t'$ is the OLS estimator of Σ_v.

Equations (3.4) and (3.7) imply that the ℓ-th eigenvalue of $\hat{\breve{\Sigma}}_v$ is equal to the element in position (ℓ, ℓ) of $\frac{1}{T}\sum_{t=1}^{T} \hat{v}_t\hat{v}_t'$. Importantly, in Section 4 we show that all estimated eigenvalues $\hat{\sigma}_\ell^2$ are strictly positive w.p.a. 1 when $T \to \infty$ and for all $\ell = 1, ..., r$, as the smallest $r - q$ eigenvalues are functions of the estimation error in the principal component estimator $\hat{\breve{f}}_t$.

3.3 Sequence of tests of hypotheses on the number of primitive shocks

As shown in Section 2, the number of primitive shocks q coincides with the number of non-zero eigenvalues of matrix $\breve{\Sigma}_v$. To develop an estimator for q, we consider the sequence of hypotheses in Table 1, which are expressed in terms of the number of non-zero eigenvalues of $\breve{\Sigma}_v$. The generic hypothesis $H(q)$ corresponds to the presence of q primitive shocks, with $1 \leq q \leq r$, and implies that the $r - q$ smallest eigenvalues of $\breve{\Sigma}_v$ are all equal to zero, while the q largest ones are strictly positive.[11]

Table 1 – Hypotheses on the number of common primitive shocks q

$H(q)$	Eigenvalues of $\breve{\Sigma}_v$
$H(1)$	$\sigma_1^2 > \sigma_2^2 = \sigma_3^2 = ... = \sigma_r^2 = 0$
$H(2)$	$\sigma_1^2 \geq \sigma_2^2 > \sigma_3^2 = ... = \sigma_r^2 = 0$
...	...
$H(q)$	$\sigma_1^2 \geq \sigma_2^2 \geq ... \geq \sigma_q^2 > \sigma_{q+1}^2 = ... = \sigma_r^2 = 0$
...	...
$H(r-1)$	$\sigma_1^2 \geq \sigma_2^2 \geq ... \geq \sigma_{r-1}^2 > \sigma_r^2 = 0$
$H(r)$	$\sigma_1^2 \geq \sigma_2^2 \geq ... \geq \sigma_{r-1}^2 \geq \sigma_r^2 > 0$

[11]Note that the assumption $1 \leq q \leq r$ implies that there exists at least one factor in our model (2.1), and therefore we do not consider the degenerate case $H(0) = \{\sigma_1^2 = \sigma_2^2 = ... = \sigma_r^2 = 0\}$, which corresponds to the absence of any common primitive shock. This degenerate case is easy to detect empirically by applying the usual approaches for the number of static factors mentioned below.

To select the number of primitive shocks q, let us consider the following sequence of tests:

$$H_0 = H(q) \qquad \text{vs.} \qquad H_1 = \bigcup_{q<s\leq r} H(s), \qquad \text{for } q = 1, 2, ..., r-1 \ .$$

Given $q = 1, 2, ..., r - 1$, testing H_0 against H_1 is based on the test statistics

$$\hat{\xi}(q) = \sum_{\ell=1}^{r-q} \hat{\sigma}^2_{r-\ell+1}, \tag{3.8}$$

which corresponds to the sum of the $r - q$ smallest eigenvalues of $\hat{\hat{\Sigma}}_v$. We reject the null $H_0 = H(q)$ when $\hat{\xi}(q)$ is positive and large, corresponding to the case where at least one of the eigenvalues in $\hat{\xi}(q)$ is significantly different from zero. Critical values of the test are obtained from the asymptotic distribution of $\hat{\xi}(q)$ derived in Section 4.2. The number of primitive shocks q is estimated by sequentially applying the test for $H(q)$, starting from $q = 1$ and increasing it if the null $H(q)$ is rejected; as described in Section 4.3, the procedure is stopped at the smallest value of q for which the $H(q)$ is not rejected.

When the true number of static factors r is unknown, the asymptotic distribution and rate of convergence for $\hat{\xi}(q)$ based on a consistent $\hat{r}$ are the same as those based on r.[12] Hence, we derive the asymptotic distribution and convergence rate of the test statistics assuming that r is known, as also done in Amengual and Watson (2007) and Bai and Ng (2007).

4 Large sample theory

In Section 4.1, we show that the OLS estimator of Σ_v has $r - q$ zero eigenvalues when factors $\breve{f}_t$ (or f_t) are observed. That is, for any finite sample of dimension T its smallest $r - q$ eigenvalues have a degenerate distribution centered around 0. This implies that testing for the number of primitive shocks is a degenerate problem when factors are observed without error. Then, in Section 4.2 we derive the large sample distribution of $\hat{\xi}(q)$ and provide an implementation of the test based on plug-in

[12]Examples of consistent estimators of r are those proposed by Bai and Ng (2002a), Alessi, Barigozzi, and Capasso (2010), Onatski (2010), Ahn and Horenstein (2013), and Trapani (2018). As also discussed in AGGR, a word of caution is warranted. It is known that pre-testing generates problems in terms of lack of uniform properties, and we therefore abstract from uniformity in this paper.

estimators of its asymptotic bias and variance. A consistent selection procedure for the number of primitive shocks q is presented in Section 4.3.

4.1 Distribution of eigenvalue estimators when static factors are observed

We first study the eigenvalues of matrices $\tilde{\breve{\Sigma}}_v$ (resp. $\tilde{\Sigma}_v$), obtained by estimating the VAR(1) in (2.2) (resp. (2.10)) by OLS from the T-dimensional sample of observed factors $\breve{f}_t$ (resp. f_t).

PROPOSITION 1. *Let $\breve{f}_t$ (resp. f_t), with $t = 1, ..., T \geq r^2$, be a T-dimensional sample of observations on $\breve{f}_t$ (resp. f_t) as given by* (2.2) *(resp.* (2.10)*), and let $\tilde{\breve{\Sigma}}_v$ (resp. $\tilde{\Sigma}_v$) be the OLS estimator of $\breve{\Sigma}_v$ (resp. Σ_v) defined in* (3.1) *(resp.* (3.3)*) based on this sample. Then:*

(i) $\tilde{\breve{\Sigma}}_v$ and $\tilde{\Sigma}_v$ have the same eigenvalues $\tilde{\sigma}^2_\ell \geq 0$, with $\ell = 1, ..., r$:

$$\tilde{\sigma}^2_1 \geq \tilde{\sigma}^2_2 \geq ... \geq \tilde{\sigma}^2_{q-1} \geq \tilde{\sigma}^2_q \geq \tilde{\sigma}^2_{q+1} = \tilde{\sigma}^2_{q+2} = ... = \tilde{\sigma}^2_r = 0. \tag{4.1}$$

(ii) The smallest $r - q$ (resp. largest q) eigenvalues of $\tilde{\breve{\Sigma}}_v$ and $\tilde{\Sigma}_v$ are equal to (resp. strictly larger than) zero w.p.a. 1, i.e. as $T \to \infty$

$$\tilde{\sigma}^2_1 \geq \tilde{\sigma}^2_2 \geq ... \geq \tilde{\sigma}^2_{q-1} \geq \tilde{\sigma}^2_q > \tilde{\sigma}^2_{q+1} = \tilde{\sigma}^2_{q+2} = ... = \tilde{\sigma}^2_r = 0 \qquad w.p.a.\ 1. \tag{4.2}$$

(iii) Assume that the largest q eigenvalues $\sigma^2_1, \ldots, \sigma^2_q$ of $\breve{\Sigma}_v$ are distinct. Then, the largest q eigenvalues of $\tilde{\breve{\Sigma}}_v$ (resp. $\tilde{\Sigma}_v$) converge in distribution to the largest q eigenvalues of $\breve{\Sigma}_v$ (resp. Σ_v) at rate $\sqrt{T}$, that is $\sqrt{T}(\tilde{\sigma}^2_\ell - \sigma^2_\ell) \xrightarrow{d} N(0, V_{asy}(\tilde{\sigma}^2_\ell))$, as $T \to \infty$, where $V_{asy}(\tilde{\sigma}^2_\ell) = e'_{q,\ell}(e'_{q,\ell} \otimes I_q) \cdot V_1 \cdot (e_{q,\ell} \otimes I_q) e_{q,\ell}$, where $V_1 := E[vec(v_t v'_t - \Sigma_v) \cdot vec(v_t v'_t - \Sigma_v)']$ and $e_{q,\ell}$ is the ℓ-th column of I_q.

Proof: see Appendix B.1.

Proposition 1 implies that when factors $\breve{f}_t$ (or f_t) are observable, there is no need the test for the number of common shocks. Indeed, the (estimated) eigenvalues of $\tilde{\breve{\Sigma}}_v$ or $\tilde{\Sigma}_v$ allow to determine the number of primitive shocks: as the q largest eigenvalues of $\tilde{\breve{\Sigma}}_v$ and $\tilde{\Sigma}_v$ will be strictly positive, while the smallest $r - q$ ones will be exactly zero. Similarly to the other cases discussed in Donald, Fortuna, and Pipiras (2014), this is another situation where testing for the rank of a p.s.d. matrix ($\tilde{\breve{\Sigma}}_v$ in our case) is a degenerate problem as the (asymptotic) variance-covariance matrix of this estimator is necessarily singular.

Notably, Proposition 1 shows that when factors are observed the estimation error of Σ_v affects only the largest q eigenvalues, but not the smallest $r-q$. Importantly, this result refines the claim in Section 2 of Bai and Ng (2007) on the eigenvalues of $\tilde{\tilde{\Sigma}}_v$: we establish that the smallest $r-q$ eigenvalues are exactly equal to 0 for any finite sample size $T \geq r^2$, while they claim that these eigenvalues converge to 0 as $T \to \infty$. In contrast, the next Section 4.2 shows that when factors f_t are estimated by PCA and matrix Σ_v is estimated using the estimated factors instead of the true ones, all its r eigenvalues are strictly larger than 0 (w.p.a. 1) for any finite sample, as they are contaminated by PCs estimation error, and converge to 0 only asymptotically when $N, T \to \infty$.

4.2 Distribution of test statistics when static factors are estimated

We consider the joint asymptotics $N, T \to \infty$ and assume that:

$$\sqrt{T}/N = o(1), \qquad N/T^{3/2} = o(1),$$

which correspond to Assumption A.1 in the list of regularity conditions in Appendix A. To derive the large sample distribution of the test statistic for the number of primitive shocks we deploy the refined asymptotic expansion for the estimated PCs derived by AGGR. This expansion extends results in Bai and Ng (2002a), Stock and Watson (2002), Bai (2003), and Bai and Ng (2006), and is reported for convenience as Proposition B.1 in Appendix B. For $t = 1, \ldots, T$ the estimate $\hat{\hat{f}}_t$ is asymptotically equivalent (see details in Proposition B.1), up to negligible terms, to $\hat{\mathcal{H}} W_v \left(f_t + \frac{1}{\sqrt{N}} u_t \right)$, where $u_t = \left(\frac{1}{N} \sum_{i=1}^N \lambda_i \lambda_i' \right)^{-1} \frac{1}{\sqrt{N}} \sum_{i=1}^N \lambda_i \varepsilon_{i,t}$, and $\hat{\mathcal{H}}$ is a nonsingular stochastic factor rotation matrix.[13] The zero-mean term u_t drives the randomness in factor estimates conditional on factor path.

Let $\tilde{\Sigma}_{u,s}(h|\mathcal{F}_t) = Cov(u_s, u_{s-h}|\mathcal{F}_t)$, with $s \leq t$, be the conditional covariance between u_s and u_{s-h} conditional on the sigma field $\mathcal{F}_t = \sigma(f_\tau, \tau \leq t)$ generated by current and past factor values f_t, i.e.

$$\tilde{\Sigma}_{u,s}(h|\mathcal{F}_t) = \left(\frac{1}{N} \sum_{i=1}^N \lambda_i \lambda_i' \right)^{-1} \frac{1}{N} \sum_{i=1}^N \sum_{\ell=1}^N \lambda_i \lambda_\ell' Cov(\varepsilon_{i,s}, \varepsilon_{\ell,s-h}|\mathcal{F}_t) \left(\frac{1}{N} \sum_{i=1}^N \lambda_i \lambda_i' \right)^{-1},$$

and $\tilde{\Sigma}_{u,s}(-h|\mathcal{F}_t) = \tilde{\Sigma}_{u,s}(h|\mathcal{F}_t)'$, for $h = 0, 1, \ldots$. We set $\tilde{\Sigma}_{u,s} \equiv \tilde{\Sigma}_{u,s}(0|\mathcal{F}_t)$, and define $\Sigma_{u,s}(h|\mathcal{F}_t) =$

[13] Vector u_t depends on sample sizes but, for convenience, we omit the indices N, T. Moreover, as shown in, e.g., Bai and Ng (2002a), $\hat{\mathcal{H}} := \left(\frac{\Lambda'\Lambda}{N} \right)^{-1} \left(\frac{\hat{F}'F}{T+1} \right) \hat{V}^{-1}$, where $\hat{V}$ is the is the (r, r) diagonal matrix containing the r largest eigenvalues of matrix $\frac{1}{N(T+1)} YY'$. See also the proof of Proposition C.1 in the OA of AGGR for further details on $\hat{\mathcal{H}}$.

$\operatorname*{plim}_{N\to\infty} \tilde{\Sigma}_{u,s}(h|\mathcal{F}_t)$ and $\Sigma_\lambda = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^N \lambda_i\lambda_i'$.

THEOREM 1. *Under Assumptions A.1 - A.7, and the null hypothesis $H_0 = H(q)$ of q primitive shocks, we have:*

$$\tilde{\xi}^u(q) := N\sqrt{T}\Omega_U^{-1/2}\left(\hat{\xi}(q) - \frac{1}{N}tr\left\{\tilde{C}^{-1}\tilde{B}_U\right\}\right) \xrightarrow{d} N(0,1), \tag{4.3}$$

with $\Omega_U = 2\sum_{h=-\infty}^{\infty} E\left[tr\left\{\Sigma_{U,t}(h)\Sigma_{U,t}(h)'\right\}\right]$, $\tilde{B}_U = \frac{1}{T}\sum_{t=1}^T \tilde{B}_{U,t}$,

$$\begin{aligned}
\Sigma_{U,t}(h) &:= \Sigma_{u,t,LL}(h|\mathcal{F}_t) - \Phi_{LH}\Sigma_{u,t-1,LH}(h-1|\mathcal{F}_t)' - \Phi_{LL}\Sigma_{u,t-1,LL}(h-1|\mathcal{F}_t)' \\
&\quad -\Sigma_{u,t,LH}(h+1|\mathcal{F}_t)\Phi_{LH}' + \Phi_{LH}\Sigma_{u,t-1,HH}(h|\mathcal{F}_t)\Phi_{LH}' + \Phi_{LL}\Sigma_{u,t-1,LH}(h|\mathcal{F}_t)\Phi_{LH}' \\
&\quad -\Sigma_{u,t,LL}(h+1|\mathcal{F}_t)\Phi_{LL}' + \Phi_{LH}\Sigma_{u,t-1,HL}(h|\mathcal{F}_t)\Phi_{LL}' + \Phi_{LL}\Sigma_{u,t-1,LL}(h|\mathcal{F}_t)\Phi_{LL}', \\
\tilde{B}_{U,t} &:= \tilde{\Sigma}_{u,t,LL}(0|\mathcal{F}_t) - \Phi_{LH}\tilde{\Sigma}_{u,t-1,LH}(-1|\mathcal{F}_t)' - \Phi_{LL}\tilde{\Sigma}_{u,t-1,LL}(-1|\mathcal{F}_t)' \\
&\quad -\tilde{\Sigma}_{u,t,LH}(1|\mathcal{F}_t)\Phi_{LH}' + \Phi_{LH}\tilde{\Sigma}_{u,t-1,HH}(0|\mathcal{F}_t)\Phi_{LH}' + \Phi_{LL}\tilde{\Sigma}_{u,t-1,LH}(0|\mathcal{F}_t)\Phi_{LH}' \\
&\quad -\tilde{\Sigma}_{u,t,LL}(1|\mathcal{F}_t)\Phi_{LL}' + \Phi_{LH}\tilde{\Sigma}_{u,t-1,HL}(0|\mathcal{F}_t)\Phi_{LL}' + \Phi_{LL}\tilde{\Sigma}_{u,t-1,LL}(0|\mathcal{F}_t)\Phi_{LL}', \\
\tilde{C} &= W_{v,r-q}'\hat{\mathcal{H}}^{-1}(\hat{\mathcal{H}}')^{-1}W_{v,r-q}.
\end{aligned}$$

Proof: See Appendix B.2.

The estimation error u_t originating from the PC estimation of the factors, and its lagged value u_{t-1}, determines the asymptotic distribution of the infeasible statistic $\tilde{\xi}^u(q)$. In fact, matrix $\Sigma_{U,t}(h)$ appearing in the variance is the limit (with respect to N) covariance matrix between the $(r-q)$-dimensional vector $U_t = u_{Lt} + \Phi_{LH}u_{Ht-1} + \Phi_{LL}u_{Lt-1}$ and its h-th lag $U_{t-h} = u_{Lt-h} + \Phi_{LH}u_{Ht-h-1} + \Phi_{LL}u_{Lt-h-1}$. Remarkably, even if we are testing for parameters at the boundary of their domain, i.e. the eigenvalues of a positive semi-definite matrix to be zero, a Gaussian asymptotic distribution is obtained because the non-negative test-statistic $\hat{\xi}(q)$ is re-centered by subtracting a strictly (a.s.) positive asymptotic bias term of order N^{-1} which is also a function of the covariance between U_t and U_{t-h}. Importantly, our Theorem 1 implies that the (the sum of the) smallest $r-q$ eigenvalues of $\hat{\hat{\Sigma}}_v$ converge to zero at non-standard rate $N\sqrt{T}$, while the remaining ones converge to their true values at the slower "usual" (for time averages of functions of PC estimates) rate $\min(\sqrt{N}, \sqrt{T})$. The latter result, can be obtained by using our Proposition 1 and the same arguments in Bai and Ng (2007).

To get a feasible distributional result for the statistic $\tilde{\xi}^u(q)$, we need consistent estimators of the bias $tr\{\tilde{C}^{-1}\tilde{B}_U\}$ and variance Ω_U terms in (4.3). To estimate these terms we make the simplifying assumptions that the errors $\varepsilon_{i,t}$ are (i) uncorrelated across individuals i, at all leads and lags, and (ii) a conditionally homoskedastic martingale difference sequence for each individual i, conditional on the factor path, i.e.

$$\begin{aligned} Cov(\varepsilon_{i,t}, \varepsilon_{j,t-h}|\mathcal{F}_t) &= 0, \qquad \text{if} \qquad i \neq j, \\ E[\varepsilon_{i,t}|\{\varepsilon_{i,t-h}\}_{h\geq 1}, \mathcal{F}_t] &= 0, \qquad E[\varepsilon_{i,t}^2|\{\varepsilon_{i,t-h}\}_{h\geq 1}, \mathcal{F}_t] = \gamma_{ii}, \end{aligned}$$

for all i, t, h (see Assumption A.9). These assumptions imply $\tilde{\Sigma}_{u,s}(h|\mathcal{F}_t) = \Sigma_{u,s}(h|\mathcal{F}_t) = 0$ for all $h \neq 0$, and $\tilde{\Sigma}_{u,j\ell}(h|\mathcal{F}_t) \equiv \tilde{\Sigma}_{u,s,j\ell}(h|\mathcal{F}_t)$ for all dates s, i.e. these matrices do not depend on time, for $j, \ell = L, H$. Matrices $\tilde{\Sigma}_u \equiv \tilde{\Sigma}_{u,s}(0|\mathcal{F}_t)$, $\Sigma_u \equiv \Sigma_{u,s}(0|\mathcal{F}_t)$ and $\tilde{\Sigma}_U(h) \equiv \tilde{\Sigma}_{U,t}(h)$ also do not depend on time for any h.[14] Therefore, we have:

$$\begin{aligned} \tilde{B}_U &= \tilde{\Sigma}_{u,LL} + \Phi_{LH}\tilde{\Sigma}_{u,HH}\Phi'_{LH} + \Phi_{LL}\tilde{\Sigma}_{u,LH}\Phi'_{LH} + \Phi_{LH}\tilde{\Sigma}_{u,HL}\Phi'_{LL} + \Phi_{LL}\tilde{\Sigma}_{u,LL}\Phi'_{LL}\ , \qquad (4.4) \\ \Sigma_U(0) &\equiv \Sigma_{U,t}(0) \\ &= \Sigma_{u,LL} + \Phi_{LH}\Sigma_{u,HH}\Phi'_{LH} + \Phi_{LL}\Sigma_{u,LH}\Phi'_{LH} + \Phi_{LH}\Sigma_{u,HL}\Phi'_{LL} + \Phi_{LL}\Sigma_{u,LL}\Phi'_{LL}\ , \\ \Sigma_U(1) &\equiv \Sigma_{U,t}(1) \ = \ -\Phi_{LH}\Sigma'_{u,LH} - \Phi_{LL}\Sigma'_{u,LL}, \\ \Sigma_U(-1) &\equiv \Sigma_{U,t}(-1) \ = \ -\Sigma_{u,LH}\Phi'_{LH} - \Sigma_{u,LL}\Phi'_{LL}, \end{aligned}$$

and $\Sigma_{U,t}(h) \ = \ 0$, for all $h \neq -1, 0, 1$, implying that Ω_U simplifies to

$$\Omega_U \ = \ 2tr\left\{\Sigma_U(0)\Sigma_U(0)' + \Sigma_U(1)\Sigma_U(1)' + \Sigma_U(-1)\Sigma_U(-1)'\right\}. \qquad (4.5)$$

In Theorem 2 below, we replace matrices $\tilde{C}^{-1}\tilde{B}_U$ and Ω_U by consistent estimators, namely $\hat{B}_U$ and $\hat{\Omega}_U$. Importantly, we show that the estimation error for the bias adjustment term $\frac{1}{N}tr\{\tilde{C}^{-1}\tilde{B}_U\}$ is of order $o_p\left(\frac{1}{N\sqrt{T}}\right)$, implying that the asymptotic distribution of the feasible statistic remains standard normal, and has the same convergence rate as the infeasible one of Theorem 1.

[14] If the errors are weakly correlated across series and/or time, consistent estimation of $\tilde{\Sigma}_U$ and Ω_U requires thresholding of estimated cross-sectional covariances and/or HAC-type estimators.

THEOREM 2. *Let*

$$\hat{\Sigma}_u = \left(\frac{1}{N}\hat{\Lambda}'\hat{\Lambda}\right)^{-1}\left(\frac{1}{N}\hat{\Lambda}'\hat{\Gamma}\hat{\Lambda}\right)\left(\frac{1}{N}\hat{\Lambda}'\hat{\Lambda}\right)^{-1} = \begin{bmatrix} \hat{\Sigma}_{u,HH} & \hat{\Sigma}_{u,HL} \\ \hat{\Sigma}_{u,LH} & \hat{\Sigma}_{u,LL} \end{bmatrix}, \tag{4.6}$$

where $\hat{\Lambda}$ are the loadings estimators defined in equation (3.5), $\hat{\Gamma} = diag(\hat{\gamma}_{ii},\ i = 1, ..., N)$ with $\hat{\gamma}_{ii} = \frac{1}{T+1}\sum_{t=0}^{T}\hat{\varepsilon}_{i,t}^2$, and $\hat{\varepsilon}_{i,t} = y_{i,t} - \hat{\lambda}_i'\hat{f}_t$. Let $\hat{\Phi}$ be the estimator of Φ defined in (3.6). *Define also:*

$$\hat{B}_U = \hat{\Sigma}_{u,LL} + \hat{\Phi}_{LH}\hat{\Sigma}_{u,HH}\hat{\Phi}_{LH}' + \hat{\Phi}_{LL}\hat{\Sigma}_{u,LH}\hat{\Phi}_{LH}' + \hat{\Phi}_{LH}\hat{\Sigma}_{u,HL}\hat{\Phi}_{LL}' + \hat{\Phi}_{LL}\hat{\Sigma}_{u,LL}\hat{\Phi}_{LL}' ,$$

and

$$\hat{\Omega}_U = 2tr\left\{\hat{\Sigma}_U(0)\hat{\Sigma}_U(0)' + \hat{\Sigma}_U(1)\hat{\Sigma}_U(1)' + \hat{\Sigma}_U(-1)\hat{\Sigma}_U(-1)'\right\},$$

where

$$\begin{aligned} \hat{\Sigma}_U(0) &= \hat{\Sigma}_{u,LL} + \hat{\Phi}_{LH}\hat{\Sigma}_{u,HH}\hat{\Phi}_{LH}' + \hat{\Phi}_{LL}\hat{\Sigma}_{u,LH}\hat{\Phi}_{LH}' + \hat{\Phi}_{LH}\hat{\Sigma}_{u,HL}\hat{\Phi}_{LL}' + \hat{\Phi}_{LL}\hat{\Sigma}_{u,LL}\hat{\Phi}_{LL}' , \\ \hat{\Sigma}_U(1) &= -\hat{\Phi}_{LH}\hat{\Sigma}_{u,LH}' - \hat{\Phi}_{LL}\hat{\Sigma}_{u,LL}', \\ \hat{\Sigma}_U(-1) &= -\hat{\Sigma}_{u,LH}\hat{\Phi}_{LH}' - \hat{\Sigma}_{u,LL}\hat{\Phi}_{LL}'. \end{aligned}$$

The feasible test statistic is

$$\tilde{\xi}(q) := N\sqrt{T}\hat{\Omega}_U^{-1/2}\left(\hat{\xi}(q) - \frac{1}{N}tr\left\{\hat{B}_U\right\}\right). \tag{4.7}$$

Let Assumptions A.1 - A.9 hold, then: (i) under the null hypothesis $H_0 = H(q)$ of q primitive shocks, with $1 \le q \le r-1$, we have $\tilde{\xi}(q) \xrightarrow{d} N(0,1)$; (ii) under the alternative hypothesis $H_1 = \bigcup_{q<s\le r} H(s)$, $\tilde{\xi}(q) \xrightarrow{p} +\infty$.

Proof: See Appendix B.3.

The feasible asymptotic distribution in Theorem 2 is the building block for a one-sided test of the null hypothesis of q primitive shocks. The rejection region for the test based on $\tilde{\xi}(q)$ at significance level α is $\tilde{\xi}(q) > z_{1-\alpha}$, where $z_{1-\alpha}$ is the $(1-\alpha)$-quantile of the standard Gaussian distribution for

$\alpha \in (0, 1)$. Similarly, we can define an *acceptance region* for the same one-sided test at level α directly in terms of the values of $\hat{\xi}(q)$, that is the sum of the smallest $r - q$ eigenvalues, instead of the standardized statistics $\tilde{\xi}(q)$, as

$$AR_\alpha = \left\{ x \in \mathbb{R} \, : \, 0 \leq x \leq \frac{z_{1-\alpha}}{N\sqrt{T}} \sqrt{\hat{\Omega}_U} + \frac{1}{N} \hat{B}_U \right\}, \tag{4.8}$$

so that we cannot reject the null of q common primitive shocks as long as $\hat{\xi}(q) \in AR_\alpha$. Theorem 2 *(ii)* implies that the test is consistent.

An intermediate result required for the proof of Theorem 2 is the asymptotic expansion of factors' estimator $\hat{f}_t$, which is provided by Lemma C.4 in OA, Section C.9. A direct consequence of this Lemma is that the estimator $\hat{\breve{\Phi}}$ ($\hat{\Phi}$ resp.) is consistent for the auto-regressive matrix Φ ($\breve{\Phi}$ resp.). This implies that we can consistently estimate the common primitive shocks $v_{H,t}$ with the estimator $\hat{v}_{H,t}$ presented in Definition 1. All these results hold up to a rotation and, for the factors, also a change of sign.

4.3 Sequential tests for the number of common primitive shocks q

One way to estimate the number primitive shocks consists in sequentially testing the null hypothesis $H_0 = H(k)$, against the alternative $H_1 = \bigcup_{k<\ell\leq r} H(\ell)$, using the test statistic $\tilde{\xi}(k)$ of Theorem 2. A "naive" estimation procedure is initiated by testing the null of $k = 1$, proceeds by increasing k by one unit and performing the test of the null $k = 2$, and so on, for $k = 1, ..., r-1$. The estimation procedure is stopped at the smallest integer $\hat{q}_{naive} = k$ such that the null $H(k)$ cannot be rejected by performing a one-sided test with significance level α, i.e. at the first k such that $\tilde{\xi}(k) \leq z_{1-\alpha}$. Otherwise, set $\hat{q}_{naive} = r$ if the test rejects the null $H(k)$ for all $k = 1, ..., r - 1$. This "naive" procedure does not return a consistent estimator of q, as there exists an asymptotic probability $\alpha > 0$ of overestimating q due to the type I error when testing $H(q_0)$ against $\bigcup_{q_0<\ell\leq r} H(\ell)$.

Building on the results in Pötscher (1983), Cragg and Donald (1997), and Robin and Smith (2000), a consistent estimator of q is obtained by implementing the above naive procedure but allowing the asymptotic size α of the test to go to zero as $N, T \to \infty$. The following Proposition 2 formalizes this consistent estimation procedure.

PROPOSITION 2. *Let $\alpha_{N,T}$ be a sequence of scalars defined on $(0,1)$ for any N, T, such that: (i) $\alpha_{N,T} \to 0$ and (ii) $(N\sqrt{T})^{-1} z_{1-\alpha_{N,T}} \to 0$ for $N, T \to \infty$, with $z_{1-\alpha_{N,T}} > 0$. Consider the estimator $\hat{q}$ for the number primitive shocks q defined as:*

$$\hat{q} = \min\left\{k : \ 1 \leq k \leq r-1, \ \tilde{\xi}(k) \leq z_{1-\alpha_{N,T}}\right\},$$
$$\textit{or } \hat{q} = r \textit{ if } \ \tilde{\xi}(k) > z_{1-\alpha_{N,T}} \textit{ for all } k = 1, ..., r-1.$$

Then, under Assumptions A.1 - A.9, the estimator $\hat{q}$ is consistent, i.e. $P(\hat{q} = q) \longrightarrow 1$ under $H(q)$, for any integer $q \in \{1, \ldots, r\}$.

The proof of Proposition 2 is a standard proof of consistency for sequential testing procedures, see, e.g. Robin and Smith (2000) for a general proof and AGGR for a particular case with similar rates for $\alpha_{N,T}$. Condition *(i)* ensures an asymptotically zero probability of type I error when testing $H(q_0)$ against $\bigcup_{q_0 < \ell \leq r} H(\ell)$. Condition *(ii)* is a lower bound on the convergence rate to zero of the asymptotic size, and is used to keep asymptotically zero probability of type II error at each step of the procedure. The conditions in Proposition 2 are satisfied when $\alpha_{N,T}$ is such that:

$$z_{1-\alpha_{N,T}} = c(N\sqrt{T})^{\gamma}, \tag{4.9}$$

for constants $c > 0$ and $0 < \gamma < 1$.

5 Bootstrap

Starting from the works of Gonçalves and Perron (2014, 2020) and Gonçalves, Koh, and Perron (2024) we propose a residual-based wild bootstrap implementation of the test in Theorem 1. This approach relies on PCA estimation of the static factors on N_b bootstrapped panels of observations $Y^{(b)} = \left[y_0^{(b)}, y_1^{(b)}, \ldots, y_T^{(b)}\right]'$ with $b = 1, \ldots, N_b$. Differently from Cavaliere, Gonçalves, Nielsen, and Zanelli (2024), we do not use the bootstrap to overcome the issue that the bias and/or the variance of our test statistics cannot be consistently estimated, as we do have derived their consistent estimators. Instead, we use the bootstrap to improve the small sample properties of our test based on the rescaled (using the expression for the variance) and recentered (using the expression for the bias) feasible statistic in equation (4.7). Monte Carlo experiments reported in Section 6 show that this bootstrap approach delivers better small sample properties than the test of Theorem 2. In particular, its actual size is much closer to the nominal one, especially when sample sizes are relatively small and comparable to values often encountered in macro-financial applications, e.g., $N = 100$ and $T = 100$.[15]

5.1 Bootstrap data generating process, estimation and testing procedure

This section describes the non-parametric bootstrap implementation of our test for the number of common primitive shocks. This testing procedure relies on a wild bootstrap resampling scheme and can be implemented as a three-step methodology.

- ***Step (1)***: Estimate the r static factors f_t through the PCA estimator $\hat{f}_t$ of Section 3.2. Estimate a VAR(1) model on $\hat{f}_t$, and let $\hat{\Phi}$ be the estimated autoregressive matrix. Construct the vector of estimated VAR residuals $\hat{v}_t = \hat{f}_t - \hat{\Phi}\hat{f}_{t-1}$. Additionally, use the estimated loadings $\hat{\Lambda}$ (see equation (3.5)) to obtain the estimated residuals $\hat{\varepsilon}_t = y_t - \hat{\Lambda}\hat{f}_t$.

- ***Step (2)***: For each value of $q = 1, \ldots, r-1$, define a new r-dimensional vector:

$$\hat{v}_t^{H_0(q)} := \left[\hat{v}'_{H,t}, 0'_{(r-q,1)}\right]', \tag{5.1}$$

  where $\hat{v}_{H,t}$ is the upper q-dimensional subvector of $\hat{v}_t$, and consider the next steps:

  – ***Step (2.a)***: For each bootstrap iteration $b = 1, ..., N_b$, with N_b large, construct a bootstrap sample $\varepsilon_t^{(b)} = [\varepsilon_{1t}^{(b)}, ..., \varepsilon_{Nt}^{(b)}]'$ from $\hat{\varepsilon}_t$ using a wild bootstrap scheme:

$$\varepsilon_{it}^{(b)} = \hat{\varepsilon}_{it} \cdot \eta_{\varepsilon,it}, \qquad i = 1, ..., N, \; t = 0, \ldots, T, \tag{5.2}$$

  where $\eta_{\varepsilon,it}$ is a zero-mean and unit-variance "external" random variable that is i.i.d. across all individuals and dates.[16] Starting from the variables in (5.1) and (5.2), construct the

[15]Unreported results, available upon requests, showed that our bootstrap procedure has better size and power than one which simply bootstraps the eigenvalue sum $\hat{\xi}(q)$ without estimating the bias and the variance terms, which corresponds to the suggestions of Cavaliere, Gonçalves, Nielsen, and Zanelli (2024) and Gonçalves, Koh, and Perron (2024).

[16]We assume that $\eta_{\varepsilon,it} \sim iiN(0,1)$ in the Monte Carlo analysis and empirical applications.

following bootstrap analogous of the DGP in (2.9)-(2.12) for all $t = 0, 1, ..., T$:

$$y_t^{(b)} = \hat{\Lambda} f_t^{(b)} + \varepsilon_t^{(b)}, \tag{5.3}$$

$$f_t^{(b)} = \hat{\Phi} f_{t-1}^{(b)} + \hat{v}_t^{H_0(q)}, \tag{5.4}$$

where the VAR(1) is initialized at $f_0^{(b)} = \hat{f}_0$.

- ***Step (2.b)***: As detailed in Section D.1 of the OA, use bootstrapped data $Y^{(b)} = [y_0^{(b)}, y_1^{(b)}, ..., y_T^{(b)}]'$ to construct the sum of the smallest $r - q$ eigenvalues $\hat{\sigma}_{q+1}^{2(b)}, \ldots, \hat{\sigma}_r^{2(b)}$ of the estimator for the variance covariance matrix of $\hat{v}_t^{H_0(q)}$:

$$\hat{\xi}^{(b)}(q) = \sum_{\ell=1}^{r-q} \hat{\sigma}_{r-\ell+1}^{2(b)}, \tag{5.5}$$

and

$$\hat{\Sigma}_u^{(b)} = \left(\frac{1}{N}\hat{\Lambda}^{(b)\prime}\hat{\Lambda}^{(b)}\right)^{-1} \left(\frac{1}{N}\hat{\Lambda}^{(b)\prime}\hat{\Gamma}^{(b)}\hat{\Lambda}^{(b)}\right) \left(\frac{1}{N}\hat{\Lambda}^{(b)\prime}\hat{\Lambda}^{(b)}\right)^{-1} = \begin{bmatrix} \hat{\Sigma}_{u,HH}^{(b)} & \hat{\Sigma}_{u,HL}^{(b)} \\ \hat{\Sigma}_{u,LH}^{(b)} & \hat{\Sigma}_{u,LL}^{(b)} \end{bmatrix}, \tag{5.6}$$

where $\hat{\Gamma}^{(b)} = diag(\hat{\gamma}_{ii}^{(b)},\ i = 1, ..., N)$ with $\hat{\gamma}_{ii}^{(b)} = \frac{1}{T+1}\sum_{t=0}^{T} \hat{\varepsilon}_{i,t}^{(b)2}$, for $\hat{\varepsilon}_{i,t}^{(b)}$ the estimator of $\varepsilon_t^{(b)}$ based on the b-th bootstrap sample.[17] Using $\hat{\Sigma}_u^{(b)}$ to derive bootstrap equivalents of $\hat{B}_U$ and $\hat{\Omega}_U$ in Theorem 2, one obtains the bootstrap test statistic

$$\tilde{\xi}^{(b)}(q) := N\sqrt{T}\left(\hat{\Omega}_U^{(b)}\right)^{-1/2}\left[\hat{\xi}^{(b)}(q) - \frac{1}{N}tr\left\{\hat{B}_U^{(b)}\right\}\right]. \tag{5.7}$$

- ***Step (2.c)***: Iterating Steps (2.a) and (2.b) N_b times yields N_b bootstrapped values of the feasible test statistic under the null hypothesis of q primitive shocks. Using these values, one can evaluate the cumulative distribution function of $\tilde{\xi}(q)$ under the bootstrap probability measure at any $c^* \in \mathbb{R}$ as $\hat{F}_{\tilde{\xi}}^B(c^*; q) := \frac{1}{N_b}\sum_{b=1}^{N_b} \mathbb{1}\left\{\tilde{\xi}^{(b)}(q) \leq c^*\right\}$, where $\mathbb{1}\left\{\tilde{\xi}^{(b)}(q) \leq c^*\right\} = 1$ if $\tilde{\xi}^{(b)}(q) \leq c^*$, and 0 otherwise. For any $\alpha \in (0, 1)$, the α-percentile of the bootstrapped distributions of $\tilde{\xi}(q)$ is $\hat{p}_\alpha^B(q) := \inf\left\{p\ :\ \hat{F}_{\tilde{\xi}}^B(p; q) \geq \alpha\right\}$, from which

[17] Expressions for all quantities can be found in Section D.1 of the OA.

we can also construct the bootstrap-based *acceptance rejection* as

$$AR_{\alpha}^{B}=\left\{x\in\mathbb{R}\,:\,0\leq x\leq\frac{1}{N\sqrt{T}}\sqrt{\hat{\Omega}_U}\hat{p}_{1-\alpha}^{B}(q)+\frac{1}{N}\hat{B}_U\right\}, \tag{5.8}$$

for $\hat{\Omega}_U$ and $\hat{B}_U$ computed as in Theorem 2, and which implies not rejecting the null of q common primitive shocks when $\hat{\xi}(q)\in AR_{\alpha}^{B}$.

- ***Step (3)***: Define the bootstrap-based estimator of q as in Proposition 2, this time replacing the adjusted critical value $z_{1-\alpha_{N,T}}$ with the adjusted bootstrapped percentile $\hat{p}^{B}_{1-\alpha^*_{N,T}}(q)$. Hence, the bootstrap-based estimator of q is

  $$\hat{q}_B=\min\left\{k:\ 1\leq k\leq r-1,\ \tilde{\xi}(k)\leq\hat{p}^{B}_{1-\alpha^*_{N,T}}(q)\right\},$$
  $$\text{or }\ \hat{q}_B=r\text{ if }\tilde{\xi}(k)>\hat{p}^{B}_{1-\alpha^*_{N,T}}(k)\text{ for all }k=1,...,r-1,$$

  where

  $$\alpha^*_{N,T}:=\frac{\alpha}{c\left(N\sqrt{T}\right)^{\gamma}}, \tag{5.9}$$

  $c>0$ and $\gamma\in(0,1)$ as in Equation (4.9).

Two remarks are in order. First, the bootstrap DGP in (5.3)-(5.4) satisfies by construction the null hypothesis of q primitive shocks, as the VAR for $f_t^{(b)}=[\,f_{Ht}^{(b)\prime},f_{Lt}^{(b)\prime}\,]'$ reads:

$$\begin{bmatrix} f_{Ht}^{(b)} \\ f_{Lt}^{(b)} \end{bmatrix}=\begin{bmatrix} \hat{\Phi}_{HH} & \hat{\Phi}_{HL} \\ \hat{\Phi}_{LH} & \hat{\Phi}_{LL} \end{bmatrix}\begin{bmatrix} f_{Ht-1}^{(b)} \\ f_{Lt-1}^{(b)} \end{bmatrix}+\begin{bmatrix} \hat{v}_{Ht} \\ 0 \end{bmatrix},$$

showing that the innovations on the factor VAR(1) in the bootstrap DGP have reduced rank q. As discussed by Andreou, Gagliardini, Ghysels, and Rubin (2024), one can establish first order validity of a general bootstrap scheme under high level assumptions on the asymptotic behaviour of the bootstrap test statistic $\tilde{\xi}^{(b)}(q)$ (see their Assumption 2, which is in line with those of Gonçalves, Koh, and Perron, 2024).[18] Second, the functional form of $\alpha^*_{N,T}$ in (5.9) is a Bonferroni-type correction as in Trapani

[18] Establishing primitive conditions for validity of our wild bootstrap scheme is beyond the scope of this paper, and we leave this task to future research. Nevertheless, our MC experiments confirm that the bootstrap has the excellent size and power in both finite samples.

(2018), and ensures that $\left(N\sqrt{T}\right)^{-1} p^{B}_{1-\alpha^{*}_{N,T}} \to 0$ when the proposed bootstrap is valid. If this is the case, one can show that $P\left(\hat{q}_B = q\right) \to 1$ as $N, T, B \to \infty$ using arguments similar to those of Proposition 2. As discussed in Section 4.3, adjusting the probability level of the bootstrap percentile so that $\alpha^{*}_{N,T} \to 0$ is necessary for consistency of $\hat{q}_B$.

6 Monte Carlo simulation analysis

The objectives of the Monte Carlo (MC) simulation study are: i) assessing the adequacy of the asymptotic standard Gaussian distribution of $\tilde{\xi}(q)$ to approximate its small sample counterpart; ii) evaluating the size and power properties of the test in Theorem 1 both under the plug-in and the bootstrap implementations, and iii) comparing the asymptotic and bootstrap-based estimator of q with the alternatives suggested in the literature.

6.1 Simulation Design

We simulate the observation $y_{i,t}$ for $i = 1, ..., N$, $t = 0, 1, ..., T$ from the following factor model:

$$y_{i,t} = \breve{\lambda}_i' \breve{f}_t + \varepsilon_{i,t}.$$

The N-dimensional vectors of idiosyncratic innovations $\{\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_T\}$ with $\varepsilon_t = [\varepsilon_{1,t}, ..., \varepsilon_{i,t}, ...\varepsilon_{N,t}]'$ are i.i.d. draws from a Gaussian random variable $N(0_{N,1}, I_N)$. We resample these innovations in each MC simulation. For each individual i, the loadings are drawn from N independent Gaussian distributions, $\breve{\lambda}_i \sim i.i.N.(\ 0,\ I_r)$ for $i = 1, \ldots, N$.

The r-dimensional vector $\breve{f}_t$ follows the stationary VAR(1) process: $\breve{f}_t = \breve{\Phi}\breve{f}_{t-1} + \breve{v}_t$, and $\breve{v}_t = G\eta_t$, where $\breve{\Phi}$ is an (r, r) autoregressive matrix. The (r, q) matrix G links the q primitive shocks to the r factor innovations $\breve{v}_t$, and is simulated at each iteration as in Section 5 of Bai and Ng (2007). To generate G, we start from an (r, r) diagonal matrix S whose first q non-zero elements are drawn from q independent uniform distributions $U(.01, 0.31)$ so that S has rank q. Second, we consider an arbitrary orthonormal matrix R, i.e. $RR' = I_r$, that we obtain in Matlab through "R = *orth(rand(r, r))*" at each MC iteration. Third, we set $G = RSR'$ and keep it constant across all i and t for a given MC sample. Note that the variance-covariance matrix of $\breve{v}_t$ is $\breve{\Sigma}_v = RS^2R'$ and has rank q.

Consistently with the number of static factors often documented in macroeconomic studies (see Onatski, 2010), we consider a DGP based on $r = 7$ static factors and $q_0 = 5$ primitive shocks.[19] The autoregressive matrix of the VAR(1) process is given by

$$\breve{\Phi} = diag(0.2, 0.2875, 0.375, 0.55, 0.725, 0.8125, 0.9),$$

while the q_0 primitive shocks η_t are always simulated as $\eta_t \sim i.i.d.N(0, I_{q_0})$. Results using other DGPs are reported in Section F.1 of the OA. We always consider $M = 2000$ Monte Carlo samples.

6.2 Asymptotic and bootstrap distribution; size and power properties

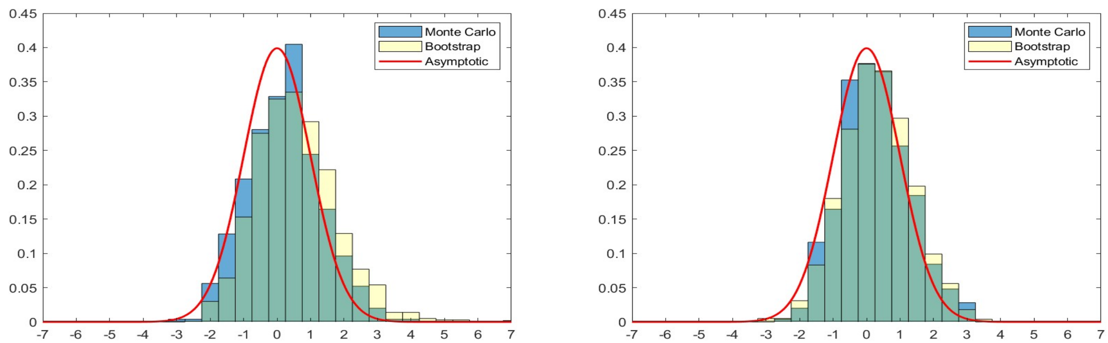
First, we study whether the asymptotic Gaussian distribution and the one based on N_b bootstrap samples provide a good approximation to the small-sample distribution of the "plug-in" test statistics $\tilde{\xi}(q)$ in Theorem 2. Blue histograms in Figure 1 display the empirical distribution of $\tilde{\xi}(q)$ under the null hypothesis $q = q_0$, while red solid lines represent the density of the asymptotic $N(0, 1)$ distribution. Though the empirical distribution is shifted to the right with respect to the asymptotic one when $(N, T) = (100, 100)$, the two become much closer when $(N, T) = (400, 600)$. Yellow histograms visualize the distribution of $\tilde{\xi}^{(b)}(q_0)$ across $N_b = 499$ bootstrap replicates for the first Monte Carlo sample. The bootstrap distribution is very close to the empirical one, thus suggesting that the bootstrap-based test will outperform the asymptotic one, at least for the smallest sample sizes.

Table 2 presents empirical sizes and powers of the test based on $\tilde{\xi}(q)$ in Theorem 2. Results are presented for both the plug-in version (left panel) and the bootstrap one (right panel). For the plug-in test, we consider significance levels $\alpha \in \{0.01, 0.05, 0.1\}$ and reject the null $H(q) = H(5)$ when the test statistics is larger than the $(1 - \alpha)$-quantile of the standard Gaussian distribution. In the bootstrap case, we look at the same percentiles but computed from the bootstrapped distribution of the test statistic. The latter is based on $N_b = 499$ bootstrap iterations for each MC sample. Empirical powers are computed as the rejection frequency of the null hypotheses $H_0 = H(3)$ and $H_0 = H(4)$, against the alternatives $q > 3$ and $q > 4$, respectively. We report powers for the 5% significance level.

The plug-in version is oversized even when $N = 400$ and $T = 600$, which are values larger than what usually encountered in macro-financial analyses, e.g. the FRED-MD database of monthly

[19]The notation q_0 highlights that this is the true number of primitive shocks that we want to estimate.

Figure 1 – Small sample and bootstrapped distribution of the test statistic $\tilde{\xi}(q_0)$.



(a) $N = 100$, and $T = 100$. **(b)** $N = 400$, and $T = 600$.

Blue histograms report the empirical distribution of the test statistic $\tilde{\xi}(q_0)$ for $(N, T) = (100, 100)$ and $(N, T) = (400, 600)$ across $M = 2000$ Monte Carlo samples. Red solid lines correspond to the density of the asymptotic distribution $N(0, 1)$ of the re-centered and re-scaled statistic. Yellow histograms visualize the bootstrap distribution of the test statistic for the first Monte Carlo sample.

macroeconomic indicators of the US economy. The bootstrap test corrects this over-rejection of the null, particularly when $N < 200$. The asymptotic test has unit power irrespectively of $H(q)$ and of the combination of N and T.[20] The bootstrap implementation also returns good power results. Lastly, Table 10 in the OA shows that actual sizes of the plug-in implementation become close to the nominal ones only when N and T are very large, e.g. $N \geq 1000$, and $T \geq 600$.

6.3 Estimation of the number of primitive shocks

We now assess the finite sample properties of our estimators of q. To understand the finite sample improvements of the adjustment discussed in Proposition 2 and Step 3 of the bootstrap algorithm, we implement the sequential testing procedures using both adjusted and unadjusted critical values, e.g. $z_{1-\alpha_{N,T}}$ and $z_{1-\alpha}$ in the asymptotic case. Critical values are adjusted as in Equations (4.9) and (5.9), for which we fix $c = 0.95$ and $\gamma = 0.1$.[21] We compare performances of our estimators with those of several competitors: $\hat{q}_3$ and $\hat{q}_4$ introduced by Bai and Ng (2007, BN henceforth); $\hat{q}_{aw,A}$ and $\hat{q}_{aw,B}$ from Amengual and Watson (2007, AW), and $\hat{q}_{bp}$ of Breitung and Pigorsch (2013, BP). Like our approach, these estimators – summarized in Appendix E – were developed to determine the number of common shocks in factor models as that in (2.1) - (2.2). Section F.3 of the OA shows that our estimators perform

[20] The same conclusion holds when the size-adjusted power is considered, results are available upon request.

[21] These are the same values adopted by AGGR and provide good finite sample results for all DGPs that we consider.

Table 2 – Empirical sizes and powers of the plug-in and of the bootstrap versions of the test of the number of primitive shocks

		Plug-in: Th. 2					*Bootstrap: Th. 2*				
		size			*power*		*size*			*power*	
N	T	1%	5%	10%	$H(3)$	$H(4)$	1%	5%	10%	$H(3)$	$H(4)$
100	100	0.07	0.17	0.25	1.00	0.99	0.03	0.09	0.15	0.98	0.95
100	200	0.13	0.28	0.39	1.00	1.00	0.02	0.08	0.14	0.99	0.96
200	100	0.03	0.09	0.13	1.00	1.00	0.02	0.08	0.14	0.99	0.98
200	200	0.03	0.12	0.19	1.00	1.00	0.02	0.07	0.12	0.99	0.99
200	300	0.05	0.15	0.22	1.00	1.00	0.01	0.06	0.12	0.99	0.99
400	100	0.02	0.06	0.09	1.00	1.00	0.02	0.07	0.13	1.00	1.00
400	200	0.02	0.07	0.13	1.00	1.00	0.02	0.06	0.12	1.00	1.00
400	300	0.02	0.07	0.13	1.00	1.00	0.01	0.06	0.12	1.00	1.00
400	600	0.02	0.11	0.18	1.00	1.00	0.01	0.06	0.11	1.00	0.99

Empirical sizes and powers of the test for the null hypothesis of q primitive shocks. Results in the left panel are based on the feasible test statistic in Theorem 2. Those in the right panel pertain to the bootstrap implementation. Simulated data come from the DGP of Section 6.1 with $r = 7$ and $q_0 = 5$. Empirical sizes are assessed at significance levels $\alpha \in \{0.01, 0.05, 0.1\}$, while powers represent the empirical rejection frequency of the null hypotheses $H_0 = H(3)$ and $H_0 = H(4)$ under the alternatives $q > 3$ and $q > 4$, respectively. These powers are assessed at the 5% significance level. Results are based on $M = 2000$ MC simulations.

well also with respect to those of Hallin and Liska (2007), which were originally developed for the estimation of q in generalized dynamic factor models, and, in our simulation designs, consistently outperform the procedure of Onatski (2009), that we do not report. We always generate $M = 2000$ Monte Carlo samples and, for our tests, set $\alpha = 0.05$ in line with the power results from the previous analysis.

Table 3 reports the average estimated number of shocks using the five approaches. The third and fourth columns present results for the estimators of BN. For all sample sizes, our four estimators improve upon those of Bai and Ng (2007), which consistently underestimate q_0. Estimators $\hat{q}_{aw,A}$ and $\hat{q}_{aw,B}$ underestimate q_0 when N and T are small but their performance improves as both sizes increase. While both bootstrap-based estimators outperform those of AW when $N < 400$ and $T \leq 200$, only the adjusted one consistently performs better than – or on par with – $\hat{q}_{aw,A}$ and $\hat{q}_{aw,B}$. The approach of BP underestimates q_0 and is always outperformed by all our estimators improving. The asymptotic sequential procedure based on standard Gaussian quantiles (eight column, labelled by z_α) always overestimates q, mostly because the test is oversized for these sample sizes.

Results substantially improve when we adjust the critical value of the test (ninth column, labeled by $z_{\alpha_{N,T}}$). Finally, the bootstrap-based estimator (tenth column, labeled by p^B_α) delivers the best results when $N = 100$ but is outperformed by its adjusted counterpart ($p^B_{\alpha^*_{N,T}}$) for all the other sample sizes.

This is a consequence of adjusting the level of the test, as the non-zero asymptotic type I error probability makes the sequential procedure reject the null even when it holds, which implies overestimating the number of common primitive shocks.

Table 3 – Comparison of estimators of q

N	T	$\hat{q}_3$	$\hat{q}_4$	$\hat{q}_{aw,A}$	$\hat{q}_{aw,B}$	$\hat{q}_{bp}$	z_α	$z_{\alpha_{N,T}}$	p^B_α	$p^B_{\alpha^*_{N,T}}$
100	100	4.48	4.49	4.86	4.88	4.23	5.17	5.12	5.01	4.97
100	200	4.49	4.50	4.92	4.93	4.39	5.30	5.20	5.03	4.97
200	100	4.45	4.46	4.90	4.92	4.34	5.07	5.04	5.06	5.01
200	200	4.63	4.63	4.94	4.95	4.55	5.12	5.06	5.05	5.01
200	300	4.64	4.64	4.96	4.96	4.63	5.17	5.06	5.05	5.01
400	100	4.44	4.45	4.92	4.93	4.40	5.04	5.01	5.07	5.03
400	200	4.62	4.62	4.96	4.97	4.63	5.06	5.02	5.06	5.02
400	300	4.70	4.70	4.97	4.98	4.71	5.08	5.02	5.06	5.02
400	400	4.74	4.74	4.99	4.99	4.79	5.12	5.03	5.05	5.02

Average estimates of q when $r = 7$ and $q_0 = 5$. The third and the fourth columns present results for estimators $\hat{q}_3$ and $\hat{q}_4$ of Bai and Ng (2007). The fifth and sixth columns consider $\hat{q}_{aw,A}$ and $\hat{q}_{aw,B}$ by Amengual and Watson (2007), while the seventh one is based on $\hat{q}_{bp}$ of Breitung and Pigorsch (2013). The eighth and ninth columns show results for our estimator $\hat{q}$ based on the asymptotic sequential testing procedure. The former is based on the 95% quantile of the asymptotic $N(0,1)$ distribution while the latter considers quantiles adjusted for a consistent selection procedure. The last two columns are based on the bootstrap version of the test performed at the 5% significance level: the first one considers unadjusted bootstrap percentiles (i.e. p^B_α) and the second adjusted ones (i.e. $p^B_{\alpha^*_{N,T}}$). We adjust the significance level α using equations (4.9) and (5.9), where we always set $c = 0.95$ and $\gamma = 0.1$. The number of MC samples is always $M = 2000$ MC.

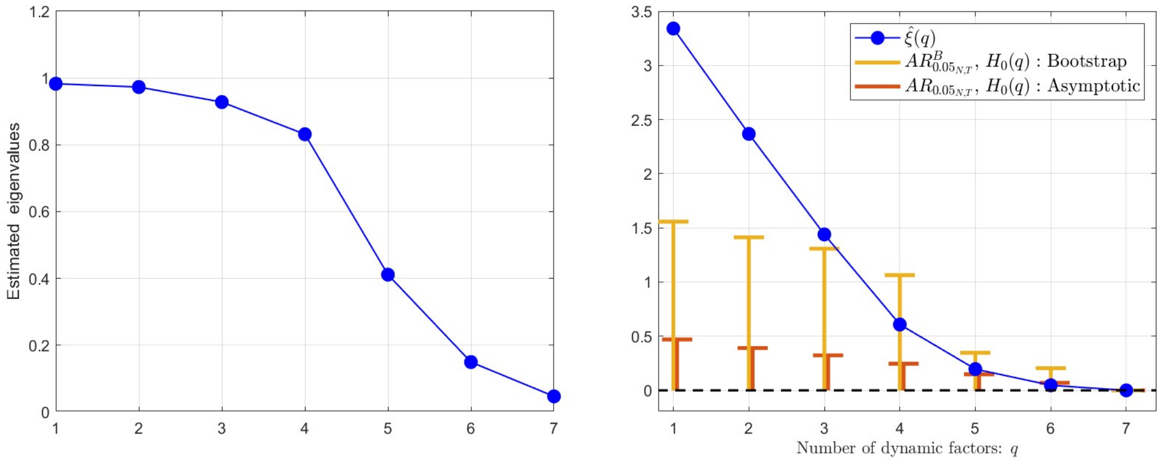
7 Common shocks in volatility and macro-financial panels

7.1 Common shocks in US macro-financial data

We consider a macro-financial application where we test the number of common shocks in the FRED-MD monthly dataset of McCracken and Ng (2016). In particular, we work with a balanced panel of $N = 120$ monthly indicators of the US economic and financial system ranging between January 1960 and December 2019 ($T = 720$). This is the longest dataset not contaminated by the COVID-crisis. We consider the September 2022 vintage and make all the series stationary following the suggestions of McCracken and Ng (2016). As recommended when using this dataset, we remove outliers following the procedure of McCracken and Ng (2016).

Information criteria IC_{p1} and IC_{p2} of Bai and Ng (2002a), and their modifications by Alessi, Barigozzi, and Capasso (2010), suggest the presence of $r = 7$ static factors. We use this value as

Figure 2 – Eigenvalue analysis for the covariance matrix of VAR innovations $\breve{v}_t$ when $\hat{r} = 7$ in US macroeconomic data



(a) Estimated eigenvalues when $\hat{r} = 7$.

(b) $\hat{\xi}(q)$ when $\hat{r} = 7$.

Left panel: estimated eigenvalues of the covariance matrix of factors' VAR(1) when $\hat{r} = 7$. Right panel: sum of the smallest $r - q$ eigenvalues $\hat{\xi}(q)$ (blue solid line) when $\hat{r} = 7$ for multiple values of q. In the right panel, vertical bars denote the acceptance region when testing the null hypothesis of q primitive shocks, i.e. $H_0(q)$, at the adjusted 5% significance level. Yellow bars pertain to the bootstrap-based test while orange ones come from the asymptotic version of the test.

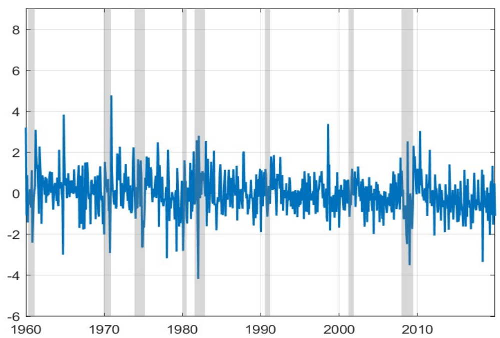
starting point for our sequential testing procedure. Following results of the Monte Carlo analysis, we use the adjusted bootstrap estimator of q, run the testing procedure with $N_b = 999$ bootstrap samples, and consider 5% level of significance, which we adjust as in Equation (5.9). The null of $q = 4$ common primitive shocks is the first one not to be rejected. The same result holds when the procedure is run at the (adjusted) 1% level of significance, and if $N_b = 499$ and $N_b = 1499$ bootstrap samples are considered. Thus, we conclude that the common temporal variation in US macro-financial system is fully characterized by four primitive shocks.[22]

Figure 2(a) shows the estimated eigenvalues when $\hat{r} = 7$. Estimates for the first four eigenvalues of the VAR innovations' covariance matrix $\breve{\Sigma}_v$ range between 0.98 and 0.83. We then observe a sharp decrease in the magnitude of the eigenvalues, with the fifth one taking value 0.41. Our test signals that we cannot reject the hypothesis that this eigenvalue is zero at adjusted significance level 5%. Remaining eigenvalues are estimated at 0.15 and 0.05. The blue solid line in Figure 2(b) represents the sum of the smallest $r - q$ eigenvalues, i.e. $\hat{\xi}(q)$, when $r = 7$ and q varies between one and six. Vertical bars denote the acceptance region when testing the null hypothesis of q shocks, i.e. $H_0(q)$,

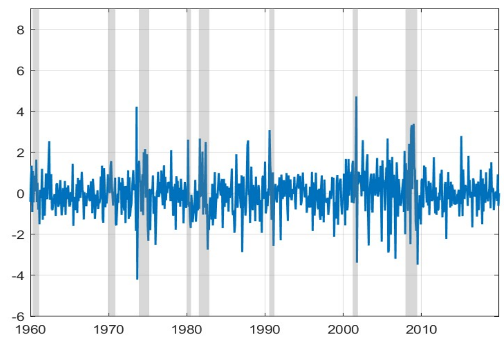
[22]Note that Bai and Ng (2007) reach same conclusions in terms of (IC-based) $\hat{r}$ and $\hat{q}$ when analyzing the monthly dataset of Stock and Watson (2005).

at the adjusted 5% significance level. Orange ones denote the plug-in version (see equation (4.8), replacing $z_{1-\alpha}$ with $z_{1-\alpha_{N,T}}$), while yellow bars are for the bootstrap-based implementation of the test (see equation (5.8) with $\hat{p}^B_{1-\alpha}$ replaced by $\hat{p}^B_{1-\alpha^*_{N,T}}$). In line with Monte Carlo results, acceptance regions for the plug-in implementation are narrower than those for the bootstrap-based version, so that the former estimates five instead of four common shocks.

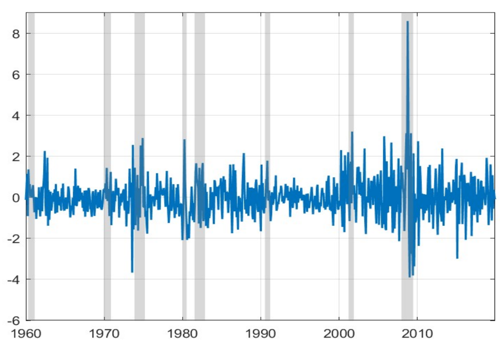
Figure 3 – Estimated non-redundant factors on US macroeconomic data



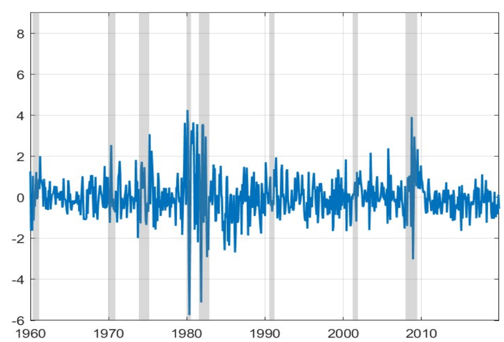
(a) First factor



(b) Second factor



(c) Third factor



(d) Fourth factor

Monthly values of the estimated non-redundant factors $\hat{f}_{H,t}$ from the panel of macroeconomic data in FRED-MD between January 1960 and December 2019. Grey shaded areas denote official NBER recession dates. Non-redundant factors are estimated as in Definition 1.

Figure 3 plots the four non-redundant factors between January 1960 and December 2019, while Table 4 reports the ten observable macro-financial variables that exhibit the highest absolute correlation with each of these factors. The first factor is positively correlated with time series that characterise the output of the US economy.[23] Hence, this is a cyclical factor that approximates the state of the

[23] All this discussion is based on the eight groups of variables constructed by McCracken and Ng (2016). These are: Output and Income; Labor Market; Housing; Consumption, orders, and inventories; Money and Credit; Interest and Exchange Rates; Prices; Stock Market.

Table 4 – Variables exhibiting the highest absolute correlation with the non-redundant macroeconomic factors

Factor 1		Factor 2	
IP: Final Products	0.74	S&P 500	-0.64
IP: Consumer Goods	0.73	S&P Index: Industrials	-0.62
IP: Final Products and Nonindustrial Supplies	0.71	S&P Index: Dividend Yield	0.61
IP: Total Index	0.67	CPI: All Items Less Shelter	0.57
IP: Manufacturing (SIC)	0.65	CPI: All Items	0.56
CU: Manufacturing	0.64	CPI: Commodities	0.56
IP: Durable Consumer Goods	0.64	PCE: Non-durable good	0.56
IP: Materials	0.52	CPI: All Items Less Medical Care	0.55
IP: Durable Goods Materials	0.51	S&P Index: Price-Earnings Ratio	-0.53
IP: Business Equipment	0.50	CPI: All Items Less Food	0.48

Factor 3		Factor 4	
CPI: Commodities	-0.69	5-Year Treasury Rate	0.76
PCE: Non-durable goods	-0.69	1-Year Treasury Rate	0.76
CPI: All Items Less Shelter	-0.68	10-Year Treasury Rate	0.74
PCE: Chain Index	-0.67	6-Month Treasury Bill	0.72
CPI: All Items	-0.66	AAA Corporate Bond Yield	0.72
CPI: Transportation	-0.56	BAA Corporate Bond Yield	0.64
CPI: All Items Less Medical Care	-0.65	3-Month Treasury Bill	0.64
CPI: All Items Less Food	-0.67	Effective Fed Funds Rate	0.39
PPI: Intermediate Materials	-0.60	CHF/USD ForEx rate	0.35
PPI: Finished Consumer Goods	-0.57	IP: Consumer Goods	0.33

For each of the four non-redundant factors extracted from the panel of macroeconomic data in FRED-MD between January 1960 and December 2019, we report the ten macroeconomic variables in FRED-MD characterised by the highest absolute correlation with the factor itself, and computed on the same sample period. For each variable, the value of the correlation coefficient is reported.

US economy. The second factor is exposed to fluctuations in price indexes and in the stock market, while the third one is solely driven by the level of prices in the economy. Thus, the second and the third factors gauge the behaviour of month-over-month inflation in the US. Finally, the fourth factor is influenced by interest and exchange rates, and therefore behaves as an indicator of the US financial system: the higher its value the worsen the financial outlook, especially for what regards the funding market.

7.2 Common shocks in volatility measures

We now study common shocks in a panel of volatility measures for the constituents of the S&P 500 index. We consider the S&P 500 composition of December 2023 and obtain daily prices from the *Datastream* platform for a period ranging between December 28, 2018 and December 28, 2023 ($T = 1256$). Following Brownlees and Gallo (2010) and Barigozzi, Cho, and Owens (2023), we measure volatility for the i-th stock on the t-th day using the high-low range:

$$\sigma_{i,t}^2 = 0.361 \left(p_{i,t}^{high} - p_{i,t}^{low} \right)^2 ,$$

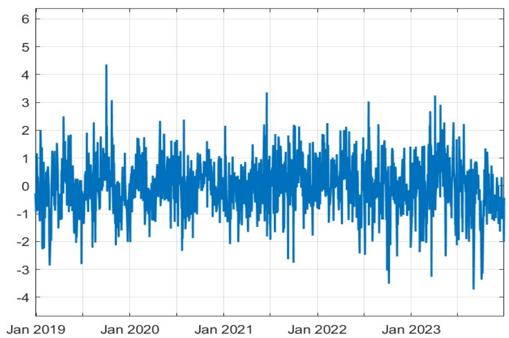
where $p_{i,t}^{high}$ ($p_{i,t}^{low}$) is the highest (lowest) log-price on day t for stock i. In what follows, $y_{i,t} = \log \left(\sigma_{i,t}^2 \right)$.

The same information criteria of Section 7.1 suggest the presence of $r = 7$ static factors, while our bootstrap-based procedure estimates $q = 4$ common shocks.[24] Figure 4 plots the four non-redundant factors estimated on the sample of interest, while table 5 reports the ten stocks which exhibit the highest absolute correlation with these factors. We also report linear correlation coefficients between the stocks and the factor of interest. The first factor is negatively related to providers of electricity and natural gas.[25] These firms experienced periods of higher volatility during both the COVID pandemic and the energy crisis driven by the Russian invasion of Ukraine. The second factor is negatively related to firms that deal with oil extraction, while it exhibits a positive correlation with healthcare providers. Volatility on oil firms peaked during the COVID pandemic as a consequence of extremely low oil prices. The third factor is highly correlated with technology firms that are heavily reliant on microchips. The global shortage of semi-conductors that occurred between 2020 and 2023 explains the importance of these firms for our dataset. Such a shortage was driven by a combination of the COVID pandemic, and of a trade war between the US and China. Finally, the fourth factor correlates with US commercial banks, particularly regional ones. These firms were extremely volatile during the spring 2023 amid the failure of three US commercial banks and the rescue of Credit Suisse.

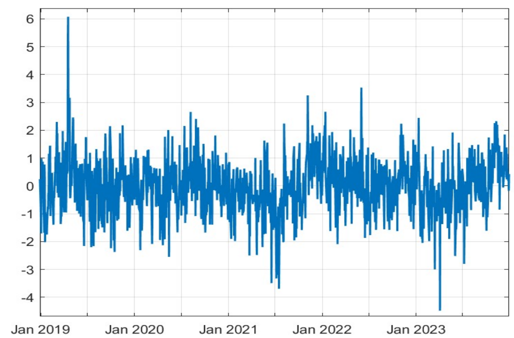
[24] As in the previous application, we consider $N_b = 999$ bootstrap samples and check for the robustness of our findings with respect to this value. The bootstrap test is run at the adjusted 5% level of significance and we check robustness with to the respect to the adjusted 1% level of significance.

[25] All stocks are categorised according to the sub-industry code of the Global Industry Classification Standard.

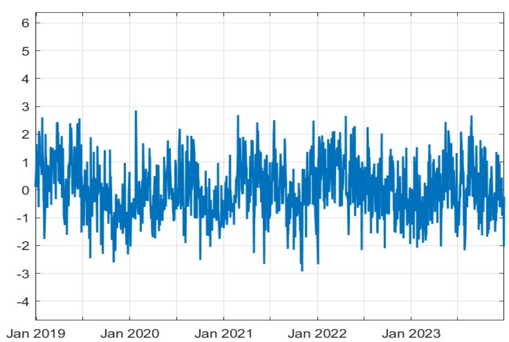
Figure 4 – Estimated non-redundant factors estimated from a panel of individual stocks volatilities



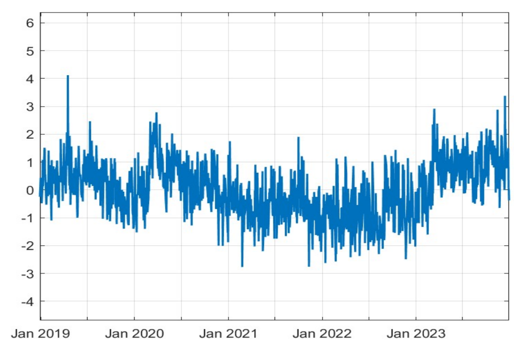
(a) First factor



(b) Second factor



(c) Third factor



(d) Fourth factor

Daily values of the estimated non-redundant factors $\hat{f}_{H,t}$ from a panel of volatility proxies for S&P 500 constituents, between December 28, 2018 and December 28, 2023. Non-redundant factors are estimated as in Definition 1.

Table 5 – Stocks exhibiting the highest absolute correlation with the non-redundant volatility factors

Factor 1		Factor 2	
Alliant Energy	-0.35	Halliburton	-0.31
Ameren	-0.35	Schlumberger	-0.28
WEC Energy	-0.34	Marathon Oil	-0.27
CMS Energy	-0.34	Devon Energy	-0.26
Consolidated Edison	-0.33	APA Corporation	-0.26
Duke Energy	-0.33	ConocoPhillips	-0.25
American Electric Power	-0.33	Agilent Technologies	0.25
Eversource	-0.31	Diamondback Energy	-0.24
NiSource	-0.31	Pioneer Natural Resources	-0.24
Xcel Energy	-0.30	Idexx Laboratories	0.24

Factor 3		Factor 4	
Nvidia	0.39	KeyCorp	0.50
Applied Materials	0.38	Comerica	0.46
Micron Technology	0.37	Citizens Financial Group	0.44
Lam Research	0.35	Zions Bancorporation	0.44
Advanced Micro Devices	0.35	Truist	0.44
Microchip Technology	0.34	U.S. Bank	0.42
Teradyne	0.33	Huntington Bancshares	0.42
Analog Devices	0.32	Fifth Third Bank	0.41
Broadcom Inc.	0.32	PNC Financial Services	0.40
Skyworks Solutions	0.31	Regions Financial Corporation	0.39

For each of the four non-redundant factors extracted from the panel of volatility proxies for S&P 500 constituents between December 28, 2018 and December 28, 2023 we report the ten stocks characterised by the highest absolute correlation with the factor itself, and computed on the same sample period. For each variable, the value of the correlation coefficient is reported.

8 Conclusions

We present new tests and estimators for the number of common primitive shocks in a large (restricted) dynamic factor model, where factors can be estimated by PCA. The starting point of our testing procedure is a static factor representation of the model where r factors evolve as a VAR(1) with innovations having rank-deficient covariance matrix Σ_v. In particular, its rank $q \leq r$ coincides with the number of common primitive shocks in the data. Hence, we test the number of such shocks by testing the rank of Σ_v. In doing it, we are the first to develop a test for the rank of a finite dimensional positive semi-definite matrix in panel data where the number of both cross-sectional entities and observations over time diverge. We are able to overcome the well known problems of testing the rank of a semi-definite matrix, and construct a test statistic whose asymptotic distribution under the null of q primitive shocks is Gaussian, and has non-standard convergence rate $N\sqrt{T}$. This is done by exploiting the estimation error of the principal component estimator of the r static factors and of related quantities, e.g. Σ_v and its eigenvalues/eigenvectors. We propose two implementations of the test: one is based on the asymptotic distribution of a consistent plug-in estimator for the test statistic, while the other relies on a wild bootstrap scheme. We also introduce estimators of the number of common shocks and non-redundant factors based on both implementations.

An analysis of the factor structure of the FRED-MD dataset suggests that output measures and price indexes explain most of the temporal variation in the US macro-financial system between January 1960 and December 2019. An application to volatility measures of US stocks shows that the COVID pandemic and the bank crisis of March 2023 were key drivers of volatility between January 2019 and January 2024.

The tests and estimators of this paper can be naturally extended to the Factor Augmented VAR (FAVAR) model of Bernanke, Boivin, and Eliasz (2005) where both latent factors estimated by PCA and observable factors follow a singular VAR model with a smaller number q of primitive shocks. This extension is in our current research agenda.

References

AHN, S. C., AND A. R. HORENSTEIN (2013): "Eigenvalue Ratio Test for the Number of Factors," *Econometrica*, 81, 1203–1227.

ALDOUS, D., AND G. EAGLESON (1963): "On Mixing and Stability of Limit Theorem," *Annals of Probability*, 6, 325–331.

ALESSI, L., M. BARIGOZZI, AND M. CAPASSO (2010): "Improved penalization for determining the number of factors in approximate factor models," *Statistics & Probability Letters*, 80(23-24), 1806–1813.

AMENGUAL, D., AND M. W. WATSON (2007): "Consistent Estimation of the Number of Dynamic Factors in a Large N and T Panel," *Journal of Business and Economic Statistics*, 25, 91–96.

ANDREOU, E., P. GAGLIARDINI, E. GHYSELS, AND M. RUBIN (2019): "Inference in Group Factor Models with an Application to Mixed-frequency Data," *Econometrica*, 87, 1267–1305.

ANDREOU, E., P. GAGLIARDINI, E. GHYSELS, AND M. RUBIN (2024): "Spanning latent and observable factors," *Journal of Econometrics*, p. 105743.

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

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

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

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

——— (2007): "Determining the Number of Primitive Shocks in Factor Models," *Journal of Business and Economic Statistics*, 25, 52–60.

BARIGOZZI, M., H. CHO, AND D. OWENS (2023): "FNETS: Factor-adjusted network estimation and forecasting for high-dimensional time series," *Journal of Business & Economic Statistics*, pp. 1–13.

BARIGOZZI, M., AND M. HALLIN (2024): "Dynamic factor models: A genealogy," in *Partial Identification in Econometrics and Related Topics*, pp. 3–24. Springer.

BARIGOZZI, M., AND M. LUCIANI (2019): “Quasi maximum likelihood estimation and inference of large approximate dynamic factor models via the EM algorithm,” *arXiv preprint arXiv:1910.03821*.

BERNANKE, B. S., J. BOIVIN, AND P. ELIASZ (2005): “Measuring the Effects of Monetary Policy: A Factor-Augmented Vector Autoregressive (FAVAR) Approach,” *Quarterly Journal of Economics*, 120(1), 387–422.

BILLINGSLEY, P. (1995): *Probability and Measure*. Wiley.

BREITUNG, J., AND U. PIGORSCH (2013): “A Canonical Correlation Approach for Selecting theNumber of Dynamic Factors,” *Oxford Bulletin of Economics and Statistics*, 75(1), 23–36.

BROWNLEES, C. T., AND G. M. GALLO (2010): “Comparison of volatility measures: a risk management perspective,” *Journal of Financial Econometrics*, 8(1), 29–56.

CAVALIERE, G., S. GONÇALVES, M. Ø. NIELSEN, AND E. ZANELLI (2024): “Bootstrap inference in the presence of bias,” *Journal of the American Statistical Association*, 119(548), 2908–2918.

CRAGG, J. G., AND S. G. DONALD (1996): “On the Asymptotic Properties of LDU-Based Tests of the Rank of a Matrix,” *Journal of the American Statistical Association*, 91, 1301–1309.

——— (1997): “Inferring the Rank of a Matrix,” *Journal of Econometrics*, 76, 223–250.

DAVIDSON, J. (1994): *Stochastic Limit Theory*. Oxford University Press.

DONALD, S. G., N. FORTUNA, AND V. PIPIRAS (2007): “On Rank Estimation in Symmetric Matrices: The Case of Indefinite Matrix Estimators,” *Econometric Theory*, 23, 1103–1123.

——— (2010): “On Rank Estimation in Semidefinite Matrices,” *Unpublished Manuscript, CEF.UP Working Paper 2-2010*.

DONALD, S. G., N. FORTUNA, AND V. PIPIRAS (2014): “On estimating the rank of semidefinite matrix,” Discussion Paper, UNC Chapel Hill.

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., AND L. REICHLIN (1998): “Let’s Get Real: A Factor Analytical Approach to Disaggregated Business Cycle Dynamics,” *Review of Economic Studies*, 65, 453–473.

FORTIN, A.-P., P. GAGLIARDINI, AND O. SCAILLET (2023a): "Eigenvalue Tests for the Number of Latent Factors in Short Panels*," *Journal of Financial Econometrics*, p. nbad024.

——— (2023b): "Latent Factor Analysis in short panels," *arXiv preprint arXiv:2306.14004*.

GILL, L., AND A. LEWBEL (1992): "Testing the Rank and Definiteness of Estimated Matrices with Applications to Factor, State-space and ARMA Models," *Journal of the American Statistical Association*, 87, 766–776.

GONÇALVES, S., J. KOH, AND B. PERRON (2024): "Bootstrap inference for group factor models," *Journal of Financial Econometrics*, p. nbae020.

GONCALVES, S., AND B. PERRON (2014): "Bootstrapping Factor-augmented Regression Models," *Journal of Econometrics*, 182, 156–173.

GONÇALVES, S., AND B. PERRON (2020): "Bootstrapping factor models with cross sectional dependence," *Journal of Econometrics*, 218(2), 476–495.

HALL, P., AND C. HEYDE (1980): *Martingale Limit Theory and Its Application*. Academic Press.

HALLIN, M., AND R. LISKA (2007): "Determining the Number of Factors in the General Dynamic Factor Model," *Journal of the American Statistical Association*, 102, 603–617.

KAPETANIOS, G. (2010): "A Testing Procedure for Determining theNumber of Factors in Approximate Factor-Models With Large Datasets," *Journal of Business & Economic Statistics*, 28(3), 397–409.

KATO, T. (1966): *Perturbation Theory for Linear Operators*. Springer, berlin.

KLEIBERGEN, F., AND R. PAAP (2006): "Generalized Reduced Rank Tests Using the Singular Value Decomposition," *Journal of Econometrics*, 133, 97–126.

KUERSTEINER, G., AND I. PRUCHA (2013): "Limit Theory for Panel Data Models with Cross Sectional Dependence and Sequential Exogeneity," *Journal of Econometrics*, 174, 107–126.

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

MCCRACKEN, M., AND S. NG (2016): "FRED-MD: A Monthly Database for Macroeconomic Research," *Journal of Business and Economic Statistics*, 36, 574–589.

ONATSKI, A. (2009): “Testing Hypotheses About The Number of Factors in Large Factor Models,” *Econometrica*, 77, 1447–1479.

——— (2010): “Determining the Number of Factors from Empirical Distribution of Eigenvalues,” *Review of Economics and Statistics*, 92, 1004–1016.

PÖTSCHER, B. M. (1983): “Order Estimation in ARMA-Models by Lagrangian Multiplier Tests,” *Annals of Statistics*, 11, 872–885.

RENYI, A. (1963): “On Stable Sequences of Events,” *Sankhya*, 25, 293–302.

ROBIN, J. M., AND R. SMITH (2000): “Tests of Rank,” *Econometric Theory*, 16, 151–175.

RUYMGAART, F. H., AND S. YANG (1997): “Some Applications of Watson’s Perturbation Approach to Random Matrices,” *Journal of Multivariate Analysis*, 60, 48–60.

SCHOTT, J. R. (2017): *Matrix Analysis for Statistics, Third Edition*. Wiley, New York, third edn.

STOCK, J. H., AND M. W. WATSON (2002): “Forecasting Using Principal Components from a Large Number of Predictors,” *Journal of the American Statistical Association*, 97, 1167 – 1179.

STOCK, J. H., AND M. W. WATSON (2005): “Implications of dynamic factor models for VAR analysis,” .

TRAPANI, L. (2018): “A randomized sequential procedure to determine the number of factors,” *Journal of the American Statistical Association*, 113(523), 1341–1349.

WATSON, G. S. (1983): *Statistics on Spheres*. Wiley, New York.

Appendices

We use the following notation. $\otimes$ denotes the Kronecker product. $\|A\| = \sqrt{tr(A'A)}$ denote the Frobenius norm of matrix A. We denote by $\|Z\|_p = (E[\|Z\|^p])^{1/p}$ the L^p-norm of random matrix Z, for $p > 0$. We denote by $\xrightarrow{d}$ convergence in distribution. For a sigma-field $\mathcal{F}$, we denote by $Z_n \xrightarrow{d} Z$ ($\mathcal{F}$-stably) the stable convergence on $\mathcal{F}$ of a sequence of random vectors, that is, $P(Z_n \in A, U) \to P(Z \in A, U)$ as $n \to \infty$, for any Borel set A with $P(Z \in \partial A) = 0$, where ∂A is the boundary of set A, and any measurable set $U \in \mathcal{F}$ (see e.g. Renyi (1963), Aldous and Eagleson (1963), Hall and Heyde (1980), Kuersteiner and Prucha (2013)). In particular, for a symmetric positive definite random matrix Ω measurable with respect to $\mathcal{F}$, by $Z_n \xrightarrow{d} N(0, \Omega)$ ($\mathcal{F}$-stably) we mean $Z_n \xrightarrow{d} \Omega^{1/2}\varepsilon$ ($\mathcal{F}$-stably), where $\varepsilon \sim N(0, I)$ is independent of $\mathcal{F}$.

A Assumptions

We make the following assumptions:

Assumption A.1. *$N, T \to \infty$ with $\sqrt{T}/N = o(1)$ and $N/T^{3/2} = o(1)$.*

Assumption A.2. *The "rotated" factor VAR $f_t = \Phi f_{t-1} + v_t$ is stationary and satisfies the normalization restrictions*

$$E(f_t) = 0 \qquad \text{and} \qquad V(f_t) = E(f_t f_t') = I_r;$$

its innovations are such that $v_t \sim iid(0, \Sigma_v)$, where the (r,r) covariance matrix is defined as $\Sigma_v = diag(\sigma_1^2, \ldots, \sigma_q^2, 0, \ldots, 0)$, with $\sigma_1^2 \geq \sigma_2^2 \geq \cdots \geq \sigma_q^2 > 0$, and $E[\|v_t\|^4] \leq M$, for a constant $M < \infty$.

Assumption A.3. *The loadings matrix $\Lambda = [\ \lambda_1,\ \ldots,\ \lambda_N\ ]'$ is such that $\lim_{N\to\infty} \frac{1}{N}\Lambda'\Lambda = \Sigma_\lambda$, where Σ_λ is a positive-definite (r,r) matrix with distinct eigenvalues.*

Assumption A.4. *The error terms $\varepsilon_{i,t}$ and the factors f_t are such that for all $i, t \geq 1$: a) $E[\varepsilon_{i,t}|\mathcal{F}_t] = 0$ and $E[\varepsilon_{i,t}^2|\mathcal{F}_t] \leq M$, a.s., where $\mathcal{F}_t = \sigma(F_s, s \leq t)$, b) $E[\varepsilon_{i,t}^8] \leq M$ and $E[\|v_t\|^{2r\vee 8}] \leq M$, for a constant $M < \infty$, where $r > 2$ is defined in Assumption A.5 b).*

Assumption A.5. *Define the variables $\xi_t = \frac{1}{\sqrt{N}}\sum_{i=1}^N \lambda_i \varepsilon_{i,t}$ and $\kappa_t = \frac{1}{\sqrt{N}}\sum_{i=1}^N (\varepsilon_{i,t}^2 - \eta_t^2)$, indexed by N, where $\eta_t^2 = \operatorname*{plim}_{N\to\infty} \frac{1}{N}\sum_{i=1}^N E[\varepsilon_{i,t}^2|\mathcal{F}_t]$. a) For any $t \geq 1$ and $h \geq 0$ have:*

$$[\xi_t', \xi_{t-h}']' \xrightarrow{d} N(\,0\,, \mathcal{V}_t(h)\,), \quad (\mathcal{F}_t\text{-stably}),$$

as $N \to \infty$, where the asymptotic variance matrix is:

$$\mathcal{V}_t(h) = \begin{bmatrix} \Omega_t(0|\mathcal{F}_t) & \Omega_t(h|\mathcal{F}_t) \\ & \Omega_{t-h}(0|\mathcal{F}_t) \end{bmatrix},$$

for $\Omega_t(h|\mathcal{F}_t) = \operatorname*{plim}_{N\to\infty} \frac{1}{N}\sum_{i=1}^N\sum_{\ell=1}^N \lambda_i\lambda_\ell' cov(\varepsilon_{i,t}, \varepsilon_{\ell,t-h}|\mathcal{F}_t)$, for any h.
Moreover, for $N \geq 1$ we have: b) $E(\|\xi_t\|^{2r}|\mathcal{F}_t) \leq M$, a.s., and c) $E\left[|\kappa_t|^4\right] \leq M$, for constants $M < \infty$ and $r > 2$.

Assumption A.6. *Innovations ϵ_t and v_t are such that: a) The triangular array processes $V_t \equiv V_{N,t} = [f_t', \xi_t']'$ and $V_t^* \equiv V_{N,t}^* = [\kappa_t, \eta_t^2]'$ are strong mixing of size $-\frac{r}{r-2}$, uniformly in $N \geq 1$.* [26] *Moreover, b) $\|E(\xi_t\xi_t'|\mathcal{F}_t) - E(\xi_t\xi_t'|F_t, ..., F_{t-m})\|_2 = O(m^{-\psi})$, $\|E(\xi_{t-1}\xi_t'|\mathcal{F}_t) - E(\xi_{t-1}\xi_t'|F_t, ..., F_{t-m})\|_2 = O(m^{-\psi})$ and $\|E(\xi_{t-1}\xi_t'|\mathcal{F}_t) - E(\xi_{t-1}\xi_t'|F_t, ..., F_{t-m})\|_2 = O(m^{-\psi})$ as $m \to \infty$, uniformly in $N \geq 1$, and with $\psi > 1$.*

Assumption A.7. *Innovations ϵ_t and v_t are such that: a) $\frac{1}{T}\sum_{t=1}^T\sum_{s=1}^{t-1} E[\eta_{ts}^4] \leq M$, $E\left[\left(\frac{1}{\sqrt{N}}\sum_{i=1}^N(\varepsilon_{i,t}\varepsilon_{i,s} - \eta_{ts}^2)\right)^2\right] \leq M$, for any $s < t$ and a constant M, where $\eta_{ts}^2 = \operatorname*{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^N E[\varepsilon_{i,t}\varepsilon_{i,s}|\mathcal{F}_t]$; b) $\frac{1}{\sqrt{T}}\sum_{t=1}^T(1+\eta_t^2)f_t\alpha_t' = O_p(1)$, $\frac{1}{T}\sum_{t=1}^T \xi_t\alpha_t' = o_p(1)$, $E[\|\alpha_t\|^2] = O(1)$, where $\alpha_t = \frac{1}{\sqrt{NT}}\sum_{i=1}^N\sum_{s=1,s\neq t}^T \varepsilon_{i,t}\varepsilon_{i,s}f_s$;*
c) $E[\|\beta_t\|^2] = O(1)$ and $E[\|\bar{\beta}_t\|^2] = O(1)$, where $\beta_t = \frac{1}{\sqrt{NT}}\sum_{i=1}^N\sum_{s=1,s\neq t}^T \varepsilon_{i,t}(\varepsilon_{i,s}\zeta_s - E[\varepsilon_{i,s}\zeta_s])$ and $\bar{\beta}_t = \frac{1}{T}\sum_{i=1}^N\sum_{s=1,s\neq t}^T \varepsilon_{i,t}E[\varepsilon_{i,s}\zeta_s]$, where $\zeta_t = (\eta_t^2 f_t', \kappa_t f_t', \xi_t', \alpha_t')'$.

[26] That is, $\alpha(h) = O(h^{-\phi})$ for some $\phi > \frac{r}{r-2}$, where $\alpha(h) = \sup_{N\geq 1}\sup_{t\geq 1}\sup_{A\in\mathcal{V}_{-\infty}^t, B\in\mathcal{V}_{t+h}^\infty} |P(A\cap B) - P(A)P(B)|$, where $\mathcal{V}_{t-m}^{t+m} = \sigma(V_s, t-m \leq s \leq t+m)$, and similarly for V_t^*.

Assumption A.8. *Innovations ϵ_t and v_t are such that: a) $P[\|f_t\| \geq \delta] \leq c_1 \exp(-c_2\delta^b)$, for large δ; b) $\sum_{\ell=1:\ell\neq i}^{N} E[\varepsilon_{\ell,t}\varepsilon_{i,t}] \leq M$, for all $i \geq 1$; c) $P[\|\frac{1}{T}\sum_{t=1}^{T} z_{i,t}\| \geq \delta] \leq c_1 T \exp(-c_2\delta^2 T^{\eta}) + c_3 T\delta^{-1} \exp(-c_4 T^{\bar{\eta}})$, for all $i \geq 1$ and $\delta > 0$, where either $z_{i,t} = f_t\varepsilon_{i,t}$, or $z_{i,t} = \varepsilon_{i,t}^2 - E[\varepsilon_{i,t}^2]$, or $z_{i,t} = \frac{1}{\sqrt{N}}\sum_{\ell=1:\ell\neq i}^{N} \varepsilon_{\ell,t}\varepsilon_{i,t} - E[\frac{1}{\sqrt{N}}\sum_{\ell=1:\ell\neq i}^{N} \varepsilon_{\ell,t}\varepsilon_{i,t}]$; d) $\|\lambda_i\| \leq M$, for all $i \geq 1$; where $b, c_1, c_2, c_3, c_4, \eta, \bar{\eta}, M > 0$ are constants, and $\eta \geq 1/2$.*

Assumption A.9. *The error terms are such that: a) $Cov(\varepsilon_{i,t}, \varepsilon_{\ell,t-h}|\mathcal{F}_t) = 0$, if $i \neq \ell$, b) $E[\varepsilon_{i,t}|\{\varepsilon_{i,t-h}\}_{h\geq 1}, \mathcal{F}_t] = 0$, c) $E[\varepsilon_{i,t}^2|\{\varepsilon_{i,t-h}\}_{h\geq 1}, \mathcal{F}_t] = \gamma_{ii}$, say, where $\gamma_{ii} > 0$, for all i, t, h.*

Assumption A.1 defines the asymptotic scheme. Assumption A.2 concerns the stationarity of the factors' VAR, as well as their first and second moments. Note that the same properties hold true also in the unrotated factor space. That is, the factor VAR $\breve{f}_t = \breve{\Phi}\breve{f}_{t-1} + \breve{v}_t$ is stationary and such that:

$$E(\breve{f}_t) = 0 \qquad \text{and} \qquad V(\breve{f}_t) = E(\breve{f}_t\breve{f}_t') = I_r.$$

Moreover, VAR innovations $\breve{v}_t$, are such that $\breve{v}_t \sim iid(0, \breve{\Sigma}_v)$, for $\breve{\Sigma}_v$ as defined in Section 2, and $E[\|\breve{v}_t\|^4] \leq M$, for a constant $M < \infty$. All remaining assumptions are the same as in AGGR, and we refer to their Appendix A for a detailed discussion of each of them and their relationship with analogous assumptions made in the literature.

Assumption A.9 simplifies the derivation of the feasible asymptotic distribution of the statistic in Theorem 2. This condition excludes correlation of the error terms across individuals and time (conditional on the factors), as well as conditional heteroschedasticity, and implies a "strict factor model" for each group. In that sense, it is more restrictive than Assumptions A.5, A.6, A.7 and A.8 *b)-c)*. Moreover, under Assumption A.9, the matrix $\Omega_t(0|\mathcal{F}_t)$ in Assumption A.5 *a)* simplifies to $\Omega = \lim_{N\to\infty}(1/N)\sum_{i=1}^{N} \lambda_i\lambda_i'\gamma_{ii}$, while $\Omega_t(h) = 0$ if $h \neq 0$. Note that Assumption A.9 simplifies substantially the proof of Theorems 2 but is not needed to prove Theorem 1.

B Proofs

Section B.1 presents the proof of Proposition 1, Section B.2 and Section B.3 present the proof of Theorem 2 and Theorem 1, respectively. The proofs of technical Lemmas are reported in Section C.1 of the OA.

Let us provide some fundamental moments of the rotated static factors f_t which will turn out to be useful in the following proofs. We define $V_{11} := E(f_{t-1}f_{t-1}')$, $V_{22} := E(f_tf_t')$, and $V_{12} := E(f_{t-1}f_t') = V_{21}'$. Stationarity of the process for factors $\breve{f}_t$ from Assumption A.2 i) implies that also the factor process of f_t in (2.10) is stationary, and that $V_{22} = V_{11}$, irrespectively of the normalization in (2.5). Moreover, as W is an orthonormal (r, r) eigenvector matrix, from the normalization in (2.5) it follows that $V_{11} = I_r$.

B.1 Proof of Proposition 1

In Section B.1.1 we characterize the eigenvalues and eigenvectors of the population covariance matrix $\Sigma_v = V(v_t)$ and its OLS estimator $\tilde{\Sigma}_v = \sum_{t=1}^{T} \tilde{v}_t\tilde{v}_t'/T$. In particular, we show that their smallest $r-q$ eigenvalues are all equal to zero. Then, in Section B.1.2 we prove that the $r-q$ smallest eigenvalues of the sample covariance matrix $\tilde{\tilde{\Sigma}}_v$ are equal to zero. These two results imply part (i) of Proposition 1. Section B.1.3 shows that the largest q eigenvalues converge, as $T \to \infty$, to the q non-zero eigenvalues of $\tilde{\Sigma}_v$, which, together with the result in part (i), proves part (ii) of Proposition 1. Finally, we derive the asymptotic distribution of the largest q eigenvalues of $\tilde{\tilde{\Sigma}}_v$, which corresponds to part (iii) of Proposition 1.

B.1.1 Eigendecomposition of Σ_v and $\tilde{\Sigma}_v$

Define the following two matrices:

$$\underset{(r\times q)}{E_H} = \begin{bmatrix} I_q \\ 0_{(r-q,q)} \end{bmatrix}, \qquad \underset{(r\times(r-q))}{E_L} = \begin{bmatrix} 0_{(q,r-q)} \\ I_{r-q} \end{bmatrix}. \tag{B.1}$$

The columns of E_H and E_L span $\mathbb{R}^r$. Since $\Sigma_v = E[v_t v_t'] = diag(\sigma_1^2, ..., \sigma_q^2, 0, ..., 0)$, the eigenvectors associated with the $r-q$ zero eigenvalues of Σ_v are spanned by the columns of matrix E_L. Analogously, the eigenvectors associated with the q non-zero eigenvalues of Σ_v are spanned by the columns of matrix E_H. If we make the additional assumption that the q non-zero eigenvalues of Σ_v are distinct, that is $\sigma_1^2 > \sigma_2^2 > ...... > \sigma_q^2 > 0$, then the orthonormal eigenvectors associated with these eigenvalues correspond to the columns of E_H.

To characterize matrix $\tilde{\Sigma}_v$ and its sorted eigenvalues $\tilde{\sigma}_1^2 \geq \tilde{\sigma}_2^2 \geq ... \geq \tilde{\sigma}_r^2$, we first define:

$$\tilde{V}_{11} := \frac{1}{T}\sum_{t=1}^{T} f_{t-1}f_{t-1}', \qquad \tilde{V}_{22} := \frac{1}{T}\sum_{t=1}^{T} f_t f_t', \qquad \tilde{V}_{12} := \frac{1}{T}\sum_{t=1}^{T} f_{t-1}f_t', \qquad \tilde{V}_{21} := \tilde{V}_{12}', \tag{B.2}$$

and

$$\tilde{V}_{v,11} := \frac{1}{T}\sum_{t=1}^{T} v_{t-1}v_{t-1}', \qquad \tilde{V}_{v,22} := \frac{1}{T}\sum_{t=1}^{T} v_t v_t', \qquad \tilde{V}_{vf,21} := \frac{1}{T}\sum_{t=1}^{T} v_t f_{t-1}', \tag{B.3}$$

From the definition of $\tilde{\Sigma}_v$ in (3.3), and that of the OLS residuals $\tilde{v}_t = f_t - \tilde{\Phi} f_{t-1}$ obtained by OLS estimation of the VAR(1) model in equation (2.10) from the T-dimensional sample of true factors f_t, we get:

$$\tilde{\Sigma}_v = \frac{1}{T}\sum_{t=1}^{T} \tilde{v}_t\tilde{v}_t' = \frac{1}{T}\sum_{t=1}^{T}(f_t - \tilde{\Phi} f_{t-1})(f_t - \tilde{\Phi} f_{t-1})' = \tilde{V}_{22} - \tilde{V}_{21}\tilde{V}_{11}^{-1}\tilde{V}_{12}, \tag{B.4}$$

where the third equality follows from (B.2) and the definition in equation (3.2) of the OLS estimator $\tilde{\Phi}$ of Φ, i.e. $\tilde{\Phi} = \tilde{V}_{21}\tilde{V}_{11}^{-1}$. By straightforward matrix algebra, we get the next Lemma.

LEMMA B.1. *The matrix $\tilde{\Sigma}_v = \tilde{V}_{22} - \tilde{V}_{21}\tilde{V}_{11}^{-1}\tilde{V}_{12}$ is such that:*

$$\tilde{\Sigma}_v = \tilde{V}_{v,22} - \tilde{V}_{vf,21}\tilde{V}_{11}^{-1}\tilde{V}_{vf,21}' = \begin{bmatrix} \tilde{\Sigma}_{v,HH} & 0_{(q,r-q)} \\ 0_{(r-q,q)} & 0_{(r-q,r-q)} \end{bmatrix}, \tag{B.5}$$

where the (q,q) matrix $\tilde{\Sigma}_{v,HH}$ is:

$$\tilde{\Sigma}_{v,HH} = \frac{1}{T}\sum_{t=1}^{T} v_{H,t}v_{H,t}' - \left[\frac{1}{T}\sum_{t=1}^{T} v_{H,t}f_{t-1}'\left(\frac{1}{T}\sum_{t=1}^{T} f_{t-1}f_{t-1}'\right)^{-1}\frac{1}{T}\sum_{t=1}^{T} f_{t-1}v_{H,t}'\right]. \tag{B.6}$$

From Lemma B.1 it immediately follows that the $r-q$ smallest eigenvalues of $\tilde{\Sigma}_v$ are $\tilde{\sigma}_{q+1}^2 = ... = \tilde{\sigma}_r^2 = 0$. This is our first non-trivial result, which shows that the smallest $r-q$ eigenvalues of $\tilde{\Sigma}_v$ are exactly equal to the $r-q$ zero eigenvalues of the population covariance matrix Σ_v for any finite sample of dimension $T \geq r^2$. From (B.5), it follows that the orthonormal eigenvectors associated to the $r-q$ zero estimated eigenvalues of $\tilde{\Sigma}_v$ are spanned by the columns of matrix E_L. Let $\tilde{W}_{v,r-q}$ be the matrix having as columns each of the $r-q$

associated eigenvectors, then $\tilde{W}_{v,r-q} = E_L \cdot A$, where A is an $(r-q, r-q)$ orthonormal matrix, implying: $\tilde{\Sigma}_v \tilde{W}_{v,r-q} = \tilde{W}_{v,r-q} \cdot 0_{(r-q,r-q)}$, and $\tilde{W}'_{v,r-q} \tilde{W}_{v,r-q} = I_{r-q}$.

Moreover, the (q,q) matrix $\tilde{\Sigma}_{v,HH}$ in (B.6) is the sample covariance matrix of the residuals obtained from a multivariate regression of $v_{H,t}$ on (lagged) factors f_{t-1}. In fact, the estimated matrix of regression coefficients of such multivariate regression is $\sum_{t=1}^{T} v_{H,t} f'_{t-1} (\sum_{t=1}^{T} f_{t-1} f_{t-1})^{-1} = \tilde{V}_{vf,12} \tilde{V}_{11}^{-1}$, the residuals are $v_{H,t} - \tilde{V}_{vf,12} \tilde{V}_{11}^{-1} f_{t-1}$, so that the OLS residuals' covariance matrix is

$$\frac{1}{T} \sum_{t=1}^{T} (v_{H,t} - \tilde{V}_{vf,12} \tilde{V}_{11}^{-1} f_{t-1})(v_{H,t} - \tilde{V}_{vf,12} \tilde{V}_{11}^{-1} f_{t-1})' = \tilde{\Sigma}_{v,HH},$$

which is positive semi-definite, thus having all non-negative eigenvalues. Hence, the largest q eigenvalues of $\tilde{\Sigma}_v$ are such that $\tilde{\sigma}_1^2 \geq ... \geq \tilde{\sigma}_q^2 \geq 0$, which implies

$$\tilde{\sigma}_1^2 \geq ... \geq \tilde{\sigma}_q^2 \geq \tilde{\sigma}_{q+1}^2 = ...\tilde{\sigma}_r^2 = 0. \tag{B.7}$$

Equation (B.5) entails that the orthonormal eigenvectors associated to the q largest eigenvalues of $\tilde{\Sigma}_v$ are spanned by the columns of matrix E_H. Let $\tilde{W}_{v,q}$ be the matrix having as columns each of the q associated eigenvectors, then $\tilde{W}_{v,q} = E_H \cdot B$, where B is a (q,q) orthogonal matrix, implying: $\tilde{\Sigma}_v \tilde{W}_{v,q} = \tilde{W}_{v,q} \tilde{\Sigma}^{eig}_{v,HH}$, $\tilde{W}'_{v,q} \tilde{W}_{v,q} = I_q$, and $\tilde{\Sigma}^{eig}_{v,HH} := diag(\tilde{\sigma}_1^2, ..., \tilde{\sigma}_q^2)$.

Let $\tilde{\Sigma}^{eig}_v$ be the (r,r) diagonal matrix collecting the ordered eigenvalues of $\tilde{\Sigma}_v$:

$$\tilde{\Sigma}^{eig}_v := diag(\tilde{\sigma}_1^2, ..., \tilde{\sigma}_r^2) = diag(\tilde{\sigma}_1^2, ..., \tilde{\sigma}_q^2, 0, ..., 0), \tag{B.8}$$

and let $\tilde{W}_v := [\tilde{W}_{v,q}\ ,\ \tilde{W}_{v,r-q}]$ be the (r,r) matrix collecting the associated orthonormal eigenvectors. Then,

$$\tilde{\Sigma}_v \tilde{W}_v = \tilde{W}_v \tilde{\Sigma}^{eig}_v\ , \qquad \text{with} \qquad \tilde{W}'_v \tilde{W}_v = \tilde{W}_v \tilde{W}'_v = I_r. \tag{B.9}$$

B.1.2 Eigendecomposition of $\breve{\tilde{\Sigma}}_v$

Let us define

$$\breve{\tilde{V}}_{11} \ := \ \frac{1}{T} \sum_{t=1}^{T} \breve{f}_{t-1} \breve{f}'_{t-1}\ , \qquad \breve{\tilde{V}}_{22} \ := \ \frac{1}{T} \sum_{t=1}^{T} \breve{f}_t \breve{f}'_t\ , \qquad \breve{\tilde{V}}_{12} \ := \ \frac{1}{T} \sum_{t=1}^{T} \breve{f}_{t-1} \breve{f}'_t = \breve{\tilde{V}}'_{12}$$

From the definition of $\breve{\tilde{\Sigma}}_v$ in Section 3.1, and that of the OLS residuals $\breve{\tilde{v}}_t = \breve{f}_t - \breve{\tilde{\Phi}} \breve{f}_{t-1}$ obtained by estimating the VAR(1) model in equation (2.10) from the T-dimensional sample of true factors $\breve{f}_t$, we get:

$$\breve{\tilde{\Sigma}}_v \ = \ \frac{1}{T} \sum_{t=1}^{T} \breve{\tilde{v}}_t \breve{\tilde{v}}'_t \ = \ \frac{1}{T} \sum_{t=1}^{T} (\breve{f}_t - \tilde{\Phi} \breve{f}_{t-1})(\breve{f}_t - \tilde{\Phi} \breve{f}_{t-1})' = \breve{\tilde{V}}_{22} - \breve{\tilde{V}}_{21} \breve{\tilde{V}}_{11}^{-1} \breve{\tilde{V}}_{12},$$

which follows by writing $\breve{\tilde{\Phi}} = \breve{\tilde{V}}_{21} \breve{\tilde{V}}_{11}^{-1}$. Recalling that $f_t = W'_v \breve{f}_t$ and $W'_v W_v = I_r$, we get $\breve{f}_t = W_v f_t$, which implies:

$$\breve{\tilde{V}}_{11} \ = \ W'_v \tilde{V}_{11} W_v\ , \qquad \breve{\tilde{V}}_{22} \ = \ W'_v \tilde{V}_{22} W_v\ , \qquad \breve{\tilde{V}}_{12} \ = \ W'_v \tilde{V}_{12} W_v = \breve{\tilde{V}}'_{21},$$

and

$$\tilde{\tilde{\Sigma}}_v \;=\; W_v\breve{V}_{22}W_v' - W_v\tilde{V}_{21}W_v'W_v\tilde{V}_{11}^{-1}W_v'W_v\tilde{V}_{12}W_v' \;=\; W_v\tilde{\Sigma}_vW_v'. \tag{B.10}$$

Equation (B.9) implies $\tilde{\Sigma}_v = \tilde{W}_v\tilde{\Sigma}_v^{eig}\tilde{W}_v'$, and therefore $\tilde{\tilde{\Sigma}}_v = W_v\tilde{W}_v\tilde{\Sigma}_v^{eig}\tilde{W}_v'W_v'$. As $\tilde{W}_v'W_v'W_v\tilde{W}_v = I_r$, we get that

$$\tilde{\tilde{\Sigma}}_v(W_v\tilde{W}_v) \;=\; (W_v\tilde{W}_v)\tilde{\Sigma}_v^{eig},$$

i.e. $\tilde{\Sigma}_v^{eig}$ is the diagonal matrix containing the sorted eigenvalues of $\tilde{\tilde{\Sigma}}_v$, with associated orthonormal eigenvectors being the r columns of matrix $W_v\tilde{W}_v$. As $\tilde{\Sigma}_v^{eig}$ is also the matrix of the sorted eigenvalues of $\tilde{\Sigma}_v$, this proves part (i) of Proposition 1.

B.1.3 Convergence of the eigenvalues of $\tilde{\Sigma}_v$

Under Assumption A.2, a Central Limit Theorem (CLT) for iid random variables with finite second moment implies that

$$\sqrt{T}\cdot vec\left(\frac{1}{T}\sum_{t=1}^{T} v_{H,t}v_{H,t}' - \Sigma_{v,HH}\right) \xrightarrow{d} N(0,\mathcal{V}_1^*),\;\text{ where }\; \mathcal{V}_1^* := E[vec(v_tv_t'-\Sigma_v)\cdot vec(v_tv_t'-\Sigma_v)'], \tag{B.11}$$

as $T\to\infty$. Hence, $\sum_{t=1}^{T} v_{H,t}v_{H,t}'/T - \Sigma_{v,HH} = O_p(1/\sqrt{T})$. The same assumptions entails

$$\frac{1}{T}\sum_{t=1}^{T} v_{H,t}f_{t-1}' \;=\; O_p(1/\sqrt{T}), \quad \text{and} \quad \frac{1}{T}\sum_{t=1}^{T} f_{t-1}f_{t-1}' \;=\; V_{11} + O_p(1/\sqrt{T}).$$

Substituting these equations into (B.6) we get $\tilde{\Sigma}_{v,HH} = \Sigma_{v,HH} + O_p(1/\sqrt{T})$ and, by substituting the last result in (B.5), also

$$\tilde{\Sigma}_v \;=\; \begin{bmatrix} \Sigma_{v,HH} + O_p(1/\sqrt{T}) & 0_{(q,r-q)} \\ 0_{(r-q,q)} & 0_{(r-q,r-q)} \end{bmatrix}. \tag{B.12}$$

We know from the results in Subsection B.1.1 that the eigenspace associated with the zero eigenvalues of $\tilde{\Sigma}_v$ has dimension $r-q$ and is spanned by the columns of matrix E_L. Therefore, from (B.8), (B.9) and (B.12), we can write the following expansions for the eigenvalue matrix $\tilde{\Sigma}_{v,HH}^{eig}$ and the associated eigenvector matrix $\tilde{W}_{v,q}$:

$$\tilde{W}_{v,q} = E_H\mathcal{U}_{v,q}, \qquad \tilde{\Sigma}_{v,HH}^{eig} = \Sigma_{v,HH} + M_{v,q} = diag(\tilde{\sigma}_1^2,...,\tilde{\sigma}_q^2) + M_{v,q},$$

where E_H is defined in equation (B.1), the stochastic (q,q) matrix $\mathcal{U}_{v,q}$ is nonsingular with probability approaching (w.p.a.) 1 and stochastic matrix $M_{v,q}$ is diagonal. By the continuity of the matrix eigenvalue and eigenfunction mappings, as $\tilde{\Sigma}_v = \Sigma_v + O_p(1/\sqrt{T})$, the largest q eigenvalues and the associated eigenvectors converge at the same rate to the eigenvalues and eigenvectors of Σ_v. Therefore, $M_{v,q}$ converges in probability to a null matrix as $T\to\infty$ at rate $O_p(1/\sqrt{T})$, that is $M_{v,q} = 0_{q,q} + O_p(1/\sqrt{T})$, which implies:

$$\tilde{\sigma}_\ell^2 = \sigma_\ell^2 + O_p(1/\sqrt{T}), \qquad \ell = 1,...,q,$$

and $\mathcal{U}_{v,q}$ converges in probability to a non-singular matrix at the same rate. Therefore, from the set of inequalities (B.7), the definition $\Sigma_{v,HH} = diag(\sigma_1^2, ..., \sigma_q^2)$, and the assumption $\sigma_1^2 \geq \sigma_2^2 \geq ... \geq \sigma_q^2 > 0$, we have

$$\tilde{\sigma}_1^2 \geq ... \geq \tilde{\sigma}_q^2 > \tilde{\sigma}_{q+1}^2 = \ ... \ = \tilde{\sigma}_r^2 = 0 \qquad w.p.a.\ 1,$$

at $T \to \infty$, which concludes the proof of part (ii) of Proposition 1.

To simplify the proof of the convergence in distribution of the largest q eigenvalues of $\tilde{\Sigma}_v$, that is the proof of part (iii) of Proposition 1, we assume that all the q non-zero eigenvalues of Σ_v are distinct, namely:

$$\sigma_1^2 > ... > \sigma_q^2 > 0.$$

This assumption implies that the orthonormal eigenvectors associated with the largest q non-zero eigenvalues of Σ_v are given exactly by the first q columns of matrix I_q. We denote each of these columns, i.e. each one of these eigenvectors, as $e_{q,\ell} := [0, ..., 0, 1, 0, ..., 0]'$, which is a q-dimensional vector of zeros, with the exception of the element in row ℓ which is equal to 1, with $\ell = 1, ..., q$. This implies $I_q = [e_{q,1}, ..., e_{q,q}]$.

By noting that

$$\begin{aligned}
\tilde{V}_{12} &= \frac{1}{T}\sum_{t=1}^{T} f_{t-1}f_t' = \frac{1}{T}\sum_{t=1}^{T} f_{t-1}(\Phi f_{t-1} + v_t)' = \tilde{V}_{11}\Phi' + \frac{1}{T}\sum_{t=1}^{T} f_{t-1}v_t' \\
&= \tilde{V}_{11}\Phi' + \left[\ \tfrac{1}{T}\textstyle\sum_{t=1}^{T} f_{t-1}v_{H,t}', \ 0_{(r,r-q)} \ \right]
\end{aligned}$$

we get

$$\tilde{\Phi} = \tilde{V}_{21}\tilde{V}_{11}^{-1} = \Phi + \begin{bmatrix} \frac{1}{T}\sum_{t=1}^{T} v_{H,t}f_{t-1}' \\ 0_{(r-q,r)} \end{bmatrix} \tilde{V}_{11}^{-1}.$$

Assumption (A.2) implies that $\frac{1}{T}\sum_{t=1}^{T} v_{H,t}f_{t-1} = O_p(1/\sqrt{T})$. Hence $\tilde{\Phi}_{HL} = \Phi_{HL} + O_p(1/\sqrt{T})$, $\tilde{\Phi}_{HH} = \Phi_{HH} + O_p(1/\sqrt{T})$, and

$$\begin{aligned}
\tilde{v}_{H,t} &= f_{H,t} - \tilde{\Phi}_{HH}f_{H,t-1} - \tilde{\Phi}_{HL}f_{L,t-1} \\
&= (f_{H,t} - \Phi_{HH}f_{H,t-1} - \Phi_{HL}f_{L,t-1}) - (\tilde{\Phi}_{HH} - \Phi_{HH})f_{H,t-1} - (\tilde{\Phi}_{HL} - \Phi_{HL})f_{L,t-1} \\
&= v_{H,t} + O_p(1/\sqrt{T}).
\end{aligned}$$

Thus, $\frac{1}{T}\sum_{t=1}^{T} \tilde{v}_{H,t}\tilde{v}_{H,t}' = \frac{1}{T}\sum_{t=1}^{T} v_{H,t}v_{H,t}' + O_p(1/T)$, which together with (B.11) implies:

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T} \tilde{v}_{H,t}\tilde{v}_{H,t}' - \Sigma_{v,HH} &= \frac{1}{T}\sum_{t=1}^{T} v_{H,t}v_{H,t}' - \Sigma_{v,HH} + \left(\frac{1}{T}\sum_{t=1}^{T} \tilde{v}_{H,t}\tilde{v}_{H,t}' - \frac{1}{T}\sum_{t=1}^{T} v_{H,t}v_{H,t}' \right) \\
&= \frac{1}{T}\sum_{t=1}^{T} \left(v_{H,t}v_{H,t}' - \Sigma_{v,HH} \right) + o_p\left(\frac{1}{\sqrt{T}} \right).
\end{aligned}$$

and

$$\sqrt{T} \cdot vec\left(\frac{1}{T}\sum_{t=1}^{T} \tilde{v}_{H,t}\tilde{v}_{H,t}' - \Sigma_{v,HH} \right) \xrightarrow{d} N(0, \mathcal{V}_1^*), \tag{B.13}$$

where $\mathcal{V}_1^*$ is defined in (B.11).

Then, from the result on the asymptotic distribution of eigenvalues and eigenvectors of symmetric random matrices in Section 1 of Ruymgaart and Yang (1997), which was originally derived in Watson (1983) using Kato's (1966) perturbation theory, the convergence result in (B.13), and the assumption that the eigenvalues σ_ℓ^2, $\ell = 1, ..., q$ are all distinct, it follows that:

$$\begin{aligned}\sqrt{T}(\tilde{\sigma}_\ell^2 - \sigma_\ell^2) &= tr\left\{\sqrt{T}\cdot\left(\frac{1}{T}\sum_{t=1}^{T}\tilde{v}_{H,t}\tilde{v}'_{H,t} - \Sigma_{v,HH}\right)e_{q,\ell}e'_{q,\ell}\right\} + o_p(1)\\ &= \sqrt{T}\cdot e'_{q,\ell}\left(\frac{1}{T}\sum_{t=1}^{T}\tilde{v}_{H,t}\tilde{v}'_{H,t} - \Sigma_{v,HH}\right)e_{q,\ell} + o_p(1)\end{aligned}$$

which can also be written as

$$\begin{aligned}\sqrt{T}(\tilde{\sigma}_\ell^2 - \sigma_\ell^2) &= \sqrt{T}e'_{q,\ell}\left(\frac{1}{T}\sum_{t=1}^{T}\tilde{v}_{H,t}\tilde{v}'_{H,t} - \Sigma_{v,HH}\right)e_{q,\ell} + o_p(1)\\ &= e'_{q,\ell}(e'_{q,\ell}\otimes I_m)\sqrt{T}\cdot vec\left(\frac{1}{T}\sum_{t=1}^{T}\tilde{v}_{H,t}\tilde{v}'_{H,t} - \Sigma_{v,HH}\right) + o_p(1), \qquad \text{(B.14)}\end{aligned}$$

where the last result follows from the equality $tr(\mathcal{ABC}) = vec(\mathcal{A}')'(\mathcal{C}'\otimes I.)vec(\mathcal{B})$, where $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ are conformable matrices. Therefore, results (B.13) and (B.14) imply

$$\sqrt{T}(\tilde{\sigma}_\ell^2 - \sigma_\ell^2) \xrightarrow{d} N(0, \mathcal{V}_2^*), \qquad \text{where} \quad \mathcal{V}_2^* := e'_{q,\ell}(e'_{q,\ell}\otimes I_q)\cdot\mathcal{V}_1^*\cdot(e_{q,\ell}\otimes I_m)e_{q,\ell}.$$

which concludes the proof of part (iii) of Proposition 1.

B.2 Proof of Theorem 1

The proof of Theorem 1 is structured as follows. We start by reporting the asymptotic expansion for PC estimator of the static factors (Subsection B.2.1). This first result result yields an asymptotic expansion for the VAR residual matrix $\hat{\Sigma}_v$ (Subsection B.2.2), and is used to obtain the asymptotic expansions of the eigenvalues and eigenvectors of matrix $\hat{\Sigma}_v$ by perturbation methods (Subsection B.2.3). This yields the asymptotic expansions of the eigenvalues and of the test statistic $\hat{\xi}(q)$ (Subsection B.2.4). Finally, the asymptotic Gaussian distribution of the test statistic follows by applying a suitable CLT for dependent triangular arrays (Subsection B.2.5).

B.2.1 Asymptotic expansion of the factor estimates $\hat{f}_t$

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

$$\hat{\breve{f}}_t = \hat{\mathcal{H}}(\breve{f}_t + \breve{\psi}_t), \qquad \breve{\psi}_t := \frac{1}{\sqrt{N}}\breve{u}_t + \frac{1}{T}\breve{b}_t + \frac{1}{\sqrt{NT}}\breve{d}_t + \breve{\vartheta}_t, \qquad \text{(B.15)}$$

for $t = 1, \dots, T$, *where* $\breve{u}_t = \left(\frac{1}{N}\sum_{i=1}^{N}\breve{\lambda}_i\breve{\lambda}'_i\right)^{-1}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\breve{\lambda}_i\varepsilon_{i,t}$,
$\breve{b}_t = \left(\frac{1}{N}\sum_{i=1}^{N}\breve{\lambda}_i\breve{\lambda}'_i\right)^{-1}\left(\frac{1}{T}\sum_{t=1}^{T}\breve{f}_t\breve{f}'_t\right)^{-1}(\eta_t^*)^2\breve{f}_t$,
$\breve{d}_t = \left(\frac{1}{N}\sum_{i=1}^{N}\breve{\lambda}_i\breve{\lambda}'_i\right)^{-1}\left(\frac{1}{T}\sum_{t=1}^{T}\breve{f}_t\breve{f}'_t\right)^{-1}\left(\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{s=1}^{T}\varepsilon_{i,s}\breve{f}_s\breve{\lambda}'_i\right)\breve{f}_t$, *and terms* $\breve{\vartheta}_t$ *are such that* $\frac{1}{T}\sum_{t=1}^{T}\left(\frac{1}{\sqrt{N}}\breve{u}_t + \frac{1}{T}\breve{b}_t + \frac{1}{\sqrt{NT}}\breve{d}_t + \breve{\vartheta}_t\right)\breve{\vartheta}'_t = o_p\left(\frac{1}{N\sqrt{T}}\right)$ *and* $\frac{1}{T}\sum_{t=1}^{T}\breve{f}_t\breve{\vartheta}'_t = O_p\left(\frac{1}{N} + \frac{1}{T^2}\right)$ *as* $N, T \to \infty$,

the matrix $\hat{\mathcal{H}}$ *converges in probability to a nonstochastic orthogonal* (r,r) *matrix, and* $\eta_t^* = \left(\underset{N\to\infty}{plim} \frac{1}{N}\sum_{i=1}^N E[\varepsilon_{i,t}^2|\mathcal{F}_t]\right)^{1/2}$.

Proposition B.1 corresponds to Proposition 3 in AGGR and the proof is omitted. From the definitions of $f_t := W_v' \breve{f}_t$ and $\lambda_i = W_v'\breve{\lambda}_i$, and the fact that W_v is an orthonormal (r,r) matrix, the next Corollary B.1 follows immediately from Proposition B.1.

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

$$\hat{\breve{f}}_t = \hat{\mathcal{H}} W_v (f_t + \psi_t), \qquad \psi_t := \frac{1}{\sqrt{N}} u_t + \frac{1}{T} b_t + \frac{1}{\sqrt{NT}} d_t + \vartheta_t, \tag{B.16}$$

for $t = 1,\dots,T$, *where* $u_t = \left(\frac{1}{N}\sum_{i=1}^N \lambda_i\lambda_i'\right)^{-1} \frac{1}{\sqrt{N}}\sum_{i=1}^N \lambda_i \varepsilon_{i,t}$,
$b_t = \left(\frac{1}{N}\sum_{i=1}^N \lambda_i\lambda_i'\right)^{-1} \left(\frac{1}{T}\sum_{t=1}^T f_t f_t'\right)^{-1} (\eta_t^*)^2 f_t$,
$d_t = \left(\frac{1}{N}\sum_{i=1}^N \lambda_i\lambda_i'\right)^{-1} \left(\frac{1}{T}\sum_{t=1}^T f_t f_t'\right)^{-1} \left(\frac{1}{\sqrt{NT}}\sum_{i=1}^N\sum_{s=1}^T \varepsilon_{i,s} f_s \lambda_i'\right) f_t$, *and terms* ϑ_t *are such that* $\frac{1}{T}\sum_{t=1}^T \left(\frac{1}{\sqrt{N}}u_t + \frac{1}{T}b_t + \frac{1}{\sqrt{NT}}d_t + \vartheta_t\right)\vartheta_t' = o_p\left(\frac{1}{N\sqrt{T}}\right)$ *and* $\frac{1}{T}\sum_{t=1}^T f_t\vartheta_t' = O_p\left(\frac{1}{N}+\frac{1}{T^2}\right)$ *as* $N,T\to\infty$, *and* η_t^* *is defined in Proposition B.1.*

B.2.2 Asymptotic expansion of matrix $\hat{\breve{\Sigma}}_v$

We can re-write matrix $\hat{\breve{\Sigma}}_v$ by using the following quantities:

$$\hat{\breve{V}}_{11} = \frac{1}{T}\sum_{t=1}^T \hat{\breve{f}}_{t-1}\hat{\breve{f}}_{t-1}' = \frac{1}{T}\sum_{t=1}^T \hat{\mathcal{H}} W_v (f_{t-1}+\psi_{t-1})(f_{t-1}+\psi_{t-1})' W_v'\hat{\mathcal{H}}' = \hat{\mathcal{H}} W_v (\tilde{V}_{11} + \hat{X}_{11}) W_v' \hat{\mathcal{H}}', \tag{B.17}$$

$$\hat{\breve{V}}_{22} = \frac{1}{T}\sum_{t=1}^T \hat{\breve{f}}_{t}\hat{\breve{f}}_{t}' = \frac{1}{T}\sum_{t=1}^T \hat{\mathcal{H}} W_v (f_{t}+\psi_{t})(f_{t}+\psi_{t}) W_v'\hat{\mathcal{H}}' = \hat{\mathcal{H}} W_v (\tilde{V}_{22} + \hat{X}_{22}) W_v' \hat{\mathcal{H}}', \tag{B.18}$$

$$\hat{\breve{V}}_{12} = \frac{1}{T}\sum_{t=1}^T \hat{\breve{f}}_{t-1}\hat{\breve{f}}_{t}' = \frac{1}{T}\sum_{t=1}^T \hat{\mathcal{H}} W_v (f_{t-1}+\psi_{t-1})(f_{t}+\psi_{t})' W_v'\hat{\mathcal{H}}' = \hat{\mathcal{H}} W_v (\tilde{V}_{12} + \hat{X}_{12}) W_v' \hat{\mathcal{H}}',$$
$$\hat{\breve{V}}_{21} = \hat{\breve{V}}_{12}', \tag{B.19}$$

where $\tilde{V}_{ij}$ for $i,j=1,2$ are defined in (B.2) and

$$\hat{X}_{11} = \frac{1}{T}\sum_{t=1}^T (f_{t-1}\psi_{t-1}' + \psi_{t-1}f_{t-1}') + \frac{1}{T}\sum_{t=1}^T \psi_{t-1}\psi_{t-1}', \tag{B.20}$$

$$\hat{X}_{22} = \frac{1}{T}\sum_{t=1}^T (f_{t}\psi_{t}' + \psi_{t}f_{t}') + \frac{1}{T}\sum_{t=1}^T \psi_{t}\psi_{t}', \tag{B.21}$$

$$\hat{X}_{12} = \frac{1}{T}\sum_{t=1}^T (f_{t-1}\psi_{t}' + \psi_{t-1}f_{t}') + \frac{1}{T}\sum_{t=1}^T \psi_{t-1}\psi_{t}', \qquad \hat{X}_{21} = \hat{X}_{12}'. \tag{B.22}$$

From the definition of matrix $\hat{\hat{\Sigma}}_v = \frac{1}{T}\sum_{t=1}^{T} \hat{\hat{v}}_t \hat{\hat{v}}_t'$ and $\hat{\hat{\Phi}} := \hat{\hat{V}}_{21}\hat{\hat{V}}_{11}^{-1}$ from definitions (B.17)-(B.19) we get:

$$\hat{\hat{\Sigma}}_v \;=\; \frac{1}{T}\sum_{t=1}^{T} \hat{\hat{v}}_t \hat{\hat{v}}_t' \;=\; \frac{1}{T}\sum_{t=1}^{T} (\hat{\hat{f}}_t - \hat{\hat{\Phi}}\hat{\hat{f}}_{t-1})(\hat{\hat{f}}_t - \hat{\hat{\Phi}}\hat{\hat{f}}_{t-1})' = \hat{\hat{V}}_{22} - \hat{\hat{V}}_{21}\hat{\hat{V}}_{11}^{-1}\hat{\hat{V}}_{12}.$$

Then, using $\hat{\hat{V}}_{11}^{-1} = (\hat{\mathcal{H}}')^{-1} W_v (\tilde{V}_{11} + \hat{X}_{11})^{-1} W_v' \hat{\mathcal{H}}^{-1} = (\hat{\mathcal{H}}')^{-1} W_v \tilde{V}_{11}^{-1}\left(I_r + \hat{X}_{11}\tilde{V}_{11}^{-1}\right)^{-1} W_v'\hat{\mathcal{H}}^{-1}$, we get:

$$\hat{\hat{\Sigma}}_v \;=\; \hat{\mathcal{H}} W_v \hat{\Sigma}_v W_v' \hat{\mathcal{H}}', \tag{B.23}$$

where:

$$\hat{\Sigma}_v \;:=\; \tilde{V}_{22} + \hat{X}_{22} - (\tilde{V}_{21} + \hat{X}_{21})\tilde{V}_{11}^{-1}\left(I_r + \hat{X}_{11}\tilde{V}_{11}^{-1}\right)^{-1}(\tilde{V}_{12} + \hat{X}_{12}). \tag{B.24}$$

When factors are estimated by PCA, the OLS estimators $\hat{\hat{\Sigma}}_v$ and $\hat{\Sigma}_v$ depend on: i) estimators of the factors' covariance matrices in the observable case through the terms $\tilde{V}_{ij}$ for $i, j = 1, 2$; and ii) the principal components' estimation error through the terms $\hat{X}_{ij}$ for $i, j = 1, 2$.

By using the definition of ψ_t in Corollary B.1, the next Lemma provides a stochastic rate for terms $\hat{X}_{jk}$.

LEMMA B.2. *Under Assumptions A.1-A.4, A.5 b)-c), A.6 a) and A.7 we have $\hat{X}_{jk} = O_p(\delta_{N,T})$, for $j, k = 1, 2$, where $\delta_{N,T} := (\min\{N, T\})^{-1}$.*

LEMMA B.3. *Under Assumptions A.1-A.4, A.5 b)-c), A.6 a) and A.7, the second-order asymptotic expansion of matrix $\hat{\hat{\Sigma}}_v$ is:*

$$\hat{\hat{\Sigma}}_v = \hat{\mathcal{H}} W_v (\tilde{\Sigma}_v + \hat{\Psi}) W_v' \hat{\mathcal{H}}' + O_p(\delta_{N,T}^2), \tag{B.25}$$

where $\tilde{\Sigma}_v$ is defined in equation (B.4)*, and $\hat{\Psi} := \hat{X}_{22} - \tilde{\Phi}\hat{X}_{12} - \hat{X}_{21}\tilde{\Phi}' + \tilde{\Phi}\hat{X}_{11}\tilde{\Phi}'$, with $\tilde{\Phi} \;=\; \tilde{V}_{21}\tilde{V}_{11}^{-1}$.*

Equation (B.25) represents matrix $\hat{\hat{\Sigma}}_v$ as a function of i) the estimated VAR errors' matrix $\breve{\tilde{\Sigma}} \;=\; W_v \tilde{\Sigma}_u W_v'$ defined in (B.10) and computed with the true rotated factor values $\breve{f}_t \;=\; W_v f_t$; ii) a second term $\hat{\Psi}$ and a reminder $O_p(\delta_{N,T}^2)$, both originating from the PC estimation error.

B.2.3 Eigenvalues and eigenvectors of matrix $\hat{\hat{\Sigma}}_v$ obtained by perturbation methods

The estimators of the smallest $r-q$ zero eigenvalues of $\hat{\hat{\Sigma}}_v$ are denoted as $\hat{\sigma}_\ell^2$, for $\ell = q+1, ..., r$. We now derive their asymptotic expansion under the null hypothesis $H(q)$ using perturbations arguments applied to equation (B.25). Let $\hat{W}_{v,r-q}^*$ be a $(r, r-q)$ matrix whose columns are eigenvectors of matrix $\hat{\hat{\Sigma}}_v$ associated with the eigenvalues $\hat{\sigma}_\ell^2$, with $\ell = q+1, ..., r$. We have:

$$\hat{\hat{\Sigma}}_v \hat{W}_{v,r-q}^* \;=\; \hat{W}_{v,r-q}^* \hat{\Lambda}_v, \tag{B.26}$$

where $\hat{\Lambda}_v \;=\; diag(\hat{\sigma}_\ell^2, \ell = q+1, ..., r)$ is the $(r-q, r-q)$ diagonal matrix containing the $r-q$ smallest eigenvalues of $\hat{\hat{\Sigma}}_v$.

We know from Subsection B.1.1 that the eigenspace associated with the zero eigenvalues of $\tilde{\Sigma}_v$ has dimension $r-q$ and is spanned by the columns of matrix E_L, which implies that the eigenspace associated with the zero eigenvalues of $\hat{\mathcal{H}} W_v \tilde{\Sigma}_v W_v' \hat{\mathcal{H}}'$ has also dimension $r-q$ and is spanned by the columns of matrix $(\hat{\mathcal{H}}')^{-1} W_v E_L$.[27] Since the columns of E_L and E_H span $\mathbb{R}^r$, (B.25) entails the following expansions:

$$\hat{W}^*_{v,r-q} = (\hat{\mathcal{H}}')^{-1} W_v [E_L \hat{\mathcal{U}} + E_H \hat{\alpha}], \qquad \hat{\Lambda}_v = 0_{(r-q,r-q)} + \hat{M}, \tag{B.27}$$

where E_L and E_H are defined in equation (B.1), the stochastic $(r-q, r-q)$ matrix $\hat{\mathcal{U}}$ is nonsingular with probability approaching (w.p.a.) 1, stochastic matrix $\hat{M}$ is diagonal, and $\hat{\alpha}$ is a $(q, r-q))$ stochastic matrix. By the continuity of the matrix eigenvalue and eigenfunction mappings, and Lemma B.2, we have that $\hat{\alpha}$ and $\hat{M}$ converge in probability to null matrices as $N, T \to \infty$ at rate $O_p(\delta_{N,T})$. By substituting the expansions (B.25) and (B.27) into the eigenvalue-eigenvector equation (B.26), using the characterization of matrix $\tilde{\Sigma}_v$ obtained in Lemma B.1, and keeping terms up to order $O_p(\delta^2_{N,T})$, we get expressions for matrices $\hat{\alpha}$ and $\hat{M}$. These yield the asymptotic expansions of the smallest $r-q$ eigenvalues and associated eigenvectors of matrix $\hat{\Sigma}_v$ provided in the next Lemma.

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

$$\hat{\Lambda}_v \;=\; \hat{\mathcal{U}}^{-1} \tilde{C}^{-1} \hat{\Psi}_{LL} \hat{\mathcal{U}} \;+ O_p(\delta^2_{N,T}) \tag{B.28}$$

$$\hat{W}^*_{v,r-q} \;=\; (\hat{\mathcal{H}}')^{-1} W_v \left(E_L - E_H \hat{A} \right) \hat{\mathcal{U}} \;+ O_p(\delta^2_{N,T}), \tag{B.29}$$

where $\hat{A} := [E_H' W_v' \hat{\mathcal{H}}' \hat{\mathcal{H}} W_v \tilde{\Sigma}_v E_H]^{-1} E_H' W_v' \hat{\mathcal{H}}' \hat{\mathcal{H}} W_v \hat{\Psi} E_L$*, matrix* $\hat{\Psi}$ *is defined in Lemma B.3, and*

$$\tilde{C} \;:=\; W'_{v,r-q} \hat{\mathcal{H}}^{-1} (\hat{\mathcal{H}}')^{-1} W_{v,r-q}\,. \tag{B.30}$$

Note that the approximation in (B.28) holds for the terms in the main diagonal, as matrix $\hat{\Lambda}_v$ has been defined to be diagonal.

B.2.4 Asymptotic expansion of $\sum_{\ell=1}^{r-q} \hat{\sigma}^2_{r-\ell+1}$

Let us now derive an asymptotic expansion for the sum of the $r-q$ smallest eigenvalues $\sum_{\ell=1}^{r-q} \hat{\sigma}^2_{r-\ell+1}$. Using $\sum_{\ell=1}^{r-q} \hat{\sigma}^2_{r-\ell+1} = tr\left\{\hat{\Lambda}_v\right\}$, we get:

$$\sum_{\ell=1}^{r-q} \hat{\sigma}^2_{r-\ell+1} \;=\; tr\left\{\tilde{C}^{-1} \cdot \hat{\Psi}_{LL}\right\} \;+ O_p(\delta^2_{N,T}), \tag{B.31}$$

by the cyclic property of the trace and including second-order terms in $O_p(\delta^2_{N,T})$. Lemma B.5 provides asymptotic expansions of the terms within curly brackets in (B.31). These are derived starting from the expressions for $\hat{\Psi}_{LL}$ and its components from Lemma B.3, and noting that from Lemma C.1 in the OA we get:

$$\tilde{\Phi} \;=\; \left(\sum_{t=1}^{T} f_t f'_{t-1}\right)\left(\sum_{t=1}^{T} f_{t-1} f'_{t-1}\right)^{-1} \;=\; \tilde{V}_{21} \tilde{V}_{11}^{-1} \;=\; \begin{bmatrix} \tilde{\Phi}_{HH} & \tilde{\Phi}_{HL} \\ \Phi_{LH} & \Phi_{LL} \end{bmatrix}.$$

[27]This can be easily seen by noting that

$$(\hat{\mathcal{H}} W_v \tilde{\Sigma}_v W_v' \hat{\mathcal{H}}') \cdot (\hat{\mathcal{H}}')^{-1} W_v E_L = \hat{\mathcal{H}} W_v \tilde{\Sigma}_v E_L = 0_{(r,r-q)} = (\hat{\mathcal{H}}')^{-1} W_v E_L \cdot 0_{(r-q,r-q)},$$

where the first equality follows from $W_v' W_v = I_r$, and the second one from equation (B.12) and $E_L = [0_{(r-q,q)}, I_{r-q}]'$.

Importantly, the last equation shows that $\tilde{\Phi}_{HL} = \Phi_{HL}$ and $\tilde{\Phi}_{LL} = \Phi_{LL}$, i.e. that the bottom blocks of matrix Φ are estimated without error if the true factors are observed without error, which is critical to derive our asymptotic distribution below.

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

$$\begin{aligned}\sum_{\ell=1}^{r-q} \hat{\sigma}^2_{r-\ell+1} &= \frac{1}{N} tr \left\{ \tilde{C}^{-1} \cdot \frac{1}{T} \sum_{t=1}^{T} E[(u_{Lt} - \Phi_{LH} u_{Ht-1} - \Phi_{LL} u_{Lt-1})(u_{Lt} - \Phi_{LH} u_{Ht-1} - \Phi_{LL} u_{Lt-1})' | \mathcal{F}_t] \right\} \\ &+ \frac{1}{N\sqrt{T}} tr \left\{ \frac{1}{\sqrt{T}} \sum_{t=1}^{T} [(u_{Lt} - \Phi_{LH} u_{Ht-1} - \Phi_{LL} u_{Lt-1})(u_{Lt} - \Phi_{LH} u_{Ht-1} - \Phi_{LL} u_{Lt-1})' \right. \\ &\left. - E[(u_{Lt} - \Phi_{LH} u_{Ht-1} - \Phi_{LL} u_{Lt-1})(u_{Lt} - \Phi_{LH} u_{Ht-1} - \Phi_{LL} u_{Lt-1})' | \mathcal{F}_t]] \right\} \\ &+ O_p(\delta^2_{N,T}) + o_p(\epsilon_{N,T}), \end{aligned} \tag{B.32}$$

where $\epsilon_{N,T} := \frac{1}{N\sqrt{T}}$ *and the terms in curly brackets are* $O_p(1)$.

From the definition of $\epsilon_{N,T} = \frac{1}{N\sqrt{T}}$ in Lemma B.5 and of $\delta_{N,T} := (\min\{N,T\})^{-1}$ in Lemma B.2, and the condition $\sqrt{T} \ll N \ll T^{3/2}$ in Assumption A.1, we have $\delta^2_{N,T} = o(\epsilon_{N,T})$. Therefore the leading stochastic terms in $\sum_{\ell=1}^{r-q} \hat{\sigma}^2_{r-\ell+1}$ are of order $O_p\left(\frac{1}{N}\right)$ and $O_p\left(\frac{1}{N\sqrt{T}}\right)$. Let us define the process

$$U_t := u_{Lt} - \Phi_{LH} u_{Ht-1} - \Phi_{LL} u_{Lt-1}. \tag{B.33}$$

Process U_t depends on N, but we do not make this dependence explicit for expository purpose. Then, from Lemma B.5 we get:

$$\sum_{\ell=1}^{r-q} \hat{\sigma}^2_{r-\ell+1} - \frac{1}{N} tr\left\{\tilde{C}^{-1} \tilde{B}_U\right\} = \frac{1}{N\sqrt{T}} \left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T} [U_t' U_t - E(U_t' U_t | \mathcal{F}_t)] \right) + o_p(\epsilon_{N,T}), \tag{B.34}$$

where:

$$\tilde{B}_U := \frac{1}{T} \sum_{t=1}^{T} E(U_t U_t' | \mathcal{F}_t).$$

Under our assumptions, term $\frac{1}{\sqrt{T}} \sum_{t=1}^{T} [U_t' U_t - E(U_t' U_t | \mathcal{F}_t)]$ is $O_p(1)$ as, in the next subsection, we show that this term is asymptotically Gaussian distributed. Under our condition $\sqrt{T} \ll N \ll T^{3/2}$ the remainder term $o_p(\epsilon_{N,T})$ in the r.h.s. of (B.34) is negligible with respect to the first term in the r.h.s.[28]

B.2.5 Asymptotic distribution of the test statistic under the null hypothesis $H(q)$

From the asymptotic expansion (B.34) we obtain the asymptotic distribution of $\hat{\xi}(q) = \sum_{\ell=1}^{r-q} \hat{\sigma}^2_{r-\ell+1}$ under the null hypothesis $H(q)$ of q common shocks. First, we apply a CLT for weakly dependent triangular array data to

[28] If $N \gtrapprox T^{3/2}$, then, similarly to what happens in the asymptotic expansion AGGR, additional higher order terms appear in the asymptotic distribution of $\sum_{\ell=1}^{r-q} \hat{\sigma}^2_{r-\ell+1}$, not necessarily negligible with respect to the term of order $\frac{1}{N\sqrt{T}}$, driven by non-linear functions of the higher order term b_t of the expansion of $\hat{\hat{f}}$ in Corollary B.1. Under the Assumptions of Theorem 2 these terms are zero.

prove the asymptotic normality of $\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \mathcal{Z}_{N,t}$ as $N, T \to \infty$, where

$$\mathcal{Z}_{N,t} := U_t' U_t - E(U_t' U_t | \mathcal{F}_t)$$

depends on N via process U_t defined in (B.33). Let process $V_{N,t} \equiv V_t$ be as defined in Assumption A.6, and let $\mathcal{V}_{t-m}^{t+m} = \sigma(V_s, t-m \leq s \leq t+m)$ for any positive integer m, with $\mathcal{V}_t \equiv \mathcal{V}_{-\infty}^{t}$.

LEMMA B.6. *Under Assumptions A.3, A.4 a), b), A.5 b) and A.6 a), b) we have:*

(i) $\mathcal{Z}_{N,t}$ *is measurable w.r.t.* $\mathcal{V}_t$*, and* $E[\mathcal{Z}_{N,t}] = 0$ *for all* $t \geq 1$ *and* $N \geq 1$*,*

(ii) $\sup_{t \geq 1, N \geq 1} E\left[|\mathcal{Z}_{N,t}|^r\right] < \infty$*, for a constant* $r > 2$*,*

(iii) Process $(\mathcal{Z}_{N,t})$ *is* L^2 *Near Epoch Dependent (*L^2*-NED) of size* -1 *on process* (V_t)*, and* (V_t) *is strong mixing of size* $-r/(r-2)$*, uniformly in* $N \geq 1$*,* [29]

(iv) The limiting variance $\Omega_U := \lim_{T,N\to\infty} V\left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \mathcal{Z}_{N,t}\right)$ *is strictly positive and such that*

$$\Omega_U = \sum_{h=-\infty}^{\infty} \Gamma(h), \qquad \Gamma(h) := \lim_{N\to\infty} Cov\left(\mathcal{Z}_{N,t}, \mathcal{Z}_{N,t-h}\right). \tag{B.35}$$

Then, by an application of the univariate CLT in Corollary 24.7 in Davidson (1994) and the Cramér-Wold device, we have that:

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \mathcal{Z}_{N,t} \xrightarrow{d} N\left(0, \Omega_U\right), \tag{B.36}$$

as $T, N \to \infty$. Let us now compute the limit autocovariance function $\Gamma(h)$ explicitly. By the Law of Iterated Expectation and $E[\mathcal{Z}_{N,t}|\mathcal{F}_t] = 0$, we have:

$$\Gamma(h) = \lim_{N\to\infty} E\left[Cov\left(\mathcal{Z}_{N,t}, \mathcal{Z}_{N,t-h} | \mathcal{F}_t\right)\right]. \tag{B.37}$$

The next Lemma provides the asymptotic distribution of $(U_t', U_{t-h}')'$.

LEMMA B.7. *From Assumptions A.3 and A.5 a), the vector* $(U_t', U_{t-h}')'$ *is asymptotically Gaussian for any* h*,* t *as* $N \to \infty$*:*

$$\left(\begin{array}{c} U_t \\ U_{t-h} \end{array}\right) \xrightarrow{d} \left(\begin{array}{c} U_t^\infty \\ U_{t-h}^\infty \end{array}\right) \sim N\left(0_{(2r,1)}, \left[\begin{array}{cc} \Sigma_{U,t}(0) & \Sigma_{U,t}(h) \\ \Sigma_{U,t}(h)' & \Sigma_{U,t-h}(0) \end{array}\right]\right), \quad (\mathcal{F}_t\text{-stably}), \tag{B.38}$$

where

$$\begin{aligned} \Sigma_{U,t}(h) &= Cov(U_t^\infty, U_{t-h}^\infty | \mathcal{F}_t) = E[U_t^\infty U_{t-h}^{\infty\prime} | \mathcal{F}_t] \\ &= \Sigma_{u,t,LL}(h|\mathcal{F}_t) - \Phi_{LH}\Sigma_{u,t-1,LH}(h-1|\mathcal{F}_t)' - \Phi_{LL}\Sigma_{u,t-1,LL}(h-1|\mathcal{F}_t)' \\ &\quad -\Sigma_{u,t,LH}(h+1|\mathcal{F}_t)\Phi_{LH}' - \Sigma_{u,t,LL}(h+1|\mathcal{F}_t)\Phi_{LL}' + \Phi_{LH}\Sigma_{u,t-1,HH}(h|\mathcal{F}_t)\Phi_{LH}' \\ &\quad +\Phi_{LL}\Sigma_{u,t-1,LH}(h|\mathcal{F}_t)\Phi_{LH}' + \Phi_{LH}\Sigma_{u,t-1,HL}(h|\mathcal{F}_t)\Phi_{LL}' + \Phi_{LL}\Sigma_{u,t-1,LL}(h|\mathcal{F}_t)\Phi_{LL}', \end{aligned}$$

and $\Sigma_{U,s}(h|\mathcal{F}_t) = Cov(U_s^\infty, U_{s-h}^\infty | \mathcal{F}_t) = E[U_s^\infty U_{s-h}^{\infty\prime} | \mathcal{F}_t]$*, for all t and* h *including* $h = 0$*, and* $s \leq t$*.*

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

Using analogous arguments, we can also compute explicitly the term $\tilde{B}_U$ in the bias of $\tilde{\xi}^u(q)$:

$$\begin{aligned}\tilde{B}_U &= \frac{1}{T}\sum_{t=1}^{T}\Big\{\tilde{\Sigma}_{u,t,LL}(0|\mathcal{F}_t) - \Phi_{LH}\tilde{\Sigma}_{u,t-1,LH}(-1|\mathcal{F}_t)' - \Phi_{LL}\tilde{\Sigma}_{u,t-1,LL}(-1|\mathcal{F}_t)' \\ &\quad -\tilde{\Sigma}_{u,t,LH}(1|\mathcal{F}_t)\Phi_{LH}' - \tilde{\Sigma}_{u,t,LL}(1|\mathcal{F}_t)\Phi_{LL}' + \Phi_{LH}\tilde{\Sigma}_{u,t-1,HH}(0|\mathcal{F}_t)\Phi_{LH}' \\ &\quad +\Phi_{LL}\tilde{\Sigma}_{u,t-1,LH}(0|\mathcal{F}_t)\Phi_{LH}' + \Phi_{LH}\tilde{\Sigma}_{u,t-1,HL}(0|\mathcal{F}_t)\Phi_{LL}' + \Phi_{LL}\tilde{\Sigma}_{u,t-1,LL}(0|\mathcal{F}_t)\Phi_{LL}'\Big\} .\end{aligned}$$

The next lemma provides the expression for $\Gamma(h)$, which is necessary to compute the variance Ω_U.

LEMMA B.8. *Under Assumptions A.3 and A.5 b), we have:*

$$\Gamma(h) = E\left[Cov(U_t^{\infty\prime}U_t^{\infty}, U_{t-h}^{\infty\prime}U_{t-h}^{\infty}|\mathcal{F}_t)\right], \tag{B.39}$$

where

$$Cov(U_t^{\infty\prime}U_t^{\infty}, U_{t-h}^{\infty\prime}U_{t-h}^{\infty}|\mathcal{F}_t) = 2tr\left\{\Sigma_{U,t}(h)\Sigma_{U,t}(h)'\right\}. \tag{B.40}$$

From (B.35) and Lemma B.8 we get:

$$\Omega_U = \sum_{h=-\infty}^{\infty} 2tr\left\{E\left[\Sigma_{U,t}(h)\Sigma_{U,t}(h)'\right]\right\} . \tag{B.41}$$

Finally, from equations (B.34) and (B.41), and $N\sqrt{T}\Omega_U^{-1/2} = O\left(N\sqrt{T}\right) = O(\epsilon_{N,T}^{-1})$, under the hypothesis of q primitive shocks the statistic $\hat{\xi}(q) = \sum_{\ell=1}^{r-q}\hat{\sigma}_{r-\ell+1}^2$ is such that:

$$N\sqrt{T}\Omega_U^{-1/2}\left[\hat{\xi}(q) - \frac{1}{N}tr\left\{\tilde{C}^{-1}\tilde{B}_U\right\}\right] = \Omega_U^{-1/2}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\mathcal{Z}_{N,T}\right) + o_p(1)$$

From equation (B.36), the r.h.s. converges in distribution to a standard normal distribution, which yields Theorem 1.

B.3 Proof of Theorem 2

The derivation of the asymptotic distribution of the feasible statistic $\tilde{\xi}(q)$ in (4.6) in Theorem 2 requires to control the effect of replacing the bias term $tr\{\tilde{C}^{-1}\tilde{B}_U\}$ and variance Ω_U terms in the unfeasible statistics $\tilde{\xi}^u(q)$ by means of their estimates, which are functions of the factors and loadings estimates. In the OA Section C.9 we derive also the asymptotic expansions of factors and loadings estimators which, additionally to confirming that the estimators are consistent, are also critical to prove Theorem 2. In Subsection B.3.1 and B.3.2 we prove the statements in Part i) and in Part ii) of Theorem 2, respectively.

B.3.1 Proof of Part (i)

Under the assumptions of Theorem 2, the asymptotic distribution of the unfeasible statistic $\tilde{\xi}^u(q)$ is:

$$\tilde{\xi}^u(q) = N\sqrt{T}\Omega_U^{-1/2}\left[\hat{\xi}(q) - \frac{1}{N}tr\left\{\tilde{C}^{-1}\tilde{B}_U\right\}\right] \xrightarrow{d} N(0,1), \tag{B.42}$$

where, importantly, the expressions of $\tilde{B}_U$ and Ω_U simplify those in (4.4) and (4.5), respectively. By rewriting the feasible test statistics $\tilde{\xi}(q)$ in (4.7) as

$$\tilde{\xi}(q) \;=\; \left(\frac{\hat{\Omega}_U}{\Omega_U}\right)^{-1/2} \left\{ \tilde{\xi}^u(q) + O_p\left(\sqrt{T}\left[tr\left\{\tilde{C}^{-1}\tilde{B}_U\right\} - tr\left\{\hat{B}_U\right\}\right]\right) \right\}, \tag{B.43}$$

it can be seen that Theorem 2 part i) follows if we prove

$$tr\left\{\hat{B}_U\right\} \;=\; tr\left\{\tilde{C}^{-1}\tilde{B}_U\right\} + o_p\left(\frac{1}{\sqrt{T}}\right), \tag{B.44}$$

$$\hat{\Omega}_U \;=\; \Omega_U + o_p(1), \tag{B.45}$$

as the ratio $\hat{\Omega}_U/\Omega_U$ converges in probability to 1 from (B.45), the first term $\tilde{\xi}^u(q)$ in curly brackets converges in distribution to a standard Gaussian distribution from (B.42), and the second term in curly brackets is $o_p(1)$ from (B.44). We now prove results (B.44) and (B.45) by deriving the asymptotic expansions of $\tilde{C}^{-1}$ and all the terms appearing in the expressions of $\hat{B}_U$ and $\hat{\Omega}_u$ defined in Theorem 2, i.e. the blocks of $\hat{\Phi}$ and those of $\hat{\Sigma}_u = \left(\frac{1}{N}\hat{\Lambda}'\hat{\Lambda}\right)^{-1}\left(\frac{1}{N}\hat{\Lambda}'\hat{\Gamma}\hat{\Lambda}\right)\left(\frac{1}{N}\hat{\Lambda}'\hat{\Lambda}\right)^{-1}$, which involve the estimated factors, loadings and residuals. The asymptotic expansions of the factors and loadings from Lemma C.4 in OA Section C.9 allow to show the next result.

LEMMA B.9. *Under Assumptions A.1 - A.9, i) The asymptotic expansion of* $\hat{\Phi} = \left(\sum_{t=1}^T \hat{f}_t\hat{f}_{t-1}'\right)\left(\sum_{t=1}^T \hat{f}_{t-1}\hat{f}_{t-1}'\right)^{-1}$ *is:*

$$\hat{\Phi} = \tilde{\mathcal{H}}^{-1}\tilde{\Phi}\tilde{\mathcal{H}} + o_p\left(\frac{1}{\sqrt{T}}\right), \tag{B.46}$$

where

$$\tilde{\mathcal{H}} := \begin{bmatrix} \tilde{\mathcal{H}}_H & 0_{(q,r-q)} \\ 0_{(r-q,q)} & \tilde{\mathcal{H}}_L \end{bmatrix}, \tag{B.47}$$

and $\hat{\mathcal{H}}_H$, $\hat{\mathcal{H}}_L$ *are non-singular matrices w.p.a. 1, and such that:*

$$\tilde{\Sigma}_{f,HH}^{-1} \;=\; \left(\tilde{\mathcal{H}}_H^{-1}\right)'\tilde{\mathcal{H}}_H^{-1} + o_p\left(\frac{1}{\sqrt{T}}\right), \qquad \tilde{\Sigma}_{f,LL}^{-1} \;=\; \left(\tilde{\mathcal{H}}_L^{-1}\right)'\tilde{\mathcal{H}}_L^{-1} + o_p\left(\frac{1}{\sqrt{T}}\right),$$

with $\tilde{\Sigma}_{f,HH} := \left[\frac{1}{T+1}\sum_{t=0}^T f_t f_t'\right]_{HH}$ *and* $\tilde{\Sigma}_{f,LL} := \left[\frac{1}{T+1}\sum_{t=0}^T f_t f_t'\right]_{LL}$. *ii) The asymptotic expansion of estimator* $\hat{\Lambda}'\hat{\Lambda}/N$ *is:*

$$\frac{\hat{\Lambda}'\hat{\Lambda}}{N} \;=\; \tilde{\mathcal{H}}'\tilde{\Sigma}_\Lambda\tilde{\mathcal{H}} + o_p\left(\frac{1}{\sqrt{T}}\right), \tag{B.48}$$

where $\tilde{\Sigma}_\Lambda = \frac{1}{N}\Lambda'\Lambda$ *with* $\Lambda = [\Lambda_q \vdots \Lambda_{r-q}]$. *iii) The asymptotic expansion of* $\hat{\Lambda}'\hat{\Gamma}\hat{\Lambda}/N$ *is:*

$$\frac{1}{N}\hat{\Lambda}'\hat{\Gamma}\hat{\Lambda} = \tilde{\mathcal{H}}'\tilde{\Omega}\tilde{\mathcal{H}} + o_p\left(\frac{1}{\sqrt{T}}\right), \tag{B.49}$$

where $\tilde{\Omega} = \frac{1}{N}\Lambda'\Gamma\Lambda$, *with* $\Gamma = diag(\gamma_{ii}, i = 1, ..., N)$.

From equation (B.48) we get $\left(\frac{1}{N}\hat{\Lambda}'\hat{\Lambda}\right)^{-1} = \tilde{\mathcal{H}}^{-1}\tilde{\Sigma}_{\Lambda}^{-1}\left(\tilde{\mathcal{H}}'\right)^{-1} + o_p\left(\frac{1}{\sqrt{T}}\right)$. By substituting the latter equality into the expression of $\hat{\Sigma}_u$ in (4.6), and using (B.49), we get

$$\hat{\Sigma}_u = \tilde{\mathcal{H}}^{-1}\tilde{\Sigma}_u\left(\tilde{\mathcal{H}}'\right)^{-1} + o_p\left(\frac{1}{\sqrt{T}}\right), \tag{B.50}$$

where $\tilde{\Sigma}_u = \tilde{\Sigma}_{\Lambda}^{-1}\tilde{\Omega}\tilde{\Sigma}_{\Lambda}^{-1}$. From equation (B.50) and (B.46) we get the asymptotic expansion for $\hat{B}_U$ in (4.4):

$$\hat{B}_U = \tilde{\mathcal{H}}_L^{-1}\tilde{B}_U\left(\tilde{\mathcal{H}}_L'\right)^{-1} + o_p\left(\frac{1}{\sqrt{T}}\right). \tag{B.51}$$

From Lemma C.4 in the OA we have $\tilde{\Sigma}_{f,LL}^{-1} = \left(\tilde{\mathcal{H}}_L^{-1}\right)'\tilde{\mathcal{H}}_L^{-1} + o_p\left(\frac{1}{\sqrt{T}}\right)$. The following Lemma B.10, establishes the link between $\tilde{C}$ and $\tilde{\Sigma}_{f,LL}$.

LEMMA B.10. *Under Assumptions A.1-A.4, A.5 b), c), A.6 a), and A.7, we have* $\tilde{C} = \tilde{\Sigma}_{f,LL} + o_p\left(\frac{1}{\sqrt{T}}\right)$.

Therefore, we have $\tilde{C}^{-1} = \left(\tilde{\mathcal{H}}_L^{-1}\right)'\tilde{\mathcal{H}}_L^{-1} + o_p\left(\frac{1}{\sqrt{T}}\right)$. This equation, together with the asymptotic expansion (B.51), the cyclical property of the trace operator, and the fact that $\tilde{\Phi}_{LH} = \Phi_{LH}$ and $\tilde{\Phi}_{LL} = \Phi_{LL}$ imply equation (B.44). Similarly, the asymptotic expansions (B.46) and (B.50), and the result $\tilde{\Sigma}_u \xrightarrow{p} \Sigma_u(0)$ imply equation (B.45).

B.3.2 Proof of Part (ii)

In order to prove Theorem 2 *(ii)*, we consider the behavior of statistic $\tilde{\xi}(q)$ under the alternative hypothesis H_1 of more than $q^* > q$ primitive shocks. Specifically, let $q^* > q$ be the true number of primitive shocks in the DGP corresponding to the alternative hypothesis. The statistic is given by: $\tilde{\xi}(q) = N\sqrt{T}\hat{\Omega}_U^{-1/2} \times \left(\sum_{\ell=1}^{r-q}\hat{\sigma}_{r-\ell+1}^2 - \frac{1}{N}tr\left\{\hat{B}_U\right\}\right)$. We rely on the following Lemma to prove that $\hat{B}_U$ is asymptotically bounded in probability.

LEMMA B.11. *Under the alternative hypothesis $H(q^*)$, with $q^* > q$, we have $\|\hat{B}_U\| \leq C$, w.p.a. 1, for a constant $C > 0$.*

From Lemma B.11 and using $\sum_{\ell=1}^{r-q}\hat{\sigma}_{r-\ell+1}^2 = \sum_{\ell=1}^{r-q}\sigma_{r-\ell+1}^2 + o_p(1)$, where the $o_p(1)$ term follows from the continuity of the eigenvalues mapping, we get $\tilde{\xi}(q) = N\sqrt{T}\hat{\Omega}_U^{-1/2}\left(\sum_{\ell=1}^{r-q}\sigma_{r-\ell+1}^2 + o_p(1)\right)$. Under $H(q^*)$, we have only $r - q^* < r - q$ eigenvalues that are equal to 0, while all the other q^* eigenvalues are strictly larger than 0. Therefore, $\sum_{\ell=1}^{r-q}\sigma_{r-\ell+1}^2 = \sum_{\ell=r-q^*+1}^{r-q}\sigma_{r-\ell+1}^2 = \sum_{\ell=q+1}^{q^*}\sigma_{\ell}^2 > 0$. Then, from Lemma B.11 we get $\tilde{\xi}(q) \geq N\sqrt{T}c_1$, w.p.a. 1, for a constant $c_1 > 0$. The conclusion follows.

ONLINE APPEDIX - Not for publication

"New Tests and Estimators for Common Primitive Shocks".

Federico Carlini [1] Mirco Rubin [2] Pierluigi Vallarino [3]

This Online Appendix provides supplementary material for Carlini, Rubin and Vallarino (2025), and is structured as follows:

- **Section C** presents the proofs for Lemmas B.1-B.11 from the main paper. It also contains the proofs of intermediate results required for the those lemmas;
- **Section D** briefly presents the alternative estimators for q discussed in the Monte Carlo analysis of the main paper;
- **Section E** contains detailed derivations of the bootstrap quantities introduced in Section 5 of the main body;
- **Section F** discusses further Monte Carlo results.

C Proofs of Lemmas

C.1 Proof of Lemma B.1

By substituting the expression for f_t provided in (2.10) into the definitions of $\tilde{V}_{22}$ and $\tilde{V}_{12}$ in (B.2) we get:

$$
\begin{aligned}
\tilde{V}_{22} &= \frac{1}{T}\sum_{t=1}^{T} f_t f_t' = \frac{1}{T}\sum_{t=1}^{T}(\Phi f_{t-1} + v_t)(\Phi f_{t-1} + v_t)' \\
&= \Phi \tilde{V}_{11}\Phi' + \Phi\left(\frac{1}{T}\sum_{t=1}^{T} f_{t-1}v_t'\right) + \left(\frac{1}{T}\sum_{t=1}^{T} v_t f_{t-1}'\right)\Phi' + \frac{1}{T}\sum_{t=1}^{T} v_t v_t', \\
\tilde{V}_{12} &= \frac{1}{T}\sum_{t=1}^{T} f_{t-1} f_t' = \frac{1}{T}\sum_{t=1}^{T} f_{t-1}(\Phi f_{t-1} + v_t)' = \tilde{V}_{11}\Phi' + \frac{1}{T}\sum_{t=1}^{T} f_{t-1}v_t'.
\end{aligned}
$$

[1] LUISS Business School, Rome, Italy (fcarlini@luiss.it).
[2] EDHEC Business School, Nice, France (mirco.rubin@edhec.edu), corresponding author.
[3] Econometrics Institute, Erasmus School of Economics, Rotterdam, The Netherlands (vallarino@ese.eur.nl).

By plugging-in the last two expressions in the definition of $\tilde{\Sigma}_v = \tilde{V}_{22} - \tilde{V}_{21}\tilde{V}_{11}^{-1}\tilde{V}_{12}$, and simplifying terms we get:

$$\begin{aligned}
\tilde{\Sigma}_v &= \tilde{V}_{22} - \tilde{V}_{21}\tilde{V}_{11}^{-1}\tilde{V}_{12} \\
&= \Phi\tilde{V}_{11}\Phi' + \Phi\left(\frac{1}{T}\sum_{t=1}^{T} f_{t-1}v_t'\right) + \left(\frac{1}{T}\sum_{t=1}^{T} v_t f_{t-1}'\right)\Phi' + \frac{1}{T}\sum_{t=1}^{T} v_t v_t' \\
&\quad - \left(\Phi\tilde{V}_{11} + \frac{1}{T}\sum_{t=1}^{T} v_t f_{t-1}'\right)\tilde{V}_{11}^{-1}\left(\tilde{V}_{11}\Phi' + \frac{1}{T}\sum_{t=1}^{T} f_{t-1}v_t'\right) \\
&= \frac{1}{T}\sum_{t=1}^{T} v_t v_t' - \left(\frac{1}{T}\sum_{t=1}^{T} v_t f_{t-1}'\right)\tilde{V}_{11}^{-1}\left(\frac{1}{T}\sum_{t=1}^{T} f_{t-1}v_t'\right). \qquad \text{(C.1)}
\end{aligned}$$

By substituting the definition $v_t = [v_{Ht}' \ , \ 0_{1\times q}]'$ from (2.11) into (C.1) it follows that

$$\begin{aligned}
\tilde{\Sigma}_v &= \begin{bmatrix} \frac{1}{T}\sum_{t=1}^{T} v_{H,t}v_{H,t}' & 0_{q\times(r-q)} \\ 0_{(r-q)\times q} & 0_{(r-q)\times(r-q)} \end{bmatrix} - \begin{bmatrix} \frac{1}{T}\sum_{t=1}^{T} v_{H,t}f_{t-1}' \\ 0_{(r-q)\times r} \end{bmatrix} \tilde{V}_{11}^{-1} \begin{bmatrix} \frac{1}{T}\sum_{t=1}^{T} f_{t-1}v_{H,t}' & 0_{r\times(r-q)} \end{bmatrix} \\
&= \begin{bmatrix} \frac{1}{T}\sum_{t=1}^{T} v_{H,t}v_{H,t}' & 0_{q\times(r-q)} \\ 0_{(r-q)\times q} & 0_{(r-q)\times(r-q)} \end{bmatrix} - \begin{bmatrix} \frac{1}{T}\sum_{t=1}^{T} v_{H,t}f_{t-1}'\tilde{V}_{11}^{-1}\left(\frac{1}{T}\sum_{t=1}^{T} f_{t-1}v_{H,t}'\right) & 0_{q\times(r-q)} \\ 0_{(r-q)\times q} & 0_{(r-q)\times(r-q)} \end{bmatrix},
\end{aligned}$$

which concludes the proof. ■

C.2 Proof of Lemma B.2

The proof of Lemma B.2 follows that of Lemma B.1 in AGGR. As some of the terms in this proof will be needed in the proofs of the following Lemmas, here we report the main steps and results, adapting them to the notation and setup of our paper. We provide the bound for $\hat{X}_{12}$ only as the bounds for the other terms $\hat{X}_{11}$, $\hat{X}_{22}$ and $\hat{X}_{21}$ are obtained similarly. By substituting the definition

$\psi_t = \frac{1}{\sqrt{N}} u_t + \frac{1}{T} b_t + \frac{1}{\sqrt{NT}} d_t + \vartheta_t$ from (B.16) into (B.22) we get:

$$
\begin{aligned}
\hat{X}_{12} &= \frac{1}{T}\sum_{t=1}^{T}(f_{t-1}\psi_t' + \psi_{t-1}f_t') + \frac{1}{T}\sum_{t=1}^{T}\psi_{t-1}\psi_t' \\
&= \frac{1}{T\sqrt{N}}\sum_{t=1}^{T}(f_{t-1}u_t' + u_{t-1}f_t') + \frac{1}{NT}\sum_{t=1}^{T}u_{t-1}u_t' \\
&+ \frac{1}{T^2}\sum_{t=1}^{T}(f_{t-1}b_t' + b_{t-1}f_t') + \frac{1}{T^2\sqrt{N}}\sum_{t=1}^{T}(b_{t-1}u_t' + u_{t-1}b_t') + \frac{1}{T^3}\sum_{t=1}^{T}b_{t-1}b_t' \\
&+ \frac{1}{T\sqrt{NT}}\sum_{t=1}^{T}(f_{t-1}d_t' + d_{t-1}f_t') + \frac{1}{TN\sqrt{T}}\sum_{t=1}^{T}(d_{t-1}u_t' + u_{t-1}d_t') + \frac{1}{NT^2}\sum_{t=1}^{T}d_{t-1}d_t' \\
&+ \frac{1}{T^2\sqrt{NT}}\sum_{t=1}^{T}(b_{t-1}d_t' + d_{t-1}b_t') + \frac{1}{T}\sum_{t=1}^{T}(f_{t-1}\vartheta_t' + \vartheta_{t-1}f_t') \\
&+ \frac{1}{T}\sum_{t=1}^{T}\left[\left(\frac{1}{\sqrt{N}}u_{t-1} + \frac{1}{T}b_{t-1} + \frac{1}{\sqrt{NT}}d_{t-1} + \vartheta_{t-1}\right)\vartheta_t' + \vartheta_{t-1}\left(\frac{1}{\sqrt{N}}u_t + \frac{1}{T}b_t + \frac{1}{\sqrt{NT}}d_t + \vartheta_t\right)'\right].
\end{aligned} \tag{C.2}
$$

Under Assumptions A.2-A.4, A.5 *b)-c)*, A.6 *a)* we have

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}f_{t_1}u_{t_2}' = O_p(1), \qquad \frac{1}{T}\sum_{t=1}^{T}u_{t_1}u_{t_2}' = O_p(1) \tag{C.3}$$

$$\frac{1}{T}\sum_{t=1}^{T}f_{t_1}b_{t_2}' = O_p(1) \tag{C.4}$$

$$\frac{1}{T}\sum_{t=1}^{T}b_{t_1}u_{t_2}' = O_p\left(\frac{1}{\sqrt{T}}\right) \tag{C.5}$$

$$\frac{1}{T}\sum_{t=1}^{T}b_{t_1}b_{t_2}' = O_p(1) \tag{C.6}$$

$$\frac{1}{T}\sum_{t=1}^{T}f_{t_1}d_{t_2}' = O_p(1) \tag{C.7}$$

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}u_{t_1}d_{t_2}' = O_p(1) \tag{C.8}$$

$$\frac{1}{T}\sum_{t=1}^{T}b_{t_1}d_{t_2}' = O_p(1) \tag{C.9}$$

$$\frac{1}{T}\sum_{t=1}^{T}d_{t_1}d_{t_2}' = O_p(1) \tag{C.10}$$

for $t_1, t_2 = t, t-1$. As these bounds can be proved using the same arguments as those used in OA Section C.4 in AGGR to prove their Lemma B.1, their proof is omitted. Therefore, the first nine summation terms in the r.h.s of equation (C.2) are of order $O_P\left(\frac{1}{\sqrt{NT}}\right)$, $O_P\left(\frac{1}{N}\right)$, $O_P\left(\frac{1}{T}\right)$, $O_P\left(\frac{1}{T\sqrt{NT}}\right)$, $O_P\left(\frac{1}{T^2}\right)$, $O_P\left(\frac{1}{\sqrt{NT}}\right)$, $O_P\left(\frac{1}{NT}\right)$, $O_P\left(\frac{1}{T\sqrt{NT}}\right)$, $O_P\left(\frac{1}{NT}\right)$, respectively. From Corollary (B.1), the last two terms in the r.h.s. of equation (C.2) are of order $O_P\left(\frac{1}{N}+\frac{1}{T^2}\right)$ and $o_P\left(\frac{1}{N\sqrt{T}}\right)$, respectively. Therefore, we get $\hat{X}_{12} = O_p(\delta_{N,T})$, where $\delta_{N,T} = \max\left\{\frac{1}{N}, \frac{1}{T}\right\} = (\min\{N,T\})^{-1}$, which concludes the proofs. ∎

C.3 Proof of Lemma B.3

As $(I_r - X)^{-1} = I_r + X + O_p(\delta^2_{N,T})$ for any (r,r) matrix $X = O_p(\delta_{N,T})$, matrix $\hat{\Sigma}_v$ in (B.24) can be expressed as:

$$
\begin{aligned}
\hat{\Sigma}_v &= \tilde{V}_{22} + \hat{X}_{22} - (\tilde{V}_{21} + \hat{X}_{21})\tilde{V}_{11}^{-1}\left[I_r - \hat{X}_{11}\tilde{V}_{11}^{-1} + O_p(\delta^2_{N,T})\right](\tilde{V}_{12} + \hat{X}_{12}) \\
&= \tilde{V}_{22} + \hat{X}_{22} - (\tilde{V}_{21} + \hat{X}_{21})\tilde{V}_{11}^{-1}\left[\tilde{V}_{12} + \hat{X}_{12} - \hat{X}_{11}\tilde{V}_{11}^{-1}\tilde{V}_{12} + +O_p(\delta^2_{N,T})\right] \\
&= \tilde{V}_{22} - \tilde{V}_{21}\tilde{V}_{11}^{-1}\tilde{V}_{12} + \hat{X}_{22} - \tilde{V}_{21}\tilde{V}_{11}^{-1}\hat{X}_{12} - \hat{X}_{21}\tilde{V}_{11}^{-1}\tilde{V}_{12} + \tilde{V}_{21}\tilde{V}_{11}^{-1}\hat{X}_{11}\tilde{V}_{11}^{-1}\tilde{V}_{12} + O_p(\delta^2_{N,T}),
\end{aligned}
$$

where we have used the fact that $X_{jk} = O_p(\delta_{N,T})$ for any $(j,k) \in \{1,2\}^2$ implies $X_{jk}X_{lm} = O_p(\delta^2_{N,T})$ for any $(j,k,l,m) \in \{1,2\}^4$ in the second and the third equations. Recalling that $\tilde{\Sigma}_v = \tilde{V}_{22} - \tilde{V}_{21}\tilde{V}_{11}^{-1}\tilde{V}_{12}$ from equation (B.4), $\tilde{\Phi} = \tilde{V}_{21}\tilde{V}_{11}^{-1}$ and the definition $\hat{\Psi} := \hat{X}_{22} - \tilde{\Phi}\hat{X}_{12} - \hat{X}_{21}\tilde{\Phi}' + \tilde{\Phi}\hat{X}_{11}\tilde{\Phi}'$, we get the expansion

$$
\hat{\Sigma}_v = \tilde{\Sigma}_v + \hat{\Psi} + O_p(\delta^2_{N,T}).
$$

Substituting the last equation into (B.23) we get the asymptotic expansion in (B.25), and proves Lemma B.3. ∎

C.4 Proof of Lemma B.4

Substituting expansions (B.25) and (B.27) into the eigenvalue-eigenvector equation (B.26), we get:

$$
\left(\hat{\mathcal{H}}W_v(\tilde{\Sigma}_v + \hat{\Psi})W_v'\hat{\mathcal{H}}' + O_p(\delta^2_{N,T})\right)(\hat{\mathcal{H}}')^{-1}W_v[E_L\,\hat{\mathcal{U}} + E_H\hat{\alpha}] = (\hat{\mathcal{H}}')^{-1}W_v[E_L\,\hat{\mathcal{U}} + E_H\hat{\alpha}](0_{(r-q,r-q)} + \hat{M}).
$$

By using $\tilde{\Sigma}_v E_L = 0_{r,r-q}$ and keeping only the terms up to first order in $\delta_{N,T}$, the last equation simplifies to:

$$
\hat{\mathcal{H}}W_v\tilde{\Sigma}_v E_H\hat{\alpha} + \hat{\mathcal{H}}W_v\hat{\Psi}E_L\,\hat{\mathcal{U}} = (\hat{\mathcal{H}}')^{-1}W_v E_L\hat{\mathcal{U}}\hat{M} + O_p(\delta^2_{N,T})\,. \tag{C.11}
$$

as $\hat{\mathcal{H}}W_v\hat{\Psi}E_H\hat{\alpha} = O_p(\delta^2_{N,T})$ and $\hat{\alpha}\hat{M} = O_p(\delta^2_{N,T})$. Pre-multiplying both sides of equation (C.11) by $E_H'W_v'\hat{\mathcal{H}}'$, and noting that $E_H'E_L =_{q,r-q}$ we get:

$$
[E_H'W_v'\hat{\mathcal{H}}'\hat{\mathcal{H}}W_v\tilde{\Sigma}_v E_H]\hat{\alpha} + [E_H'W_v'\hat{\mathcal{H}}'\hat{\mathcal{H}}W_v]\hat{\Psi}E_L\,\hat{\mathcal{U}} = O_p(\delta^2_{N,T})\,,
$$

which implies:

$$\hat{\alpha} \;=\; -[E_H' W_v' \hat{\mathcal{H}}' \hat{\mathcal{H}} W_v \tilde{\Sigma}_v E_H]^{-1} E_H' W_v' \hat{\mathcal{H}}' \hat{\mathcal{H}} W_v \hat{\Psi} E_L \,\hat{\mathcal{U}} + O_p(\delta_{N,T}^2).$$

Similarly, pre-multiplying equation (C.11) by $E_L' W_v' \hat{\mathcal{H}}^{-1}$, and noting that equation (B.12) implies $\tilde{\Sigma}_{v,LH} = 0$ we get:

$$\hat{\Psi}_{LL} \hat{\mathcal{U}} \;=\; [E_L' W_v' \hat{\mathcal{H}}^{-1} (\hat{\mathcal{H}}')^{-1} W_v E_L] \hat{\mathcal{U}} \hat{M} + O_p(\delta_{N,T}^2). \tag{C.12}$$

where $\hat{\Psi}_{LL} = E_L' \hat{\Psi} E_L$, which implies:

$$\hat{M} \;=\; \hat{\mathcal{U}}^{-1} [E_L' W_v' \hat{\mathcal{H}}^{-1} (\hat{\mathcal{H}}')^{-1} W_v E_L]^{-1} \hat{\Psi}_{LL} \hat{\mathcal{U}} + O_p(\delta_{N,T}^2). \tag{C.13}$$

The proof of Lemma B.4 is concluded by substituting (C.12) and (C.13) into the expressions of $\hat{W}_{v,r-q}$ and $\hat{\Lambda}_v$ in (B.27). ■

C.5 Proof of Lemma B.5

The proof is based on the asymptotic expansion of the term $\hat{\Psi}_{LL}$ in the r.h.s. of equation (B.31). By substituting equations (B.20)-(B.22) into the expression of $\hat{\Psi} = \hat{X}_{22} - \tilde{\Phi}\hat{X}_{12} - \hat{X}_{21}\tilde{\Phi}' + \tilde{\Phi}\hat{X}_{11}\tilde{\Phi}'$ provided in Lemma B.3, we get:

$$\begin{aligned} \hat{\Psi} &= \frac{1}{T}\sum_{t=1}^{T} \Big[f_t \psi_t' + \psi_t f_t' + \psi_t \psi_t' - \tilde{\Phi} f_{t-1} \psi_t' - \tilde{\Phi} \psi_{t-1} f_t' - \tilde{\Phi} \psi_{t-1} \psi_t' \\ &\quad - \psi_t f_{t-1}' \tilde{\Phi}' - f_t \psi_{t-1}' \tilde{\Phi}' - \psi_t \psi_{t-1}' \tilde{\Phi}' + \tilde{\Phi} f_{t-1} \psi_{t-1}' \tilde{\Phi}' + \tilde{\Phi} \psi_{t-1} f_{t-1}' \tilde{\Phi}' + \tilde{\Phi} \psi_{t-1} \psi_{t-1}' \tilde{\Phi}' \Big] \\ &= \frac{1}{T}\sum_{t=1}^{T} \Big[\tilde{v}_t (\psi_t - \tilde{\Phi}\psi_{t-1})' + (\psi_t - \tilde{\Phi}\psi_{t-1}) \tilde{v}_t' + (\psi_t - \tilde{\Phi}\psi_{t-1})(\psi_t - \tilde{\Phi}\psi_{t-1})' \Big] \end{aligned} \tag{C.14}$$

where the last equality follows by collecting terms in the previous equation and recalling that

$$\tilde{v}_t = f_t - \tilde{\Phi} f_{t-1}$$

from Section 3.2. Recalling that matrix $\hat{\Psi}_{LL}$ is the bottom-right $(r-q, r-q)$ block of $\hat{\Psi}$, from equation (C.14) we get

$$\hat{\Psi}_{LL} \;=\; \frac{1}{T}\sum_{t=1}^{T} \Big\{ \Big[\tilde{v}_t (\psi_t - \tilde{\Phi}\psi_{t-1})' \Big]_{LL} + \Big[(\psi_t - \tilde{\Phi}\psi_{t-1}) \tilde{v}_t' \Big]_{LL} + \Big[(\psi_t - \tilde{\Phi}\psi_{t-1})(\psi_t - \tilde{\Phi}\psi_{t-1})' \Big]_{LL} \Big\} \tag{C.15}$$

The first two terms in the curly brackets in (C.15) can be computed by noting that:

$$
\begin{aligned}
\tilde{v}_t &= f_t - \tilde{\Phi} f_{t-1} \;=\; (\Phi f_{t-1} + v_s) - \left[\sum_{s=1}^{T} f_s f_{s-1}'\right] \cdot \left[\sum_{s=1}^{T} f_{s-1} f_{s-1}'\right]^{-1} f_{t-1} \\
&= \Phi f_{t-1} + v_t - \left[\sum_{s=1}^{T} (\Phi f_{s-1} + v_t) f_{s-1}'\right] \cdot \left[\sum_{s=1}^{T} f_{s-1} f_{s-1}'\right]^{-1} f_{t-1} \\
&= \Phi f_{t-1} + v_t - \Phi f_{t-1} - \left[\sum_{s=1}^{T} v_s f_{s-1}'\right] \tilde{V}_{11}^{-1} f_{t-1} \\
&= v_t - \left[\sum_{s=1}^{T} v_s f_{s-1}'\right] \tilde{V}_{11}^{-1} f_{t-1}.
\end{aligned}
$$

Recalling that from equation (2.11) that $v_t = [v_{Ht}' \;,\; 0_{(r-q)\times 1}']'$ for all dates $t = 1, ..., T$, we get:

$$
\tilde{v}_t(\psi_t - \tilde{\Phi}\psi_{t-1})' = \begin{bmatrix} \cdot & \cdot \\ 0_{(r-q),q} & 0_{(r-q),(r-q)} \end{bmatrix},
$$

which implies

$$
\left[\tilde{v}_t(\psi_t - \tilde{\Phi}\psi_{t-1})'\right]_{LL} = 0_{(r-q,r-q)}, \qquad \text{and} \qquad \left[(\psi_t - \tilde{\Phi}\psi_{t-1})\tilde{v}_t'\right]_{LL} = 0_{(r-q,r-q)},
$$

so that

$$
\hat{\Psi}_{LL} \;=\; \left[\frac{1}{T}\sum_{t=1}^{T} (\psi_t - \tilde{\Phi}\psi_{t-1})(\psi_t - \tilde{\Phi}\psi_{t-1})'\right]_{LL}. \tag{C.16}
$$

The r.h.s. of (C.16) is obtained by substituting $\psi_s = \frac{1}{\sqrt{N}} u_s + \frac{1}{T} b_s + \frac{1}{\sqrt{NT}} d_s + \vartheta_s$ for $s = t, t-1$. From Corollary (B.1) the contribution of the remainder terms ϑ_s, with $s = t, t-1$, is of order $o_p(\epsilon_{N,T}) = o_p\left(\frac{1}{N\sqrt{T}}\right)$. Moreover, under Assumptions A.2-A.4, A.5 *b)-c)*, and A.6 *a)* from bounds (C.8) and (C.10) we have $\frac{1}{\sqrt{T}}\sum_{t=1}^{T} u_{t_1} d_{t_2}' = O_p(1)$ and $\frac{1}{T}\sum_{t=1}^{T} d_{t_1} d_{t_2}' = O_p(1)$. Therefore, we have:

$$
\begin{aligned}
&\frac{1}{T}\sum_{t=1}^{T} (\psi_t - \tilde{\Phi}\psi_{t-1})(\psi_t - \tilde{\Phi}\psi_{t-1})' \\
&= \frac{1}{TN}\sum_{t=1}^{T} (u_t - \tilde{\Phi}u_{t-1})(u_t - \tilde{\Phi}u_{t-1})' + \frac{1}{T^2\sqrt{N}}\sum_{t=1}^{T}\left[(b_t - \tilde{\Phi}b_{t-1})(u_t - \tilde{\Phi}u_{t-1})' + (u_t - \tilde{\Phi}u_{t-1})(b_t - \tilde{\Phi}b_{t-1})'\right] \\
&+ \frac{1}{T^3}\sum_{t=1}^{T} (b_t - \tilde{\Phi}b_{t-1})(b_t - \tilde{\Phi}b_{t-1})' + \frac{1}{T^2\sqrt{NT}}\sum_{t=1}^{T}\left[(b_t - \tilde{\Phi}b_{t-1})(d_t - \tilde{\Phi}d_{t-1})' + (d_t - \tilde{\Phi}d_{t-1})(b_t - \tilde{\Phi}b_{t-1})'\right] \\
&+ o_p(\epsilon_{N,T})
\end{aligned} \tag{C.17}
$$

By recentering the first term in the r.h.s., and highlighting the convergence rates, equation (C.17) can be rewritten as:

$$
\begin{aligned}
&\frac{1}{T}\sum_{t=1}^{T}(\psi_t - \tilde{\Phi}\psi_{t-1})(\psi_t - \tilde{\Phi}\psi_{t-1})' \\
&= \frac{1}{N}\left(\frac{1}{T}\sum_{t=1}^{T} E\left[(u_t - \tilde{\Phi}u_{t-1})(u_t - \tilde{\Phi}u_{t-1})'|\mathcal{F}_t\right]\right) \\
&+ \frac{1}{N\sqrt{T}}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}(u_t - \tilde{\Phi}u_{t-1})(u_t - \tilde{\Phi}u_{t-1})' - E\left[(u_t - \tilde{\Phi}u_{t-1})(u_t - \tilde{\Phi}u_{t-1})'|\mathcal{F}_t\right]\right) \\
&+ \frac{1}{T\sqrt{NT}}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[(b_t - \tilde{\Phi}b_{t-1})(u_t - \tilde{\Phi}u_{t-1})' + (u_t - \tilde{\Phi}u_{t-1})(b_t - \tilde{\Phi}b_{t-1})'\right]\right) \\
&+ \frac{1}{T^2}\left(\frac{1}{T}\sum_{t=1}^{T}(b_t - \tilde{\Phi}b_{t-1})(b_t - \tilde{\Phi}b_{t-1})'\right) \\
&+ \frac{1}{T\sqrt{NT}}\left(\frac{1}{T}\sum_{t=1}^{T}\left[(b_t - \tilde{\Phi}b_{t-1})(d_t - \tilde{\Phi}d_{t-1})' + (d_t - \tilde{\Phi}d_{t-1})(b_t - \tilde{\Phi}b_{t-1})'\right]\right) + o_p(\epsilon_{N,T}),
\end{aligned}
\tag{C.18}
$$

where the terms in the round brackets are $O_p(1)$ from (C.5), (C.6), (C.8) and Lemma B.6, which provides a CLT for $\frac{1}{\sqrt{T}}\sum_{t=1}^{T}(u_t - \tilde{\Phi}u_{t-1})(u_t - \tilde{\Phi}u_{t-1})' - E\left[(u_t - \tilde{\Phi}u_{t-1})(u_t - \tilde{\Phi}u_{t-1})'|\mathcal{F}_t\right]$. Moreover, under our assumption A.1 we also have that $\frac{1}{T\sqrt{NT}} = o(\epsilon_{N,T})$ and $\frac{1}{T^2} = o(\epsilon_{N,T})$, which allows to simplify (C.18) as:

$$
\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T}(\psi_t - \tilde{\Phi}\psi_{t-1})(\psi_t - \tilde{\Phi}\psi_{t-1})' \;=\; & \frac{1}{N}\left(\frac{1}{T}\sum_{t=1}^{T} E\left[(u_t - \tilde{\Phi}u_{t-1})(u_t - \tilde{\Phi}u_{t-1})'|\mathcal{F}_t\right]\right) \\
+ \frac{1}{N\sqrt{T}}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}(u_t - \tilde{\Phi}u_{t-1})(u_t - \tilde{\Phi}u_{t-1})' - E\left[(u_t - \tilde{\Phi}u_{t-1})(u_t - \tilde{\Phi}u_{t-1})'|\mathcal{F}_t\right]\right) & + o_p(\epsilon_{N,T}).
\end{aligned}
$$

Substituting the last equation into (C.16) we get:

$$
\begin{aligned}
\hat{\Psi}_{LL} \;=\; & \frac{1}{N}\left[\frac{1}{T}\sum_{t=1}^{T} E\left((u_t - \tilde{\Phi}u_{t-1})(u_t - \tilde{\Phi}u_{t-1})'|\mathcal{F}_t\right)\right]_{LL} \\
& + \frac{1}{N\sqrt{T}}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}(u_t - \tilde{\Phi}u_{t-1})(u_t - \tilde{\Phi}u_{t-1})' - E\left[(u_t - \tilde{\Phi}u_{t-1})(u_t - \tilde{\Phi}u_{t-1})'|\mathcal{F}_t\right]\right)_{LL} + o_p(\epsilon_{N,T}). \\
\;=\; & \frac{1}{N}\left(\frac{1}{T}\sum_{t=1}^{T} E\left[(u_t - \tilde{\Phi}u_{t-1})_L(u_t - \tilde{\Phi}u_{t-1})'_L|\mathcal{F}_t\right]\right) \\
& + \frac{1}{N\sqrt{T}}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}(u_t - \tilde{\Phi}u_{t-1})_L(u_t - \tilde{\Phi}u_{t-1})'_L - E\left[(u_t - \tilde{\Phi}u_{t-1})_L(u_t - \tilde{\Phi}u_{t-1})'_L|\mathcal{F}_t\right]\right) + o_p(\epsilon_{N,T}).
\end{aligned}
\tag{C.19}
$$

By using

$$(u_t - \tilde{\Phi} u_{t-1})_L = u_{Lt} - [\tilde{\Phi} u_{t-1}]_L = u_{Lt} - \tilde{\Phi}_{LH} u_{Ht-1} - \tilde{\Phi}_{LL} u_{Lt-1}$$

equation (C.19) can be rewritten as:

$$\begin{aligned}
\hat{\Psi}_{LL} &= \frac{1}{N}\left(\frac{1}{T}\sum_{t=1}^{T} E\left[(u_{Lt} - \tilde{\Phi}_{LH} u_{Ht-1} - \tilde{\Phi}_{LL} u_{Lt-1})(u_{Lt} - \tilde{\Phi}_{LH} u_{Ht-1} - \tilde{\Phi}_{LL} u_{Lt-1})' | \mathcal{F}_t\right]\right) \\
&+ \frac{1}{N\sqrt{T}}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}(u_{Lt} - \tilde{\Phi}_{LH} u_{Ht-1} - \tilde{\Phi}_{LL} u_{Lt-1})(u_{Lt} - \tilde{\Phi}_{LH} u_{Ht-1} - \tilde{\Phi}_{LL} u_{Lt-1})'\right. \\
&\left. - E\left[(u_{Lt} - \tilde{\Phi}_{LH} u_{Ht-1} - \tilde{\Phi}_{LL} u_{Lt-1})(u_{Lt} - \tilde{\Phi}_{LH} u_{Ht-1} - \tilde{\Phi}_{LL} u_{Lt-1})' | \mathcal{F}_t\right]\right) + o_p(\epsilon_{N,T}).
\end{aligned} \tag{C.20}$$

By substituting (C.20) into (B.31) we get:

$$\begin{aligned}
\sum_{\ell=1}^{r-q} \hat{\sigma}^2_{r-\ell+1} &= tr\left\{\tilde{C}^{-1} \cdot \hat{\Psi}_{LL}\right\} + O_p(\delta^2_{N,T}) \\
&= \frac{1}{N} tr\left\{\tilde{C}^{-1} \frac{1}{T}\sum_{t=1}^{T} E\left[(u_{Lt} - \tilde{\Phi}_{LH} u_{Ht-1} - \tilde{\Phi}_{LL} u_{Lt-1})(u_{Lt} - \tilde{\Phi}_{LH} u_{Ht-1} - \tilde{\Phi}_{LL} u_{Lt-1})' | \mathcal{F}_t\right]\right\} \\
&+ \frac{1}{N\sqrt{T}} tr\left\{\tilde{C}^{-1}\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}(u_{Lt} - \tilde{\Phi}_{LH} u_{Ht-1} - \tilde{\Phi}_{LL} u_{Lt-1})(u_{Lt} - \tilde{\Phi}_{LH} u_{Ht-1} - \tilde{\Phi}_{LL} u_{Lt-1})'\right.\right. \\
&\left.\left. - E\left[(u_{Lt} - \tilde{\Phi}_{LH} u_{Ht-1} - \tilde{\Phi}_{LL} u_{Lt-1})(u_{Lt} - \tilde{\Phi}_{LH} u_{Ht-1} - \tilde{\Phi}_{LL} u_{Lt-1})' | \mathcal{F}_t\right]\right]\right\} \\
&+ O_p(\delta^2_{N,T}) + o_p(\epsilon_{N,T}).
\end{aligned} \tag{C.21}$$

The next lemma allows to replace the estimated quantities $\tilde{\Phi}_{LH}$ and $\tilde{\Phi}_{LL}$ with their true values Φ_{LH} and Φ_{LL}. Notably, the results is not asymptotic and only relies on the fact that the lower $(r-q)$-dimensional subvector v_{Lt} is a zero vector.

LEMMA C.1. *For any sample size N and $T > r^2$ it holds that*

$$\begin{aligned}
\tilde{\Phi}_{LH} &= \Phi_{LH} + 0_{q,(r-q)}, \\
\tilde{\Phi}_{LL} &= \Phi_{LL} + 0_{(r-q),(r-q)}.
\end{aligned}$$

Hence, we get

$$\begin{aligned}
\sum_{\ell=1}^{r-q} \hat{\sigma}^2_{r-\ell+1} &= \frac{1}{N} tr\left\{\tilde{C}^{-1} \frac{1}{T}\sum_{t=1}^{T} E\left[(u_{Lt} - \Phi_{LH} u_{Ht-1} - \Phi_{LL} u_{Lt-1})(u_{Lt} - \Phi_{LH} u_{Ht-1} - \Phi_{LL} u_{Lt-1})' | \mathcal{F}_t\right]\right\} \\
&+ \frac{1}{N\sqrt{T}} tr\left\{\tilde{C}^{-1}\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}(u_{Lt} - \Phi_{LH} u_{Ht-1} - \Phi_{LL} u_{Lt-1})(u_{Lt} - \Phi_{LH} u_{Ht-1} - \Phi_{LL} u_{Lt-1})'\right.\right. \\
&\left.\left. - E\left[(u_{Lt} - \Phi_{LH} u_{Ht-1} - \Phi_{LL} u_{Lt-1})(u_{Lt} - \Phi_{LH} u_{Ht-1} - \Phi_{LL} u_{Lt-1})' | \mathcal{F}_t\right]\right]\right\} \\
&+ O_p(\delta^2_{N,T}) + o_p(\epsilon_{N,T}).
\end{aligned} \tag{C.22}$$

which can be further simplified by using the next Lemma C.2, which provides an asymptotic expansion of matrix C defined in Theorem 1.

LEMMA C.2. *Under Assumptions A.2-A.4, A.5 b)-c), A.6 a) and A.7 we have:*

$$\hat{\mathcal{H}}'\hat{\mathcal{H}} = I_{r-q} + o_p(1), \tag{C.23}$$

which implies $C = I_{r-q} + o_p(1)$, *and*

$$\tilde{C}^{-1} = I_{r-q} + o_p(1)\ . \tag{C.24}$$

Plugging in $\tilde{C}^{-1} = I_{r-q} + o_p(1)$ from equation (C.24) into the second term in (C.22) concludes the proof of Lemma B.5, as $\frac{1}{N\sqrt{T}} \cdot o_p(1) \cdot O_p(1) = o_p\left(\frac{1}{N\sqrt{T}}\right) = o_p(\epsilon_{N,T})$. ■

C.5.1 Proof of Lemma C.1

By definition of the OLS estimator of Φ when factors are observable, and using the VAR(1) expression in rotated factor space (see equations (2.10) and (3.2)), we have that

$$\begin{aligned}
\tilde{\Phi} &= \left(\sum_{t=1}^{T} f_t f_{t-1}'\right)\left(\sum_{t=1}^{T} f_{t-1} f_{t-1}'\right)^{-1} \\
&= \Phi\left(\sum_{t=1}^{T} f_{t-1} f_{t-1}'\right)\left(\sum_{t=1}^{T} f_{t-1} f_{t-1}'\right)^{-1} + \left(\sum_{t=1}^{T} v_t f_{t-1}'\right)\left(\sum_{t=1}^{T} f_{t-1} f_{t-1}'\right)^{-1} \\
&= \Phi + \left(\sum_{t=1}^{T} v_t f_{t-1}'\right) S_f^{-1},
\end{aligned}$$

where $S_f \coloneqq \sum_{t=1}^{T} f_{t-1} f_{t-1}'$. Since $v_t = [v_{H,t}', 0_{(r-q)}']'$, we have that

$$\left(\sum_{t=1}^{T} v_t f_{t-1}'\right) = \begin{bmatrix} \cdot & \cdot \\ 0_{(r-q),q} & 0_{(r-q),(r-q)} \end{bmatrix},$$

where we omitted the upper blocks as they are irrelevant for the Lemma. Combining this results with the expansion for $\tilde{\Phi}$ we get that

$$\begin{bmatrix} \tilde{\Phi}_{HH} & \tilde{\Phi}_{HL} \\ \tilde{\Phi}_{LH} & \tilde{\Phi}_{LL} \end{bmatrix} = \begin{bmatrix} \Phi_{HH} & \Phi_{HL} \\ \Phi_{LH} & \Phi_{LL} \end{bmatrix} + \begin{bmatrix} \cdot & \cdot \\ 0_{(r-q),q} & 0_{(r-q),(r-q)} \end{bmatrix},$$

where we have again omitted irrelevant blocks in the last matrix. This concludes the proof.

C.5.2 Proof of Lemma C.2

The first result (C.23) corresponds to result in equation (C.29) in OA C.3.1 in of AGGR. The proof of that result requires first to show that all their Assumptions A.2-A.4, A.5 b)-c), A.6 a) and A.7 imply their Proposition C.2, their Assumption C.1 and their equation (C.27). Additionally, the proof of that result requires that the estimated factors are orthogonal, that is $\hat{\check{F}}'\hat{\check{F}}/(T+1) = I_r$ in our notation. As our Assumptions A.2-A.4, A.5 b)-c), A.6 a) and A.7 correspond to those with the same numbering in AGGR, and we also assume that $\hat{\check{F}}'\hat{\check{F}}/(T+1) = I_r$ (see Section 3.2), result (C.23) using the same arguments as in OA C.3.1 in of AGGR.

Equation (C.23) and the equality $(\hat{\mathcal{H}}'\hat{\mathcal{H}})^{-1} = \hat{\mathcal{H}}^{-1}(\hat{\mathcal{H}}')^{-1}$ imply

$$\hat{\mathcal{H}}^{-1}(\hat{\mathcal{H}}')^{-1} = I_{r-q} + o_p(1), \tag{C.25}$$

where the equality follows the expansion of the matrix inverse function in a neighborhood of the identity is

$$(I_{r-q} + X)^{-1} = I_{r-q} - X, \qquad \text{where} \qquad X = o_p(1). \tag{C.26}$$

Substituting equation (C.25) in the expression $\tilde{C} = W'_{v,r-q}\hat{\mathcal{H}}^{-1}(\hat{\mathcal{H}}')^{-1}W_{v,r-q}$ from Theorem 1, we get:

$$\tilde{C}^{-1} \;\;=\;\; W'_{v,r-q}W_{v,r-q} + o_p(1) \;=\; I_{r-q} + o_p(1) \tag{C.27}$$

where the second equality follows from the definition of the $(r, r-q)$ eigenvector matrix $W_{v,r-q}$ in Section 2.1, which implies $W'_{v,r-q}W_{v,r-q} = I_{r-q}$. Finally, the result $\tilde{C}^{-1} = I_{r-q} + o_p(1)$ follows substituting (C.27) the expansion of the matrix inverse function in (C.26). ■

C.6 Proof of Lemma B.6

The proof of Lemma B.6 follows the same steps as the Proof of Lemma B.6 of AGGR, with the notable difference that the definition of process U_t in their paper is different from the one in our paper, as our U_t involves not only the contemporaneous values of the terms u_t (which are function of the terms $\xi_t = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\lambda_i\varepsilon_{i,t}$ defined in Assumption A.5), but also their lagged version u_{t-1}, and $\tilde{\Phi} = \left(\sum_{t=1}^{T} f_t f'_{t-1}\right)\left(\sum_{t=1}^{T} f_{t-1}f'_{t-1}\right)^{-1}$. The proof relies on showing that the conditions in parts (i)-(iv) of Lemma B.6 hold under our Assumptions A.3, A.4 *a)*, *b)*, A.5 *b)* and A.6 *a)*, *b)*.

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

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

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

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

where the first equality follows from stationarity of the data. The series converges because the zero-mean process $\mathcal{Z}_{N,t}$ is a L^2-mixingale with size -1, [4] by Theorem 17.5 in Davidson (1994) and Conditions (ii)-(iii), which implies

$$\|Cov\left(\mathcal{Z}_{N,t},\mathcal{Z}_{N,t-h}\right)\| = \left\|E\left[E(\mathcal{Z}_{N,t}|\mathcal{V}_{t-h})\mathcal{Z}'_{N,t-h}\right]\right\| \leq \|E(\mathcal{Z}_{N,t}|\mathcal{V}_{t-h})\|_2\|\mathcal{Z}_{N,t-h}\|_2 = O\left(h^{-\psi}\right),$$

uniformly in $N \geq 1$, for some $\psi > 1$. The latter uniform bound also allows for an application of the Lebesgue Lemma to get:

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

where $\Gamma(h) = \lim_{N\to\infty} Cov\left(\mathcal{Z}_{N,t},\mathcal{Z}_{N,t-h}\right)$, which yields equation (B.35). The computations in Subsection B.2.5, and in particular Lemma B.8, show that the limit in $\Gamma(h)$ is well-defined. This concludes the proof of Lemma C.6. ■

C.6.1 Proof of Lemma C.3

Assumption A.6 *a)* gives the strong mixing condition for process V_t. Since

$$U_t = [u_t - \Phi u_{t-1}]_L = \left[\tilde{\Sigma}_\Lambda^{-1}\xi_t - \Phi\tilde{\Sigma}_\Lambda^{-1}\xi_{t-1}\right]_L$$

where $\tilde{\Sigma}_\Lambda = \Lambda'\Lambda/N$. Process U_t is a linear combination of components of the α-mixing process V_t and of a finite number of its lags. Hence, U_t is an α-mixing sequence (see Theorem 14.1 in Davidson, 1994) and the NED property for process $\mathcal{Z}_{N,t}$ can be proved by showing that process $X_{N,t} = E(U_t'U_t|\mathcal{F}_t)$ is L^2-NED on (V_t). This is the case as

$$\begin{aligned}
\|X_{N,t} - E(X_{N,t}|\mathcal{V}_{t-m}^{t+m})\|_2 &\leq \|X_{N,t} - E(X_{N,t}|F_t,...,F_{t-m})\|_2 \\
&= \|E(U_t'U_t|\mathcal{F}_t) - E(U_t'U_t|F_t,...,F_{t-m})\|_2 \\
&\leq \|E([u_t'u_t]_{LL}|\mathcal{F}_t)\|_2 + \|E([u_{t-1}'\Phi' u_t]_{LL}|\mathcal{F}_t)\|_2 + 2\|E([u_{t-1}\Phi'\Phi u_{t-1})_{LL}]_{LL}|\mathcal{F}_t)\|_2 \\
&= O(m^{-\psi}),
\end{aligned}$$

for $\psi > 1$, where the first inequality follows from the Minkowski's inequality and the Law of Iterated Expectation (LIE), the first equality follows again by applying the LIE, the second inequality follows by applying again first inequality follows from the Minkowski's inequality, and the last equality follows directly form Assumption A.6 *b)*. The conclusion follows. ■

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

C.7 Proof of Lemma B.7

Since $U_t = [u_t - \Phi u_{t-1}]_L$, and $u_t = (\frac{1}{N}\sum_{i=1}^{N} \lambda_i \lambda_i')^{-1}\xi_{i,t}$, Assumptions A.3 and A.5 a) imply that $(U_t', U_{t-h}')' \xrightarrow{d} (U_t^{\infty}{}', U_{t-h}^{\infty}{}')$, where

$$\left(\begin{array}{c} U_t \\ U_{t-h} \end{array}\right) \xrightarrow{d} \left(\begin{array}{c} U_t^{\infty} \\ U_{t-h}^{\infty} \end{array}\right) \sim N\left(\left[\begin{array}{c} E\left(U_t^{\infty}|\mathcal{F}_t\right) \\ E\left(U_{t-h}^{\infty}|\mathcal{F}_t\right) \end{array}\right], \left[\begin{array}{cc} Cov\left(U_t^{\infty}, U_t^{\infty}|\mathcal{F}_t\right) & Cov\left(U_t^{\infty}, U_{t-h}^{\infty}|\mathcal{F}_t\right) \\ Cov\left(U_t^{\infty}, U_{t-h}^{\infty}|\mathcal{F}_t\right)' & Cov\left(U_{t-h}^{\infty}, U_{t-h}^{\infty}|\mathcal{F}_t\right) \end{array}\right]\right),$$

with

$$E\left(U_{t-h}^{\infty}|\mathcal{F}_t\right) = \operatorname*{plim}_{N\to\infty} E\left[U_t|\mathcal{F}_t\right] = \operatorname*{plim}_{N\to\infty} E\left[\left(u_t - \Phi u_{t-1}\right)_L|\mathcal{F}_t\right] = 0$$

as $E\left[u_{t-h}|\mathcal{F}_t\right] = 0$ for any h, and

$$\begin{aligned}
Cov\left(U_t^{\infty}, U_{t-h}^{\infty}|\mathcal{F}_t\right) &= E\left[U_t^{\infty} U_{t-h}^{\infty\prime}|\mathcal{F}_t\right] = \operatorname*{plim}_{N\to\infty} E\left[\left(u_t - \Phi u_{t-1}\right)_L \left(u_{t-h} - \Phi u_{t-h-1}\right)_L'|\mathcal{F}_t\right] \\
&= \operatorname*{plim}_{N\to\infty} E\left[\left(u_t u_{t-h}' - \Phi u_{t-1} u_{t-h}' - u_t u_{t-h-1}'\Phi' + \Phi u_{t-1} u_{t-h-1}'\Phi'\right)_{LL}|\mathcal{F}_t\right] \\
&= \operatorname*{plim}_{N\to\infty} \left(\tilde{\Sigma}_{u,t}(h|\mathcal{F}_t) - \Phi\tilde{\Sigma}_{u,t-1}(h-1|\mathcal{F}_t) - \tilde{\Sigma}_{u,t}(h+1|\mathcal{F}_t)\Phi' + \Phi\tilde{\Sigma}_{u,t-1}(h|\mathcal{F}_t)\Phi'\right)_{LL} \\
&= \left[\Sigma_{u,t}(h|\mathcal{F}_t) - \Phi\Sigma_{u,t-1}(h-1|\mathcal{F}_t) - \Sigma_{u,t}(h+1|\mathcal{F}_t)'\Phi + \Phi\Sigma_{u,t-1}(h|\mathcal{F}_t)\Phi'\right]_{LL} \\
&= \Sigma_{u,t,LL}(h|\mathcal{F}_t) - \Phi_{LH}\Sigma_{u,t-1,HL}(h-1|\mathcal{F}_t) - \Phi_{LL}\Sigma_{u,t-1,LL}(h-1|\mathcal{F}_t) \\
&\quad -\Sigma_{u,t,LH}(h+1|\mathcal{F}_t)'\Phi_{LH}' - \Sigma_{u,t,LL}(h+1|\mathcal{F}_t)'\Phi_{LL}' + \Phi_{LH}\Sigma_{u,t-1,HH}(h|\mathcal{F}_t)\Phi_{LH} \\
&\quad +\Phi_{LL}\Sigma_{u,t-1,LH}(h|\mathcal{F}_t)\Phi_{LH}' + \Phi_{LH}\Sigma_{u,t-1,HL}(h|\mathcal{F}_t)\Phi_{LL}' + \Phi_{LL}\Sigma_{u,t-1,LL}(h|\mathcal{F}_t)\Phi_{LL}'
\end{aligned}$$

for all h, including $h = 0$. ■

C.8 Proof of Lemma B.8

The proof of result (B.39) follows similar arguments as the Proof of Lemma B.7 in AGGR. We report it for completeness adapting it our set-up and modified set of assumptions.

First, we show that we can interchange the limit $N \to \infty$ and the outer expectation in the r.h.s. of equation (B.37), i.e.:

$$\lim_{N\to\infty} E\left[Cov\left(\mathcal{Z}_{N,t}, \mathcal{Z}_{N,t-h}|\mathcal{F}_t\right)\right] = E\left[\lim_{N\to\infty} Cov\left(\mathcal{Z}_{N,t}, \mathcal{Z}_{N,t-h}|\mathcal{F}_t\right)\right]. \quad \text{(C.28)}$$

Indeed, by the Cauchy-Schwarz inequality, we have the bound $Cov\left(\mathcal{Z}_{N,t}, \mathcal{Z}_{N,t-h}|\mathcal{F}_t\right) \leq \chi_t\chi_{t-h}$, P-a.s., uniformly in $N \geq 1$, where $\chi_t := \sup_{N\geq 1} E\left[|\mathcal{Z}_{N,t}|^2|\mathcal{F}_t\right]^{1/2}$. The uniform upper bound $\chi_t\chi_{t-h}$ is integrable, because $E[\chi_t\chi_{t-h}] \leq E[\chi_t^2]^{1/2}E[\chi_{t-h}^2]^{1/2}$ by Cauchy-Schwarz, and $E[\chi_t^2] = E\left[\sup_{N\geq 1} E\left(|\mathcal{Z}_{N,t}|^2|\mathcal{F}_t\right)\right] \leq c^* \cdot E\left[\sup_{N\geq 1} E\left(|U_t|^4|\mathcal{F}_t\right)\right] < \infty$, for a constant c^*, by Assumption A.5 *b)*. Then, (C.28) follows from the Lebesgue Lemma.

We now need to show that:

$$\lim_{N\to\infty} Cov(U_t'U_t, U_{t-h}'U_{t-h}|\mathcal{F}_t) \quad = \quad Cov(U_t^{\infty\,\prime}U_t^{\infty}, U_{t-h}^{\infty\,\prime}U_{t-h}^{\infty}|\mathcal{F}_t), \quad P-a.s. \qquad \text{(C.29)}$$

We have $Cov(U_t'U_t, U_{t-h}'U_{t-h}|\mathcal{F}_t) = E[(U_t'U_t)(U_{t-h}'U_{t-h})|\mathcal{F}_t] - E[U_t'U_t|\mathcal{F}_t]E[U_{t-h}'U_{t-h}|\mathcal{F}_t]$. Let us prove that:

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

By definition of conditional expectation, this is equivalent to:

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

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

$$\lim_{N\to\infty} E\left[(U_t'U_t)(U_{t-h}'U_{t-h})1_A\right] = E\left[(U_t^{\infty\prime}U_t^{\infty})(U_{t-h}^{\infty\prime}U_{t-h}^{\infty})1_A\right]. \qquad \text{(C.30)}$$

Now, by (B.38) and stable convergence, we have $(U_t'U_t)(U_{t-h}'U_{t-h})1_A \overset{d}{\to} (U_t^{\infty\prime}U_t^{\infty})(U_{t-h}^{\infty\prime}U_{t-h}^{\infty})1_A$. Moreover, by Assumption A.5 *b)*, we have uniform integrability: $\sup_{N\geq 1} E[|(U_t'U_t)(U_{t-h}'U_{t-h})1_A|^{\rho}] < \infty$, for some $\rho > 1$. Therefore, by the Corollary of Theorem 25.12 on page 338 in Billingsley (1995), we get (C.30). By similar arguments applied to $E[U_t'U_t|\mathcal{F}_t]$ and $E[U_{t-h}'U_{t-h}|\mathcal{F}_t]$ equation (C.29) follows. Combining (C.28) and (C.29), equation (B.39) follows.

Equation (B.39) in Lemma B.8 allows to deploy the joint asymptotic Gaussian distribution of $(U_t^{\infty\,\prime}, U_{t-h}^{\infty\,\prime})'$ to compute the limit autocovariance $\Gamma(h)$. The expression of $Cov(U_t^{\infty\,\prime}U_t^{\infty}, U_{t-h}^{\infty\,\prime}U_{t-h}^{\infty}|\mathcal{F}_t)$ inside $\Gamma(h)$ in (B.40) is derived by applying Theorem 12 p. 284 in Magnus and Neudecker (2007) and Theorem 11.22 in Schott (2017) which provide the covariance between two quadratic forms of Gaussian vectors. In particular, as vector $U_t^{\infty,\star} := (U_t^{\infty\prime}, U_{t-h}^{\infty\prime})'$ in equation (B.38) is, conditionally on $\mathcal{F}_t$, normally distributed with zero mean and variance-covariance matrix:

$$V(U_t^{\infty,\star}|\mathcal{F}_t) \quad = \quad \begin{bmatrix} \Sigma_{U,t}(0) & \Sigma_{U,t}(h) \\ \Sigma_{U,t}(h) & \Sigma_{U,t-h}(0) \end{bmatrix},$$

and as

$$Cov(U_t^{\infty\,\prime}U_t^{\infty}, U_{t-h}^{\infty\,\prime}U_{t-h}^{\infty}|\mathcal{F}_t) = Cov(U_t^{\infty,\star\,\prime}M_1U_t^{\infty,\star}, U_t^{\infty,\star\,\prime}M_1U_t^{\infty,\star}|\mathcal{F}_t)$$

with

$$M_1 := \begin{bmatrix} I_{r-q} & 0_{(r-q),(r-q)} \\ 0_{(r-q),(r-q)} & 0_{(r-q),(r-q)} \end{bmatrix}, \qquad M_2 := \begin{bmatrix} 0_{(r-q),(r-q)} & 0_{(r-q),(r-q)} \\ 0_{(r-q),(r-q)} & I_{r-q} \end{bmatrix},$$

Theorem 11.22 in Schott (2017) implies:

$$Cov(U_t^{\infty\,\prime}U_t^{\infty}, U_{t-h}^{\infty\,\prime}U_{t-h}^{\infty}|\mathcal{F}_t) \quad = \quad 2\,tr\left\{M_1 \cdot V(U_t^{\infty,\star}|\mathcal{F}_t) \cdot M_2 \cdot V(U_t^{\infty,\star}|\mathcal{F}_t)\right\} = 2\,tr\left\{\Sigma_{U,t}(h)\Sigma_{U,t}(h)'\right\},$$

where the second equality follows by substituting the expressions for matrices M_1, M_2 and $V(U_t^{\infty,\star}|\mathcal{F}_t)$. Therefore, Lemma B.8 follows. ■

C.9 Proof of Lemma B.9

In order to prove Lemma B.9 we first need to define the estimator $\hat{f}_{L,t}$ of the redundant factors $f_{L,t}$, and then to derive the asymptotic expansion of the factors estimators $\hat{f}_{H,t}$ and $\hat{f}_{L,t}$ and the associated loadings, which are provided by Lemma C.4. Then, the three separate Subsections C.9.1, C.9.3 and C.9.3 provide the proofs of parts i), ii) and iii), respectively, of Lemma B.9.

DEFINITION 2. *The estimator of the redundant static factors $f_{L,t}$ is $\hat{f}_{L,t} = \hat{W}'_{v,r-q}\hat{\mathring{f}}_t$, with $\hat{W}_{v,r-q}$ defined in Section 3.2.*

LEMMA C.4. *Under Assumptions A.1-A.4, A.5 b)-c), A.6 a), A.7, A.8, the asymptotic expansions of factors and loadings estimates are*

$$\hat{f}_{L,t} = \tilde{\mathcal{H}}_L^{-1}\left[f_{L,t} + \frac{1}{\sqrt{N}}u_{L,t}\right] + o_p\left(T^{-1/2}\right), \tag{C.31}$$

$$\hat{f}_{H,t} = \tilde{\mathcal{H}}_L^{-1}\left[f_{H,t} + \frac{1}{\sqrt{N}}u_{L,t}\right] + o_p(T^{-1/2}), \tag{C.32}$$

$$\hat{\lambda}_{L,i} = \tilde{\mathcal{H}}_L'\left[\lambda_{L,i} + \frac{1}{\sqrt{T+1}}w_{L,i}\right] + o_p\left(T^{-1/2}\right), \tag{C.33}$$

$$\hat{\lambda}_{H,i} = \tilde{\mathcal{H}}_H'\left[\lambda_{H,i} + \frac{1}{\sqrt{T+1}}w_{H,i}\right] + o_p\left(T^{-1/2}\right), \tag{C.34}$$

where the $o_p(T^{-1/2})$ terms are uniform w.r.t. $0 \leq t \leq T$ and $1 \leq i \leq N$, vector $u_t = [u'_{H,t}\ ,\ u'_{L,t}]'$ is defined in Corollary B.1, $w_{L,i} = \tilde{\Sigma}_{f,LL}^{-1}\frac{1}{\sqrt{T+1}}\sum_{t=0}^{T} f_{L,t}\varepsilon_{i,t}$ and $w_{H,i} = \tilde{\Sigma}_{f,HH}^{-1}\frac{1}{\sqrt{T+1}}\sum_{t=0}^{T} f_{H,t}\varepsilon_{i,t}$, with $\tilde{\Sigma}_{f,HH} := \left[\frac{1}{T+1}\sum_{t=0}^{T} f_t f_t'\right]_{HH}$ and $\tilde{\Sigma}_{f,LL} := \left[\frac{1}{T+1}\sum_{t=0}^{T} f_t f_t'\right]_{LL}$, and matrices $\tilde{\mathcal{H}}_L$ and $\tilde{\mathcal{H}}_H$ are such that

$$\tilde{\Sigma}_{f,HH}^{-1} = \left(\tilde{\mathcal{H}}_H^{-1}\right)'\tilde{\mathcal{H}}_H^{-1} + o_p\left(\frac{1}{\sqrt{T}}\right), \qquad \tilde{\Sigma}_{f,LL}^{-1} = \left(\tilde{\mathcal{H}}_L^{-1}\right)'\tilde{\mathcal{H}}_L^{-1} + o_p\left(\frac{1}{\sqrt{T}}\right) \tag{C.35}$$

$$\tilde{\mathcal{H}}_L'\hat{\mathcal{H}}_L = I_{r-q} + o_p(1), \qquad \tilde{\mathcal{H}}_H'\hat{\mathcal{H}}_H = I_q + o_p(1), \tag{C.36}$$

By recalling that the estimator $\hat{f}_t$ of the r static factors can be written as

$$\hat{f}_t = \begin{bmatrix} \hat{f}_{H,t} \\ \hat{f}_{L,t} \end{bmatrix} = \begin{bmatrix} \hat{W}'_{v,q} \\ \hat{W}'_{v,r-q} \end{bmatrix}\hat{\mathring{f}}_t = \hat{W}_v'\hat{\mathring{f}}_t,$$

from (C.31) and (C.32) we get:

$$\hat{f}_t = \tilde{\mathcal{H}}^{-1}\left[f_t + \frac{1}{\sqrt{N}}u_t\right] + o_p(T^{-1/2}), \tag{C.37}$$

where $\tilde{\mathcal{H}}$ is the block-diagonal matrix defined in equation (B.47), namely:

$$\tilde{\mathcal{H}} := \begin{bmatrix} \tilde{\mathcal{H}}_H & 0_{(q,r-q)} \\ 0_{(r-q,q)} & \tilde{\mathcal{H}}_L \end{bmatrix},$$

which implies:

$$\tilde{\mathcal{H}}^{-1} = \left[\begin{array}{cc} \tilde{\mathcal{H}}_H^{-1} & 0_{(q,r-q)} \\ 0_{(r-q,q)} & \tilde{\mathcal{H}}_L^{-1} \end{array} \right]. \tag{C.38}$$

Finally, to prove this Lemma is useful to explicitly write the factor model in (2.9) for the i-th individual:

$$y_{i,t} = \lambda_i' f_t + \varepsilon_{i,t}, \quad i = 1, ..., N, \quad t = 0, 1, ..., T,$$

and in matrix notation as:

$$Y = F\Lambda' + \varepsilon,$$

where $Y = [y_1, ..., y_i, ...y_N]$, with $y_i = [y_{i,0}, y_{i,1}, ..., y_{i,T}]'$ is the $(T+1, N)$ matrix collecting the observations $y_{i,t}$, $F = [f_0, f_1, ..., f_T]'$ is the $(T+1, r)$ matrix of factor values, $\Lambda = [\lambda_1, ..., \lambda_N]'$ and $\varepsilon = [\varepsilon_1, ..., \varepsilon_i, ..., \varepsilon_N]$ where $\varepsilon_i = [\varepsilon_{i,0}, \varepsilon_{i,1}, ..., \varepsilon_{i,T}]'$ for all $i = 1, ..., N$. Moreover, to prove some of the next results we need the following Lemma (C.5)

LEMMA C.5. *Under Assumptions A.2 and A.3, the factors are such*

$$F'F/(T+1) = I_r + o_p(1) \qquad \text{as } T \to \infty, \tag{C.39}$$

and the loadings are such that

$$\Lambda'\Lambda/N = \Sigma_\lambda + o(1) \qquad \text{as } N \to \infty, \tag{C.40}$$

where matrix Σ_λ is positive definite.

The results in Lemma C.5 correspond to standard assumptions in the factor literature, see e.g. Bai and Ng (2002b), Stock and Watson (2002), Bai (2003).

C.9.1 Proof of Lemma B.9 Part (i)

By substituting (C.37) into the estimator $\hat{\Phi} = \left(\sum_{t=1}^T \hat{f}_t \hat{f}_{t-1}'\right)\left(\sum_{t=1}^T \hat{f}_{t-1}\hat{f}_{t-1}'\right)^{-1}$ in (3.6), we get:

$$\begin{aligned} \hat{\Phi} \;\; = \;\; & \tilde{\mathcal{H}}^{-1}\left(\frac{1}{T}\sum_{t=1}^T \left[f_t + \frac{1}{\sqrt{N}}u_t + o_p\left(T^{-1/2}\right)\right] \cdot \left[f_{t-1} + \frac{1}{\sqrt{N}}u_{t-1} + o_p\left(T^{-1/2}\right)\right]'\right)(\tilde{\mathcal{H}}^{-1})'[(\tilde{\mathcal{H}}^{-1})']^{-1} \\ & \times \left(\frac{1}{T}\sum_{t=1}^T \left[f_{t-1} + \frac{1}{\sqrt{N}}u_{t-1} + o_p\left(T^{-1/2}\right)\right] \cdot \left[f_{t-1} + \frac{1}{\sqrt{N}}u_{t-1} + o_p\left(T^{-1/2}\right)\right]'\right)^{-1}(\tilde{\mathcal{H}}^{-1})^{-1}. \end{aligned}$$

By noting that $\frac{1}{T\sqrt{N}}\sum_{t=1}^T f_t u_{t-1} = O_p\left(\frac{1}{\sqrt{NT}}\right) = o_p\left(T^{-1/2}\right)$, $\frac{1}{T\sqrt{N}}\sum_{t=1}^T f_{t-1}u_t = o_p\left(T^{-1/2}\right)$, $\frac{1}{T\sqrt{N}}\sum_{t=1}^T f_{t-1}u_{t-1} = o_p\left(T^{-1/2}\right)$, and $\frac{1}{TN}\sum_{t=1}^T u_{t-1}u_{t-1} = O_p\left(\frac{1}{N}\right) = o_p\left(T^{-1/2}\right)$ from (C.3), the assumption that $T^{1/2} \ll N$, and by using Assumptions A.2, A.5 and A.6 a) we get:

$$\begin{aligned}
\hat{\Phi} &= \tilde{\mathcal{H}}^{-1}\left(\frac{1}{T}\sum_{t=1}^{T} f_t f_{t-1}'\right)\left(\frac{1}{T}\sum_{t=1}^{T} f_{t-1} f_{t-1}'\right)^{-1}\tilde{\mathcal{H}} + o_p\left(T^{-1/2}\right) \\
&= \tilde{\mathcal{H}}^{-1}\tilde{\Phi}\tilde{\mathcal{H}} + o_p\left(T^{-1/2}\right) \\
&= \begin{bmatrix} \tilde{\mathcal{H}}_H^{-1}\tilde{\Phi}_{HH}\tilde{\mathcal{H}}_H & \tilde{\mathcal{H}}_H^{-1}\tilde{\Phi}_{HL}\tilde{\mathcal{H}}_L \\ \tilde{\mathcal{H}}_L^{-1}\tilde{\Phi}_{LH}\tilde{\mathcal{H}}_H & \tilde{\mathcal{H}}_L^{-1}\tilde{\Phi}_{LL}\tilde{\mathcal{H}}_L \end{bmatrix} + o_p\left(T^{-1/2}\right),
\end{aligned}$$

where the last equation follows from the definition of $\tilde{\mathcal{H}}^{-1}$ in (C.38).

C.9.2 Proof of Lemma B.9 Part ii)

Vectors λ_i, $\hat{\lambda}_i$ and w_i are defined as

$$\lambda_i = \begin{bmatrix} \lambda_{H,i} \\ \lambda_{L,i}' \end{bmatrix}, \qquad \hat{\lambda}_i = \begin{bmatrix} \hat{\lambda}_{H,i} \\ \hat{\lambda}_{L,i}' \end{bmatrix}, \qquad \text{and} \qquad w_i = \begin{bmatrix} w_{H,i} \\ w_{L,i} \end{bmatrix}.$$

Then, the estimator $\hat{\Lambda} = [\hat{\lambda}_1, ..., \hat{\lambda}_N]'$ of loadings matrix Λ (which is $\hat{\Lambda} = \frac{1}{T+1}Y'\hat{F}$, and is provided in (3.5)) has the following asymptotic expansion:

$$\hat{\Lambda} = \left[\Lambda + \frac{1}{\sqrt{T+1}}G\right]\tilde{\mathcal{H}} + o_p\left(T^{-1/2}\right), \tag{C.41}$$

where

$$G = \frac{1}{\sqrt{T+1}}\varepsilon' F,$$

as $\tilde{\Sigma}_{f,LL} := \left[\frac{1}{T+1}\sum_{t=0}^{T} f_t f_t'\right]_{LL} = I_{r-q} + o_p(1)$ from Lemma C.5, and $o_p(T^{-1/2})$ denotes a matrix whose rows are $(r,1)$ vectors uniformly of order $o_p(T^{-1/2})$.

To derive the asymptotic expansion of matrix $\hat{\Lambda}'\hat{\Lambda}/N$, we work with the asymptotic expansion in equation (C.41), which involves the matrix product:

$$\frac{1}{N}\left[\Lambda + \frac{1}{\sqrt{T+1}}G\right]'\left[\Lambda + \frac{1}{\sqrt{T+1}}G\right] = \frac{1}{N}\Lambda'\Lambda + \frac{1}{N\sqrt{T+1}}\left(\Lambda' G + G\Lambda\right) + \frac{1}{N(T+1)}G'G \tag{C.42}$$

The different terms in the r.h.s of the last equation can be bounded as:

$$\frac{1}{\sqrt{N}}\Lambda' G = \frac{1}{\sqrt{N(T+1)}}\Lambda'\varepsilon' F = \frac{1}{\sqrt{N(T+1)}}\sum_{i=1}^{N}\sum_{t=0}^{T}\lambda_i f_t'\varepsilon_{i,t} = O_p(1),$$

and

$$\frac{1}{N}G'G = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{\sqrt{T+1}}\sum_{t=0}^{T} f_t\varepsilon_{i,t}\right)\left(\frac{1}{\sqrt{T+1}}\sum_{t=0}^{T} f_t\varepsilon_{i,t}\right)' = O_p(1),$$

by arguments similar to the proof of Lemma B.2. Thus, by using these bounds and $\Lambda'\Lambda/N = O(1)$, from equation (C.42) we get:

$$\frac{1}{N}\left[\Lambda + \frac{1}{\sqrt{T+1}}G\right]'\left[\Lambda + \frac{1}{\sqrt{T+1}}G\right] = \frac{1}{N}\Lambda'\Lambda + O_p\left(\frac{1}{\sqrt{NT}} + \frac{1}{T}\right),$$

which implies:

$$\frac{\hat{\Lambda}'\hat{\Lambda}}{N} = \tilde{\mathcal{H}}'\left(\frac{\Lambda'\Lambda}{N}\right)\tilde{\mathcal{H}} + o_p\left(\frac{1}{\sqrt{T}}\right) = \tilde{\mathcal{H}}'\tilde{\Sigma}_\Lambda\tilde{\mathcal{H}} + o_p\left(\frac{1}{\sqrt{T}}\right),$$

where the second equality follow from the definition of $\tilde{\Sigma}_\Lambda$.

C.9.3 Proof of Lemma B.9 Part iii)

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

We start by deriving the uniform asymptotic expansion for the residuals. The asymptotic expansions in (C.31)-(C.34) allow to compute the asymptotic expansion of $\hat{\varepsilon}_{i,t}$:

$$\begin{aligned}
\hat{\varepsilon}_{i,t} &= y_{i,t} - \hat{\lambda}_i'\hat{f}_t = \varepsilon_{i,t} - \left[\hat{\lambda}_i'\hat{f}_t - \lambda_i' f_t\right] \\
&= \varepsilon_{i,t} - \left[\left(\lambda_i + \frac{1}{\sqrt{T+1}}w_i + o_p(T^{-1/2})\right)'\left(f_t + \frac{1}{\sqrt{N}}u_t + o_p(T^{-1/2})\right) - \lambda_i' f_t\right] \\
&= \varepsilon_{i,t} - \left(\frac{1}{\sqrt{N}}\lambda_i' u_t + \frac{1}{\sqrt{T+1}}w_i' f_t\right) + o_p\left(T^{-1/2}\right). \qquad \text{(C.43)}
\end{aligned}$$

Here the $o_p(T^{-1/2})$ term is uniform w.r.t. $1 \le i \le N$, $0 \le t \le T$ by the bounds in the next Lemma C.6 and Assumption A.8 *d)*.

LEMMA C.6. *Let $X = O_{p,\ell}(a_{N,T})$ mean $X = O_p[a_{N,T}(\log T)^{\bar{b}}]$ for some $\bar{b} > 0$. Under Assumption A.8 we have the following uniform bounds:*

$$\begin{aligned}
\sup_{0\le t\le T}\|f_t\| &= O_{p,\ell}(1), \\
\sup_{0\le t\le T}\|u_t\| &= O_{p,\ell}(1), \\
\sup_{1\le i\le N}\left\|\frac{1}{T+1}\sum_{t=0}^{T} f_t\varepsilon_{i,t}\right\| &= O_{p,\ell}(T^{-\eta/2}),
\end{aligned}$$

where $\eta \ge 1/2$.

Equation (C.43) allows us to compute:

$$
\begin{aligned}
\hat{\gamma}_{ii} &= \frac{1}{T+1}\sum_{t=0}^{T}\hat{\varepsilon}_{i,t}^{2} = \frac{1}{T+1}\sum_{t=0}^{T}\left[\varepsilon_{i,t}-\frac{1}{\sqrt{T+1}}w_i'f_t-\frac{1}{\sqrt{N}}\lambda_i'u_t\right]^2+o_p\left(T^{-1/2}\right) \\
&= \frac{1}{T+1}\sum_{t=0}^{T}\varepsilon_{i,t}^{2}-\frac{2}{(T+1)\sqrt{(T+1)}}\sum_{t=0}^{T}\varepsilon_{i,t}w_i'f_t-\frac{2}{(T+1)\sqrt{N}}\sum_{t=0}^{T}\varepsilon_{i,t}\lambda_i'u_t \\
&\quad +\frac{1}{(T+1)^2}\sum_{t=0}^{T}\left(w_i'f_t\right)^2+\frac{1}{(T+1)N}\sum_{t=0}^{T}\left(\lambda_i'u_t\right)^2+\frac{2}{(T+1)\sqrt{(T+1)N}}\sum_{t=0}^{T}\left(w_i'f_t\right)\left(\lambda_i'u_t\right)+o_p\left(T^{-1/2}\right).
\end{aligned}
$$

By using $w_i=\frac{1}{\sqrt{T+1}}\sum_{t=0}^{T}\varepsilon_{i,t}f_t=O_p(1)$, $\frac{1}{\sqrt{T+1}}\sum_{t=0}^{T}\varepsilon_{i,t}u_t=O_p(1)$ uniformly in $1\leq i\leq N$, we get:

$$
\hat{\gamma}_{ii} = \frac{1}{T+1}\sum_{t=0}^{T}\varepsilon_{i,t}^{2}+O_p\left(\frac{1}{N}\right)+o_p\left(T^{-1/2}\right),
$$

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

$$
\hat{\gamma}_{ii} = \frac{1}{T+1}\sum_{t=0}^{T}\varepsilon_{i,t}^{2}+o_p\left(T^{-1/2}\right) = \gamma_{ii}+\frac{1}{\sqrt{T+1}}w_i^{\varepsilon}+o_p\left(T^{-1/2}\right),
$$

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

$$
w_i^{\varepsilon} := \frac{1}{\sqrt{T+1}}\sum_{t=0}^{T}(\varepsilon_{i,t}^{2}-\gamma_{ii}).
$$

Therefore, we have:

$$
\hat{\Gamma} = \Gamma+\frac{1}{\sqrt{T+1}}W^{\varepsilon}+o_p\left(T^{-1/2}\right), \tag{C.44}
$$

where $\Gamma=diag(\gamma_{ii},\ i=1,...,N)$ and $W^{\varepsilon}=diag(w_i^{\varepsilon},\ i=1,...,N)$, for $j=1,2$.

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

From (C.41) and (C.44) we have:

$$
\frac{1}{N}\hat{\Lambda}'\hat{\Gamma}\hat{\Lambda}=\tilde{\mathcal{H}}'\hat{\Omega}^{*}\tilde{\mathcal{H}}+o_p\left(T^{-1/2}\right), \tag{C.45}
$$

where we define:

$$
\begin{aligned}
\hat{\Omega}^{*} &:= \frac{1}{N}\left(\Lambda+\frac{1}{\sqrt{T+1}}G\right)'\left(\Gamma+\frac{1}{\sqrt{T+1}}W^{\varepsilon}\right)\left(\Lambda+\frac{1}{\sqrt{T+1}}G\right) \\
&= \tilde{\Omega}+\hat{\Omega}_I^{*}+\hat{\Omega}_{II}^{*}+\hat{\Omega}_{II}^{*\prime}+\hat{\Omega}_{III}^{*}+\hat{\Omega}_{III}^{*\prime}+\hat{\Omega}_{IV}^{*}+\hat{\Omega}_V^{*},
\end{aligned}
$$

and:

$$\begin{aligned}
\tilde{\Omega} &:= \frac{1}{N}\Lambda'\Gamma\Lambda, \\
\hat{\Omega}^*_I &:= \frac{1}{N\sqrt{T+1}}\Lambda' W^{\varepsilon}\Lambda = O_p\left(\frac{1}{\sqrt{NT}}\right), \\
\hat{\Omega}^*_{II} &:= \frac{1}{N\sqrt{T+1}}G'\Gamma\Lambda = O_p\left(\frac{1}{\sqrt{NT}}\right), \\
\hat{\Omega}^*_{III} &:= \frac{1}{N(T+1)}G' W^{\varepsilon}\Lambda = O_p\left(\frac{1}{T}\right), \\
\hat{\Omega}^*_{IV} &:= \frac{1}{N(T+1)}G'\Gamma G = O_p\left(\frac{1}{T}\right), \\
\hat{\Omega}^*_{V} &:= \frac{1}{N(T+1)\sqrt{T+1}}G' W^{\varepsilon} G = O_p\left(\frac{1}{T\sqrt{T}}\right).
\end{aligned}$$

Collecting the previous results, we get:

$$\hat{\Omega}^* = \tilde{\Omega} + O_p\left(\frac{1}{\sqrt{NT}} + \frac{1}{T}\right).$$

By substituting into equation (C.45) we get:

$$\frac{1}{N}\hat{\Lambda}'\hat{\Gamma}\hat{\Lambda} = \tilde{\mathcal{H}}'\,\tilde{\Omega}\,\tilde{\mathcal{H}} + o_p\left(T^{-1/2}\right).$$

which concludes the Proof of Lemma B.9. ■

C.9.4 Proof of Lemma C.4

We start by providing some uniform bounds in Subsection C.9.4 a), that are instrumental for the rest of the proof of Lemma C.4. Then, in Subsections C.9.4 b) and c) we establish the uniform asymptotic expansions of factors and loadings up to order $o_p(\bar{N}^{-1/2})$, where $\bar{N} = \max\{N, T\}$. Finally, in Subsection C.9.4 e) we show how to get the uniform asymptotic expansions up to order $o_p(T^{-1/2})$ under a less restrictive asymptotic scheme.

a) Uniform bounds

Let $X = O_{p,\ell}(a_{N,T})$ mean $X = O_p[a_{N,T}(\log T)^{\bar{b}}]$ for some $\bar{b} > 0$. Under Assumption A.8 we have the following uniform bounds, which complement those in Lemma C.6:

$$\sup_{0\leq t\leq T} \|b_t\| = O_{p,\ell}(1), \tag{C.46}$$

$$\sup_{0\leq t\leq T} \|d_t\| = O_{p,\ell}(1), \tag{C.47}$$

$$\sup_{0\leq t\leq T} \|\hat{f}_t\| = O_{p,\ell}(1), \tag{C.48}$$

$$\sup_{1\leq i\leq N} \|\frac{1}{T+1}\sum_{t=0}^{T} \varepsilon_{i,t}^2\| = O_p(1), \tag{C.49}$$

$$\sup_{1\leq i\leq N} \frac{1}{N(T+1)} \sum_{\ell=1,\ell\neq i}^{N} \sum_{t=0}^{T} \lambda_\ell \varepsilon_{\ell,t}\varepsilon_{i,t} = O_{p,\ell}(\frac{1}{\sqrt{NT^\eta}}) + O(\frac{1}{N}), \tag{C.50}$$

where $\eta \geq 1/2$. The proofs of these uniform bounds is analogous to that in Section D.4.1 in the OA of AGGR, and therefore is omitted.

b) Asymptotic expansion of $\hat{f}_{L,t}$

Let us start by establishing the asymptotic expansion of $\hat{f}_{L,t}$ up to order $o_p(\bar{N}^{-1/2})$. The eigenvectors associated to the smallest $r-q$ eigenvalues of $\hat{\hat{\Sigma}}_v$ are spanned by $\hat{W}^*_{v,r-q}$, which has the following asymptotic expansion $\hat{W}^*_{v,r-q} = (\hat{\mathcal{H}}')^{-1} W_v \left(E_L - E_H \hat{A}\right)\hat{\mathcal{U}} \quad + O_p(\delta^2_{N,T})$ from (B.29). Hence, the normalized eigenvectors associated to the smallest $r-q$ eigenvalues of $\hat{\hat{\Sigma}}_v$ are:

$$\hat{W}_{v,r-q} = \hat{W}^*_{v,r-q}\hat{D}_1,$$

where

$$\hat{D}_1 = diag(\hat{W}^{*\,\prime}_{v,r-q}\hat{W}^*_{v,r-q})^{-1/2}.$$

Then, we get:

$$\begin{aligned}
\hat{f}_{L,t} &= \hat{W}'_{v,r-q}\hat{\hat{f}}_t = \hat{D}_1\hat{\mathcal{U}}'\left[E'_L - \hat{A}'E'_H\right] W'_v \hat{\mathcal{H}}^{-1}\hat{\mathcal{H}} W_v f_t + O_{p,l}\left(\delta^2_{N,T}\right) \\
&= \hat{D}_1\hat{\mathcal{U}}'\Bigg[f_{L,t} + \frac{1}{\sqrt{N}}u_{L,t} + \frac{1}{T}b_{L,t} + \frac{1}{\sqrt{NT}}d_{L,t} + \vartheta_{L,t} \\
&\quad -\hat{A}\left(f_{H,t} + \frac{1}{\sqrt{N}}u_{H,t} + \frac{1}{T}b_{H,t} + \frac{1}{\sqrt{NT}}d_{H,t} + \vartheta^{(s)}_{H,t}\right)\Bigg] + O_{p,l}\left(\delta^2_{N,T}\right),
\end{aligned}$$

uniformly in $1 \leq t \leq T$, where we use the expansion of the factor estimates in Corollary B.1, and (C.48). Under Assumption A.1 we have $(\log T)^{\bar{b}}\delta^2_{N,T} = o(\bar{N}^{-1/2})$, for any $\bar{b} > 0$, $\frac{1}{\sqrt{N}}\delta_{N,T} = o(\bar{N}^{-1/2})$ and $\frac{1}{T}\delta_{N,T} = o(\bar{N}^{-1/2})$. By using uniform bounds in Lemma C.6 and (C.46)-(C.47), and keeping only terms up to $o_p(\bar{N}^{-1/2})$, we get:

$$\hat{f}_{L,t} = \hat{\mathcal{H}}_L^{-1}\left[f_{L,t} + \frac{1}{\sqrt{N}}u_{L,t} - \hat{A}\cdot f_{L,t}\right] + o_p\left(\bar{N}^{-1/2}\right), \tag{C.51}$$

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

$$\hat{\mathcal{H}}_L^{-1} = \hat{D}_1\hat{\mathcal{U}}'.$$

Recalling from definition of $\hat{A}$ in Lemma B.4 that it is a function of $\hat{\Psi} = O_p(\delta_{N,T})$, we have $\hat{A} = O_p(\delta_{N,T})$, and by using Assumption A.1, equation (C.51) further simplifies to:

$$\hat{f}_{L,t} = \hat{\mathcal{H}}_L^{-1}\left[f_{L,t} + \frac{1}{\sqrt{N}}u_{L,t}\right] + o_p\left(\bar{N}^{-1/2}\right), \tag{C.52}$$

uniformly in $1 \leq t \leq T$. By using Assumption A.1 we have

$$\hat{f}_{L,t} = \hat{\mathcal{H}}_L^{-1}\left[f_{L,t} + \frac{1}{\sqrt{N}}u_{L,t}\right] + o_p\left(T^{-1/2}\right),$$

Finally, let us show the asymptotic expansion for $\hat{\mathcal{H}}_L\hat{\mathcal{H}}_L'$. Substituting the expression of $\hat{f}_{L,t}$ from equation (C.52) into the equality $\frac{1}{T+1}\sum_{t=0}^{T}\hat{f}_{L,t}\hat{f}_{L,t}' = I_{r-q}$, we get:

$$\begin{aligned}
I_{r-q} &= \hat{\mathcal{H}}_L^{-1}\frac{1}{T+1}\sum_{t=0}^{T}\left(f_{L,t} + \frac{1}{\sqrt{N}}u_{L,t}\right)\left(f_{L,t} + \frac{1}{\sqrt{N}}u_{L,t}\right)'\left(\hat{\mathcal{H}}_L^{-1}\right)' + o_p\left(\bar{N}^{-1/2}\right) \\
&= \hat{\mathcal{H}}_L^{-1}\left(\frac{1}{T+1}\sum_{t=0}^{T}f_{L,t}f_{L,t}'\right)\left(\hat{\mathcal{H}}_H^{-1}\right)' + o_p\left(\bar{N}^{-1/2}\right), \\
&= \hat{\mathcal{H}}_L^{-1}\tilde{\Sigma}_{f,LL}\left(\hat{\mathcal{H}}_L^{-1}\right)' + o_p\left(\bar{N}^{-1/2}\right),
\end{aligned}$$

using arguments similar to the proof of Lemma B.2 and Assumption A.1. Thus, we get

$$\hat{\mathcal{H}}_L\hat{\mathcal{H}}_L' = \tilde{\Sigma}_{f,LL} + o_p\left(\bar{N}^{-1/2}\right), \tag{C.53}$$

which yields the first equation in (C.35). By using (C.39) it follows:

$$\hat{\mathcal{H}}_L\hat{\mathcal{H}}_L' = I_{r-q} + O_p(T^{-1/2}). \tag{C.54}$$

c) Asymptotic expansion of $\hat{f}_{H,t}$

To find the asymptotic expansion of $\hat{f}_{H,t}$ up to order $o_p(\bar{N}^{-1/2})$, we first need to find the asymptotic expansion of the estimator $\hat{W}_{w,q}$. Let $\hat{W}_{v,q}^*$ be a (r,q) matrix whose columns are eigenvectors of matrix $\hat{\hat{\Sigma}}_v$ associated with the largest eigenvalues $\hat{\sigma}_1^2$, ..., $\hat{\sigma}_q^2$. We have:

$$\hat{\hat{\Sigma}}_v\hat{W}_{v,q}^* = \hat{W}_{v,q}^*\hat{\Lambda}_{v,1:q},$$

where $\hat{\Lambda}_{v,1:q} = diag(\hat{\sigma}_1^2, ..., \hat{\sigma}_q^2)$ is the (q, q) diagonal matrix containing the q largest eigenvalues of $\hat{\hat{\Sigma}}_v$. Using arguments analogous to those in Section B.2.3 we have

$$\hat{W}_{v,q}^{*} = (\hat{\mathcal{H}}')^{-1} W_v \left(E_H + E_L \hat{A}_2 \right) \hat{\mathcal{U}}_2 \ + O_p(\delta_{N,T}^2),$$

the stochastic (q, r) matrix $\hat{\mathcal{U}}_2$ is nonsingular with probability approaching (w.p.a.) 1, and $\hat{A}_2$ is a $(r-q, q)$ stochastic matrix such that $\hat{\alpha} = o_p(\delta_{N,T})$.

Therefore, the normalized eigenvectors associated to the smallest q eigenvalues of $\hat{\hat{\Sigma}}_v$ are:

$$\hat{W}_{v,r-q} \ = \ \hat{W}_{v,r-q}^{*} \hat{D}_2,$$

where

$$\hat{D}_2 = diag(\hat{W}_{v,r-q}^{*\,\prime} \hat{W}_{v,r-q}^{*})^{-1/2}.$$

Then, by using arguments analogous to those used to derive the asymptotic distribution of $\hat{f}_{L,t}$ in the previous Subsection C.9.4 b) we get:

$$\hat{f}_{H,t} \ = \ \hat{\mathcal{H}}_H^{-1} \left[f_{H,t} + \frac{1}{\sqrt{N}} u_{H,t} \right] + o_p\left(\bar{N}^{-1/2} \right), \tag{C.55}$$

uniformly in $0 \leq t \leq T$, where

$$\hat{\mathcal{H}}_H^{-1} = \hat{D}_2 \hat{\mathcal{U}}_2'.$$

By using Assumption A.1 we have

$$\hat{f}_{H,t} \ = \ \hat{\mathcal{H}}_H^{-1} \left[f_{H,t} + \frac{1}{\sqrt{N}} u_{H,t} \right] + o_p\left(T^{-1/2} \right),$$

By substituting the expression of $\hat{f}_{H,t}$ from equation (C.55) into the equality $\frac{1}{T+1} \sum_{t=0}^{T} \hat{f}_{H,t} \hat{f}_{H,t}' = I_{r-q}$, we get:

$$\begin{aligned} I_q \ &= \ \hat{\mathcal{H}}_H^{-1} \frac{1}{T+1} \sum_{t=0}^{T} \left(f_{H,t} + \frac{1}{\sqrt{N}} u_{H,t} \right) \left(f_{H,t} + \frac{1}{\sqrt{N}} u_{H,t} \right)' \left(\hat{\mathcal{H}}_H^{-1} \right)' + o_p\left(\bar{N}^{-1/2} \right) \\ &= \ \hat{\mathcal{H}}_H^{-1} \left(\frac{1}{T+1} \sum_{t=0}^{T} f_{H,t} f_{H,t}' \right) \left(\hat{\mathcal{H}}_H^{-1} \right)' + o_p\left(\bar{N}^{-1/2} \right), \\ &= \ \hat{\mathcal{H}}_H^{-1} \tilde{\Sigma}_{f,HH} \left(\hat{\mathcal{H}}_H^{-1} \right)' + o_p\left(\bar{N}^{-1/2} \right), \end{aligned}$$

using arguments similar to the proof of Lemma B.2 and Assumption A.1. Thus, we get

$$\hat{\mathcal{H}}_H \hat{\mathcal{H}}_H' = \tilde{\Sigma}_{f,HH} + o_p\left(\bar{N}^{-1/2} \right), \tag{C.56}$$

which yields the second equation in (C.35). By using (C.39) it follows:

$$\hat{\mathcal{H}}_H \hat{\mathcal{H}}_H' = I_q + O_p(T^{-1/2}). \tag{C.57}$$

d) Asymptotic expansion of $\hat{\lambda}_i$

We now derive the asymptotic expansion of $\hat{\lambda}_i = \left[\hat{\lambda}_{H,i}', \hat{\lambda}_{L,i}\right]'$, which is the $i-th$ row of the (N,r) matrix $\hat{\Lambda} = \frac{1}{T+1}Y'\hat{F}$ defined in (3.5). Stacking together (C.52) and (C.55) in $\hat{f}_t = [\hat{f}_{H,t}', \hat{f}_{L,t}']'$ we get:

$$\hat{f}_t = \hat{\mathcal{H}}^{-1}\left[f_t + \frac{1}{\sqrt{\bar{N}}}u_t\right] + o_p\left(\bar{N}^{-1/2}\right), \tag{C.58}$$

where $f_t = [f_{H,t}', f_{L,t}']'$, $u_t = [u_{H,t}', u_{L,t}']'$ and $\hat{\mathcal{H}}^{-1}$ is defined in equation (C.38). As $\hat{F} = [\hat{f}_1, ..., \hat{f}_T]'$, equation (C.58) implies

$$\hat{F} = \left(F + \frac{1}{\sqrt{\bar{N}}}U\right)\left(\hat{\mathcal{H}}^{-1}\right)' + o_p\left(\bar{N}^{-1/2}\right),$$

for $U = [u_1, ..., u_T]'$, so that

$$\hat{F}\hat{\mathcal{H}}' - F = \frac{1}{\sqrt{\bar{N}}}U + o_p\left(\bar{N}^{-1/2}\right),$$

where $o_p\left(\bar{N}^{-1/2}\right)$ denotes a matrix whose rows are uniformly of stochastic order $o_p\left(\bar{N}^{-1/2}\right)$. Then, by using $y_i = [y_{i,1}, ..., y_{i,T}]'$, we have:

$$\begin{aligned} \hat{\lambda}_i &= \frac{1}{T+1}\hat{F}'y_i = \frac{1}{T+1}\hat{F}'\left(F\lambda_i + \varepsilon_i\right) \\ &= \frac{1}{T+1}\hat{F}'\left(\left[\hat{F}\hat{\mathcal{H}}' - \left(\hat{F}\hat{\mathcal{H}}' - F\right)\right]\lambda_i + \varepsilon_i\right) \\ &= \hat{\mathcal{H}}'\lambda_i - \frac{1}{T+1}\hat{F}'\left(\hat{F}\hat{\mathcal{H}}' - F\right)\lambda_i + \frac{1}{T+1}\hat{F}'\varepsilon_i. \end{aligned}$$

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

$$\begin{aligned} \hat{\lambda}_i = \hat{\mathcal{H}}'\Big\{ & \lambda_i + (\hat{\mathcal{H}}')^{-1}(\hat{\mathcal{H}})^{-1}\frac{1}{T+1}F'\varepsilon_i + (\hat{\mathcal{H}}')^{-1}(\hat{\mathcal{H}})^{-1}\frac{1}{T+1}(\hat{F}\hat{\mathcal{H}}' - F)'\varepsilon_i \\ & -(\hat{\mathcal{H}}')^{-1}(\hat{\mathcal{H}})^{-1}\frac{1}{T+1}\left[F + (\hat{F}\hat{\mathcal{H}}' - F)\right]'\left(\hat{F}\hat{\mathcal{H}}' - F\right)\lambda_i\Big\}. \end{aligned} \tag{C.59}$$

Define the (r,r) block diagonal matrix

$$\tilde{\Sigma}_{f,BD} = \begin{bmatrix} \tilde{\Sigma}_{f,HH} & 0_{(q,r-q)} \\ 0_{(r-q,q)} & \tilde{\Sigma}_{f,LL} \end{bmatrix},$$

then, from (C.54) and (C.57) we have

$$
\begin{aligned}
(\hat{\mathcal{H}}')^{-1}(\hat{\mathcal{H}})^{-1} &= \begin{bmatrix} (\tilde{\mathcal{H}}'_H)^{-1} & 0_{(q,r-q)} \\ 0_{(r-q,q)} & (\tilde{\mathcal{H}}'_L)^{-1} \end{bmatrix} \begin{bmatrix} \tilde{\mathcal{H}}_H^{-1} & 0_{(q,r-q)} \\ 0_{(r-q,q)} & \tilde{\mathcal{H}}_L^{-1} \end{bmatrix} \\
&= \begin{bmatrix} (\tilde{\mathcal{H}}_H\tilde{\mathcal{H}}'_H)^{-1} & 0_{(q,r-q)} \\ 0_{(r-q,q)} & (\tilde{\mathcal{H}}_L\tilde{\mathcal{H}}'_L)^{-1} \end{bmatrix} \\
&= \tilde{\Sigma}_{f,BD}^{-1} + o_p(\bar{N}^{-1/2}) \\
&= I_r + O_p(T^{-1/2})
\end{aligned}
$$

where the third equality follows from (C.53) and (C.56), while the last equality is due to Assumption A.1. Then, as Assumptions A.1 and A.8 allow to bound the different terms in (C.59), we get:

$$
\hat{\lambda}_i = \hat{\mathcal{H}}' \left[\lambda_i + \tilde{\Sigma}_{f,BD}^{-1} \frac{1}{T+1} F'\varepsilon_i \right] + o_p(\bar{N}^{-1/2}),
$$

uniformly in $1 \leq i \leq N$. The last equation can be rewritten as

$$
\hat{\lambda}_i = \hat{\mathcal{H}}' \left[\lambda_i + \frac{1}{\sqrt{T+1}} w_i \right] + o_p\left(\bar{N}^{-1/2}\right), \tag{C.60}
$$

where:

$$
w_i := \tilde{\Sigma}_{f,BD}^{-1} \frac{1}{\sqrt{T+1}} F'\varepsilon_i = \tilde{\Sigma}_{f,BD}^{-1} \frac{1}{\sqrt{T+1}} \sum_{t=0}^{T} f_t \varepsilon_{i,t}.
$$

Equation (C.60), the fact that $\tilde{\Sigma}_{f,BD}^{-1}$ is block-diagonal (as $\tilde{\Sigma}_{f,BD}^{-1}$ is block diagonal), the definition of $\tilde{\mathcal{H}}$, and the fact that $f_t = [f'_{H,t}, f'_{L,t}]'$ imply:

$$
\begin{aligned}
\hat{\lambda}_{H,i} &= \hat{\mathcal{H}}'_H \left[\lambda_{H,i} + \frac{1}{\sqrt{T+1}} w_{H,i} \right] + o_p\left(\bar{N}^{-1/2}\right), \\
\hat{\lambda}_{L,i} &= \hat{\mathcal{H}}'_L \left[\lambda_{L,i} + \frac{1}{\sqrt{T+1}} w_{L,i} \right] + o_p\left(\bar{N}^{-1/2}\right),
\end{aligned}
$$

where $w_{L,i} = \tilde{\Sigma}_{f,LL}^{-1} \frac{1}{\sqrt{T+1}} \sum_{t=0}^{T} f_{L,t}\varepsilon_{i,t}$ and $w_{H,i} = \tilde{\Sigma}_{f,HH}^{-1} \frac{1}{\sqrt{T+1}} \sum_{t=0}^{T} f_{H,t}\varepsilon_{i,t}$. Assumption A.1 implies:

$$
\begin{aligned}
\hat{\lambda}_{H,i} &= \hat{\mathcal{H}}'_H \left[\lambda_{H,i} + \frac{1}{\sqrt{T+1}} w_{H,i} \right] + o_p\left(T^{-1/2}\right), \\
\hat{\lambda}_{L,i} &= \hat{\mathcal{H}}'_L \left[\lambda_{L,i} + \frac{1}{\sqrt{T+1}} w_{L,i} \right] + o_p\left(T^{-1/2}\right),
\end{aligned}
$$

which concludes the proof of Lemma C.4.

C.9.5 Proof of Lemma C.5

The results in Lemma C.5 are implied by implied by Assumptions A.2 and A.3. These results are is standard in the factor literature, see e.g. Bai and Ng (2002b), Stock and Watson (2002), Bai (2003).

C.9.6 Proof of Lemma C.6

The uniform bounds in Lemma C.6 can be proved by using analogous arguments to those used to prove the analogous Lemma B.10 of AGGR in Section C.11.13 of their OA.

C.10 Proof of Lemma B.10

Assumptions A.1-A.4, A.5 b), c), A.6 a), and A.7 allow to prove the asymptotic expansion (B.16) in Corollary B.1. By substituting the expansion of $\hat{\hat{f}}_t$ from (B.16) into the equality $\frac{1}{T+1}\sum_{t=0}^{T}\hat{\hat{f}}_t\hat{\hat{f}}_t' = I_r$, using arguments similar to the proof of Lemma B.2 and Assumption A.1, we get:

$$
\begin{aligned}
I_r &= \hat{\mathcal{H}}W_v\left[\frac{1}{T}\sum_{t=0}^{T}\left(f_t+\frac{1}{\sqrt{N}}u_t\right)\left(f_t+\frac{1}{\sqrt{N}}u_t\right)'\right]W_v'\hat{\mathcal{H}}' + o_p\left(\bar{N}^{-1/2}\right) \\
&= \hat{\mathcal{H}}W_v\left[\frac{1}{T+1}\sum_{t=0}^{T}f_tf_t'\right]W_v'\hat{\mathcal{H}}' + o_p\left(T^{-1/2}\right), \\
&= \hat{\mathcal{H}}W_v\tilde{\Sigma}_fW_v'\hat{\mathcal{H}}' + o_p\left(T^{-1/2}\right),
\end{aligned}
$$

where $\tilde{\Sigma}_f = \frac{1}{T+1}\sum_{t=0}^{T}f_tf_t'$. Pre-multiplying both sides by $(\hat{\mathcal{H}}W_v)^{-1}$ and post-multiplying them by $(W_v'\hat{\mathcal{H}}')^{-1}$ yields

$$
(\hat{\mathcal{H}}W_v)^{-1}(W_v'\hat{\mathcal{H}}')^{-1} = W_v'\hat{\mathcal{H}}^{-1}(\hat{\mathcal{H}}')^{-1}W_v = \tilde{\Sigma}_f + o_p\left(T^{-1/2}\right).
$$

Noting that $\tilde{C} = W_{v,r-q}'\hat{\mathcal{H}}^{-1}(\hat{\mathcal{H}}')^{-1}W_{v,r-q} = \left[W_v'\hat{\mathcal{H}}^{-1}(\hat{\mathcal{H}}')^{-1}W_v\right]_{LL}$, the last equation implies

$$
\tilde{C} = \tilde{\Sigma}_{f,LL} + o_p\left(T^{-1/2}\right).
$$

This concludes the proof of Lemma B.10. ■

C.11 Proof of Lemma B.11

Let q^* be the true number of common factors (corresponding to the alternative hypothesis), and let q denote the number of primitive shocks used in the estimation procedure (corresponding to the null hypothesis of interest $H_0(q)$). We consider the case with $q < q^* \leq r$, which implies $r-q > r-q^*$. The proofs consists in bounding the different terms of

$$
\hat{B}_U = \hat{\Sigma}_{u,LL} + \hat{\Phi}_{LH}\hat{\Sigma}_{u,HH}\hat{\Phi}_{LH}' + \hat{\Phi}_{LL}\hat{\Sigma}_{u,LH}\hat{\Phi}_{LH}' + \hat{\Phi}_{LH}\hat{\Sigma}_{u,HL}\hat{\Phi}_{LL}' + \hat{\Phi}_{LL}\hat{\Sigma}_{u,LL}\hat{\Phi}_{LL}',
$$

when a model with q primitive shocks is estimated from a DGP with q^* primitive shocks.

The factor estimator is $\hat{f}_t = \hat{W}_v' \hat{\tilde{f}}_t$, with $\hat{W}_v = [\hat{W}_{v,q}, \hat{W}_{v,r-q}]$ where $\hat{W}_{v,q}$ and is the $r \times q$ matrix whose columns are the eigenvectors of $\hat{\tilde{\Sigma}}_v$ associated with its q largest eigenvalues, and $\hat{W}_{r-q}$ and is the $r \times (r-q)$ matrix whose columns are thr eigenvectors of $\hat{\tilde{\Sigma}}_v$ associated with its $r-q$ smallest eigenvalues, normalized such that, $\hat{W}_{v,q}' \hat{W}_{v,q} = I_q$, $\hat{W}_{v,r-q}' \hat{W}_{v,r-q} = I_{r-q}$, and $\hat{W}_{v,q}' \hat{W}_{v,r-q} = 0_{(q,r-q)}$. As, in this proof, we only want to show that elements of $\hat{B}_U$ are stochastically bounded, we can, without loss of generality, simplify the exposition, and let $\hat{\mathcal{H}} = I_r$, and

$$W_v = [W_{v,q^*}, W_{v,r-q^*}] = I_r, \quad \text{with,} \qquad W_{v,q^*} = \begin{bmatrix} I_{q*} \\ 0_{(r-q^*,q^*)} \end{bmatrix}, \text{ and } \quad W_{v,r-q^*} = \begin{bmatrix} 0_{(q*,r-q*)} \\ I_{(r-q^*,r-q^*)} \end{bmatrix}.$$

Then, we have $\hat{\tilde{\Sigma}}_v = \tilde{\Sigma}_v + o_p(1)$, where

$$\tilde{\Sigma}_v = \begin{bmatrix} \tilde{\Sigma}_{v,HH} & 0_{(q,r-q^*)} \\ 0_{(r-q^*,q^*)} & 0_{(r-q^*,r-q^*)} \end{bmatrix},$$

where the $q^* \times q^*$ matrix $\tilde{\Sigma}_{v,HH}$ is:

$$\tilde{\Sigma}_{v,HH} = \frac{1}{T}\sum_{t=1}^{T} v_{H,t}v_{H,t}' - \frac{1}{T}\sum_{t=1}^{T} v_{H,t}f_{t-1}'\tilde{V}_{11}^{-1}\left(\frac{1}{T}\sum_{t=1}^{T} f_{t-1}v_{H,t}'\right).$$

The large-sample limit of $\hat{W}_{v,q}$ (resp. $\hat{W}_{v,r-q}$) is the matrix of normalized eigenvectors associated to the q (resp. $r-q$) largest (resp. smallest) eigenvalues of matrix $\tilde{\Sigma}_v$. The smallest $r-q^*$ eigenvalues of $\tilde{\Sigma}_v$ are $\sigma_{r-q}^2 = \sigma_{r-q+1}^2 = ...\ \sigma_r^2 = 0$, while the q^* largest ones of matrix $\tilde{\Sigma}_{v,HH}$ are assumed to be strictly positive and distinct, i.e. $\sigma_1^2 > \sigma_2^2 >, ..., > \sigma_{r-q^*}^2 > 0$, to simplify the proof.

Let α denote the $q^* \times (q^* - q)$ matrix whose columns are the normalized eigenvectors corresponding to the smallest $q^* - q$ eigenvalues of $\tilde{\Sigma}_{v,HH}$. Then, we have that $\hat{W}_{v,r-q} = W_1 + o_p(1)$ where

$$W_1 = \begin{bmatrix} \alpha & 0_{(q^*,r-q^*)} \\ 0_{(r-q^*,q^*-q)} & \mathcal{U} \end{bmatrix},$$

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

$$\alpha\alpha' + \beta\beta' = I_{q^*}. \tag{C.61}$$

Then, we also have: $\hat{W}_{v,q} = W_2 + o_p(1)$ where

$$W_2 = \begin{bmatrix} \beta \\ 0_{r-q^*,q} \end{bmatrix}.$$

From Corollary B.1 with $\hat{\mathcal{H}} = I_r$ and $W_v = I_r$, we have $\hat{f}_t \simeq f_t$, where symbol $\simeq$ means equality up to terms that are asymptotically negligible for determining large-sample limits. Then:

$$\hat{f}_{H,t} \simeq W_2' f_t = \beta' f_{H,t}, \qquad \text{and} \qquad \hat{f}_{L,t} \simeq W_1' f_t = \begin{bmatrix} \alpha' f_{H,t} \\ \mathcal{U}' f_{L,t} \end{bmatrix}.$$

Define the $r \times r$ matrix $\mathcal{R} := [W_2, W_1]$. From its definition it follows that $\mathcal{R}\mathcal{R}' = \mathcal{R}'\mathcal{R} = I_r$, and:

$$\hat{f}_t \simeq \mathcal{R}' f_t = \begin{bmatrix} \beta' f_{H,t} \\ \alpha' f_{H,t} \\ \mathcal{U}' f_{L,t} \end{bmatrix},$$

which implies that $\hat{f}_t$ is, asymptotically, an orthogonal transformation of f_t, and that

$$\hat{\Phi} = \left(\sum_{t=1}^{T} \hat{f}_t \hat{f}_{t-1}' \right) \left(\sum_{t=1}^{T} \hat{f}_{t-1} \hat{f}_{t-1}' \right)^{-1} \simeq \mathcal{R}\Phi\mathcal{R}'.$$

Importantly, the last equation implies that

$$\hat{\Phi}_{HH} = O_p(1), \qquad \hat{\Phi}_{HL} = O_p(1), \qquad \hat{\Phi}_{LL} = O_p(1). \tag{C.62}$$

Let us consider the estimation of the factor loadings. From (C.61) and the definition of $\mathcal{U}$, $y_{i,t}$ is given by:

$$\begin{aligned} y_{i,t} &= f_{H,t}' \lambda_{H,i} + f_{L,t}' \lambda_{L,i} + \varepsilon_{i,t} \\ &= [\beta' f_{H,t}]' [\beta' \lambda_{H,i}] + [\alpha' f_{H,t}]' [\alpha' \lambda_{H,i}] + [\mathcal{U}' f_{L,t}]' [\mathcal{U}' \lambda_{L,i}] + \varepsilon_{i,t} \\ &= \underline{f}_{H,t}' \underline{\lambda}_{H,i} + \underline{f}_{L,t}' \underline{\lambda}_{L,i} + \varepsilon_{i,t}, \end{aligned}$$

where $\underline{f}_{H,t} = \beta' f_{H,t}$ and $\underline{\lambda}_{H,i} = \beta' \lambda_{H,i}$, $\underline{f}_{L,t} = \begin{bmatrix} \alpha' f_{H,t} \\ \mathcal{U}' f_{L,t} \end{bmatrix}$, $\underline{\lambda}_{L,i} = \begin{bmatrix} \alpha' \lambda_{H,i} \\ \mathcal{U}' \lambda_{L,i} \end{bmatrix}$. The estimated factor loadings $\hat{\lambda}_{H,i}$ and $\hat{\lambda}_{L,i}$ are obtained by regressing $y_{i,t}$ on $\underline{f}_{H,t}$ and $\underline{f}_{L,t}$, which implies:

$$\hat{\lambda}_{H,i} \simeq \underline{\lambda}_{H,i} = \beta' \lambda_{H,i}, \qquad \hat{\lambda}_{L,i} \simeq \underline{\lambda}_{L,i} = \begin{bmatrix} \alpha' \lambda_{H,i} \\ \mathcal{U}' \lambda_{L,i} \end{bmatrix}$$

Since $\hat{\lambda}_i \simeq \mathcal{R}' \lambda_i$, $\hat{\lambda}_i$ is, asymptotically, an orthogonal transformation of λ_i. Using $\hat{\varepsilon}_{i,t} \simeq \varepsilon_{i,t}$, we get $\hat{\Sigma}_u \simeq \mathcal{R} \Sigma_u \mathcal{R}'$, which implies

$$\hat{\Sigma}_{u,HH} = O_p(1), \qquad \hat{\Sigma}_{u,HL} = O_p(1), \qquad \hat{\Sigma}_{u,LL} = O_p(1). \tag{C.63}$$

Results (C.62) and (C.63) imply that $\hat{B}_U = O_p(1)$, which concludes the proof of Lemma B.11. ■

D Estimation of the test statistics under the wild Bootstrap scheme

This Section provides details for the construction of the test statistics in equation (5.7) starting from the $(T+1, N)$ bootstrapped panel of observables $Y^{(b)} = \left[y_0^{(b)}, y_1^{(b)}, \ldots, y_T^{(b)} \right]'$. This panel is generated from the bootstrap DGP in equations (5.3)-(5.4), that we report here for the sake of illustration:

$$\begin{aligned} y_t^{(b)} &= \hat{\Lambda} f_t^{(b)} + \varepsilon_t^{(b)}, \\ f_t^{(b)} &= \hat{\Phi} f_{t-1}^{(b)} + \hat{v}_t^{H_0(q)}, \end{aligned}$$

for $\hat{\Lambda}$ and $\hat{\Phi}$ defined in Section 3.2, while $\varepsilon_t^{(b)}$ and $\hat{v}_t^{H_0(q)}$ are defined in Section 5.1 (equations (5.1) and (5.2), respectively).

D.1 Test based on the smallest eigenvalues

Let $\hat{\breve{F}}^{(b)} = [\hat{\breve{f}}_0^{(b)}, \hat{\breve{f}}_1^{(b)}, ..., \hat{\breve{f}}_T^{(b)}]'$ be the $(T+1, r)$ matrix of estimated Principal Components (PCs) extracted from panel $Y^{(b)}$ associated with the largest r eigenvalues of matrix $\frac{1}{N(T+1)} Y^{(b)} Y^{(b)\prime}$. That is, $\hat{\breve{F}}^{(b)}$ satisfies the usual PCA eigenvalue-eigenvector equation:

$$\frac{1}{N(T+1)} Y^{(b)} Y^{(b)\prime} \hat{\breve{F}}^{(b)} = \hat{\breve{F}}^{(b)} \hat{V}^{(b)},$$

where $\hat{V}^{(b)}$ is the (r, r) diagonal matrix of the r largest eigenvalues of matrix $\frac{1}{N(T+1)} Y^{(b)} Y^{(b)\prime}$, and the columns of matrix $\hat{\breve{F}}^{(b)}$ are the associated normalized eigenvectors such that $\frac{1}{T+1} \hat{\breve{F}}^{(b)\prime} \hat{\breve{F}}^{(b)\prime} = \frac{1}{T+1} \sum_{t=0}^{T} \hat{\breve{f}}_t^{(b)} \hat{\breve{f}}_t^{(b)\prime} = I_r$. [5]

Let $\hat{\breve{\Phi}}^{(b)} = (\sum_{t=1}^{T} \hat{\breve{f}}_t^{(b)} \hat{\breve{f}}_{t-1}^{(b)\prime})(\sum_{t=1}^{T} \hat{\breve{f}}_{t-1}^{(b)} \hat{\breve{f}}_{t-1}^{(b)\prime})^{-1}$ be the OLS estimator of $\breve{\Phi}^{(b)}$, and let $\hat{\breve{v}}_t^{(b)} = \hat{\breve{f}}_t^{(b)} - \hat{\breve{\Phi}}^{(b)} \hat{\breve{f}}_{t-1}^{(b)}$ be the VAR residuals estimated by using $\hat{\breve{\Phi}}^{(b)}$. In this case, the OLS estimator of $\hat{\breve{\Sigma}}_v^{(b)}$ is:

$$\hat{\breve{\Sigma}}_v^{(b)} = \frac{1}{T} \sum_{t=1}^{T} \hat{\breve{v}}_t^{(b)} \hat{\breve{v}}_t^{(b)\prime}.$$

Let $\hat{W}_v^{(b)}$ be the (r, r) matrix collecting the (orthonormal) eigenvectors associated to the ordered eigenvalues $\hat{\sigma}_\ell^{2(b)}$, $\ell = 1, ..., r$, of $\hat{\breve{\Sigma}}_v^{(b)}$:

$$\hat{\breve{\Sigma}}_v^{(b)} \hat{W}_v^{(b)} = \hat{W}_v^{(b)} \hat{\Sigma}_v^{(b)} ,$$

where $\hat{\Sigma}_v^{(b)} := diag(\hat{\sigma}_1^{2(b)}, ..., \hat{\sigma}_r^{2(b)})$ is the diagonal matrix collecting the ordered eigenvalues of $\hat{\breve{\Sigma}}_v^{(b)}$, and $\hat{W}_v^{(b)\prime} \hat{W}_v^{(b)} = \hat{W}_v^{(b)} \hat{W}_v^{(b)\prime} = I_r$. Let us define the estimator $\hat{f}_t^{(b)} := \hat{W}_v^{(b)\prime} \hat{\breve{f}}_t^{(b)}$ of f_t, and matrix $\hat{F}^{(b)} := [\hat{f}_0^{(b)}, \hat{f}_1^{(b)}, ..., \hat{f}_T^{(b)}]' = \hat{W}_v^{(b)\prime} \hat{\breve{F}}^{(b)}$. The (N, r) matrix of estimated loadings $\hat{\Lambda}^{(b)} = [\hat{\lambda}_1^{(b)}, ..., \hat{\lambda}_N^{(b)}]'$

[5] Let $\hat{F}^{(b)*}$ be the orthonormal eigenvectors of $\frac{1}{N(T+1)} Y^{(b)} Y^{(b)\prime}$, s.t. $\frac{1}{N(T+1)} Y^{(b)} Y^{(b)\prime} \hat{F}^{*} = \hat{F}^{(b)*} \hat{V}^{(b)}$ and $\hat{F}^{(b)*\prime} \hat{F}^{(b)*} = I_r$, then the normalized factor estimator $\hat{\breve{F}}^{(b)}$ is computed as $\hat{\breve{F}}^{(b)} = \sqrt{T+1} \cdot \hat{F}^{(b)*}$.

is computed as:

$$\hat{\Lambda}^{(b)} = Y^{(b)\prime}\hat{F}^{(b)}(\hat{F}^{(b)\prime}\hat{F}^{(b)})^{-1}$$

Define

$$\hat{\Phi}^{(b)} = \left(\sum_{t=1}^{T}\hat{f}_t^{(b)}\hat{f}_{t-1}^{(b)\prime}\right)\left(\sum_{t=1}^{T}\hat{f}_{t-1}^{(b)}\hat{f}_{t-1}^{(b)\prime}\right)^{-1} = \begin{bmatrix} \hat{\Phi}_{HH}^{(b)} & \hat{\Phi}_{HL}^{(b)} \\ \hat{\Phi}_{LH}^{(b)} & \hat{\Phi}_{LL}^{(b)} \end{bmatrix}.$$

Let also $\hat{v}_t^{(b)} = \hat{f}_t^{(b)} - \hat{\Phi}^{(b)}\hat{f}_{t-1}^{(b)}$ be the VAR residuals estimated by using $\hat{\Phi}^{(b)}$.

Consider also the estimator of $\Sigma_v^{(b)}$:

$$\hat{\Sigma}_v^{(b)} = \frac{1}{T}\sum_{t=1}^{T}\hat{v}_t^{(b)}\hat{v}_t^{(b)\prime}, \tag{D.1}$$

and let $\hat{\sigma}_\ell^{2(b)}$ be the ℓ-th largest eigenvalue of matrix $\hat{\Sigma}_v^{(b)}$. Then, the sum of the smallest $r-q$ estimated eigenvalues $\hat{\sigma}_{q+1}^{2(b)}, \dots, \hat{\sigma}_r^{2(b)}$ is what we use when constructing $\hat{\xi}^{(b)}(q)$ in equation (5.5).

Starting from $\hat{\Sigma}_u^{(b)}$ in equation (5.6), we can also define the quantities:

$$\begin{aligned}
\hat{B}_U^{(b)} &= \hat{\Sigma}_{u,LL}^{(b)} + \hat{\Phi}_{LH}^{(b)}\hat{\Sigma}_{u,HH}^{(b)}\hat{\Phi}_{LH}^{(b)\prime} + \hat{\Phi}_{LL}^{(b)}\hat{\Sigma}_{u,LH}^{(b)}\hat{\Phi}_{LH}^{(b)\prime} + \hat{\Phi}_{LH}^{(b)}\hat{\Sigma}_{u,HL}^{(b)}\hat{\Phi}_{LL}^{(b)\prime} + \hat{\Phi}_{LL}^{(b)}\hat{\Sigma}_{u,LL}^{(b)}\hat{\Phi}_{LL}^{(b)\prime}, \\
\hat{\Sigma}_U^{(b)}(0) &= \hat{\Sigma}_{u,LL}^{(b)} + \hat{\Phi}_{LH}^{(b)}\hat{\Sigma}_{u,HH}^{(b)}\hat{\Phi}_{LH}^{(b)\prime} + \hat{\Phi}_{LL}^{(b)}\hat{\Sigma}_{u,LH}^{(b)}\hat{\Phi}_{LH}^{(b)\prime} + \hat{\Phi}_{LH}^{(b)}\hat{\Sigma}_{u,HL}^{(b)}\hat{\Phi}_{LL}^{(b)\prime} + \hat{\Phi}_{LL}^{(b)}\hat{\Sigma}_{u,LL}^{(b)}\hat{\Phi}_{LL}^{(b)\prime}, \\
\hat{\Sigma}_U^{(b)}(1) &= -\hat{\Phi}_{LH}^{(b)}\hat{\Sigma}_{u,LH}^{(b)\prime} - \hat{\Phi}_{LL}^{(b)}\hat{\Sigma}_{u,LL}^{(b)\prime}, \qquad \hat{\Sigma}_U^{(b)}(-1) = -\hat{\Sigma}_{u,LH}^{(b)}\hat{\Phi}_{LH}^{(b)\prime} - \hat{\Sigma}_{u,LL}^{(b)}\hat{\Phi}_{LL}^{(b)\prime}, \\
\hat{\Omega}_{U,1}^{(b)} &= 2tr\left\{\hat{\Sigma}_U^{(b)}(0)\hat{\Sigma}_U^{(b)\prime}(0) + \hat{\Sigma}_U^{(b)}(1)\hat{\Sigma}_U^{(b)\prime}(1) + \hat{\Sigma}_U^{(b)}(-1)\hat{\Sigma}_U^{(b)\prime}(-1)\right\},
\end{aligned}$$

which are instrumental to scale and shift $\hat{\xi}^{(b)}(q)$ so as to obtain $\tilde{\xi}^{(b)}(q)$ in equation (5.7).

E Estimators of q proposed in the literature

This section describes estimators of the number of common shocks that we employed in Section 6.3.

E.1 Estimators of Bai and Ng (2007)

As in Bai and Ng (2007) we define:

$$\hat{D}_{1,k} = \left(\frac{\hat{\sigma}_{k+1}^2}{\sum_{\ell=1}^r \hat{\sigma}_\ell^2}\right)^{0.5}, \qquad \hat{D}_{2,k} = \left(\frac{\sum_{\ell=k+1}^r \hat{\sigma}_\ell^2}{\sum_{\ell=1}^r \hat{\sigma}_\ell^2}\right)^{0.5}$$

$K_3 = \left\{k \ : \ \hat{D}_{1,k} < \frac{m_3}{\min\left(N^{0.5-\delta}, T^{0.5-\delta}\right)}\right\}$, and $K_4 = \left\{k \ : \ \hat{D}_{2,k} < \frac{m_4}{\min\left(N^{0.5-\delta}, T^{0.5-\delta}\right)}\right\}$, where $s_{NT} := \min(N,T)$, with $\delta = 0.1$, implying :

$$K_3 = \left\{k \ : \ \hat{D}_{1,k} < \frac{m_3}{s_{NT}^{2/5}}\right\} \qquad K_4 = \left\{k \ : \ \hat{D}_{2,k} < \frac{m_4}{s_{NT}^{2/5}}\right\},$$

Then, the estimator of q considered by Bai and Ng (2007) are:

$$\hat{q}_{bn,3} = \min\left(k \in K_3\right), \qquad \hat{q}_{bn,4} = \min\left(k \in K_4\right) \ . \tag{E.1}$$

Bai and Ng (2007) set either $m_3 = m_4 = 1$, or $m_3 = 1.25$ and $m_4 = 2.25$. The first values are preferable when working with covariance matrices, while the latter are recommended when dealing with correlation matrices.

E.2 Estimators of Amengual and Watson (2007)

Amengual and Watson (2007) define the N-dimensional vectors $\hat{Z}_t^A = [\hat{Z}_{1t}^A, ..., \hat{Z}_{Nt}^A]'$ and $\hat{Z}_t^B = [\hat{Z}_{1t}^B, ..., \hat{Z}_{Nt}^B]'$ as:

$$\hat{Z}_t^A := Y_t - \sum_{i=1}^{p} \hat{\breve{\Lambda}}\hat{\breve{\Phi}}_i \hat{\breve{F}}_{t-i}, \qquad \hat{Z}_t^B := Y_t - \sum_{i=1}^{p} \hat{\breve{\Pi}}\hat{\breve{F}}_{t-i},$$

where $\hat{\breve{\Phi}}_1, \hat{\breve{\Phi}}_2, ..., \hat{\breve{\Phi}}_p$ denote the OLS estimators from the regression of $\hat{\breve{F}}_t$ on $(\hat{\breve{F}}_{t-1}, \hat{\breve{F}}_{t-2}, ..., \hat{\breve{F}}_{t-p})$, while $\hat{\breve{\Pi}}_1, \hat{\breve{\Pi}}_2, ..., \hat{\breve{\Pi}}_p$ are OLS estimators from regressing Y_t on $(\hat{\breve{F}}_{t-1}, \hat{\breve{F}}_{t-2}, ..., \hat{\breve{F}}_{t-p})$. Starting from these new panels, they introduce the estimators

$$\hat{q}_{aw,A} \;=\; \underset{0 \le k \le r}{\arg\min} \left\{ \ln[\hat{\sigma}^2_{\hat{Z}^A} - R(k, \hat{Z}^A)] + k \times \frac{ln[s_{NT}] \cdot (N+T)}{NT} \right\}, \tag{E.2}$$

$$\hat{q}_{aw,B} \;=\; \underset{0 \le k \le r}{\arg\min} \left\{ \ln[\hat{\sigma}^2_{\hat{Z}^B} - R(k, \hat{Z}^B)] + k \times \frac{ln[s_{NT}] \cdot (N+T)}{NT} \right\}, \tag{E.3}$$

where $\hat{\sigma}^2_{\hat{Z}^A} := \frac{1}{NT}\sum_{t=1}^{T}\sum_{i=1}^{N}(\hat{Z}_{it}^A)^2$, $R(k, \hat{Z}^A)$ is defined as

$$R(k, \hat{Z}^A) \;:=\; \sum_{\ell=1}^{k} \omega_\ell^A, \tag{E.4}$$

for ω_i^A the largest ℓ- eigenvalue of $\dfrac{1}{NT}\hat{Z}^A\hat{Z}^{A\prime}$ with $\hat{Z}^A := [\hat{Z}_1^{A\prime}, ..., \hat{Z}_T^{A\prime}]'$. Identical definitions hold for quantities indexed by $\hat{Z}^B$. Because their MC analysis is comparable to ours, we follow Amengual and Watson (2007) and set $p = 2$ in Section 6.3.

E.3 Estimator of Breitung and Pigorsch (2013)

Breitung and Pigorsch (2013) define $\hat{G}_{t-1} := [\hat{\breve{F}}'_{t-1}, \hat{\breve{F}}'_{t-2}, ..., \hat{\breve{F}}'_{t-m}]$, and consider the matrices:

$$\tilde{S}_{00} \;:=\; \sum_{t=m+1}^{T} \hat{\breve{F}}_t \hat{\breve{F}}'_t \qquad \tilde{S}_{01} \;:=\; \sum_{t=m+1}^{T} \hat{\breve{F}}_t \hat{G}'_{t-1}, \qquad \tilde{S}_{11} \;:=\; \sum_{t=m+1}^{T} \hat{G}_{t-1}\hat{G}'_{t-1}. \tag{E.5}$$

Let $\hat{\tilde{R}}_{bp} = \tilde{S}_{00}^{-1}\tilde{S}_{01}\tilde{S}_{11}^{-1}\tilde{S}'_{01}$, then the k largest eigenvalues of $\hat{\tilde{R}}_{bp}$, denoted as $\hat{\rho}^2_{bp,\ell}$, $\ell = 1, ..., k$, are the first squared sample canonical correlations between $\hat{\breve{F}}_t$ and $\hat{G}_{t-1}$, with $k \le r$. They also define the statistic:

$$\begin{aligned}\hat{\xi}_{bp}(k) &= \tilde{C}_{NT}^{2-\delta}\cdot\sum_{\ell=1}^{r-k}(1-\hat{\rho}_{bp,\ell}), \qquad k=1,...,r-1\\ \hat{\xi}_{bp}(r) &= 0\ ,\end{aligned}$$

for $\tilde{C}=\sqrt{NT}/\sqrt{N+T}$. Starting from $\hat{\xi}_{bp}(k)$, they estimate q with:

$$\hat{q}_{bp}=\min\Big(k\ :\ \hat{\xi}_{bp}(k)<\tau\Big)\,.$$

Because $\delta=1/2$ and $\tau=4.5$ deliver good results in their MC simulations, we can write their preferred estimator as:

$$\hat{q}_{bp}=\min\Big(k\ :\ \hat{\xi}^{*}_{bp}(k)<4.5\Big)\,, \tag{E.6}$$

where:

$$\hat{\xi}^{*}_{bp}(k) = \tilde{C}_{NT}^{3/2}\cdot\sum_{\ell=1}^{r-k}(1-\hat{\rho}_{bp,\ell}), \qquad k=1,...,r-1 \tag{E.7}$$

$$\hat{\xi}_{bp}(r) = 0\ . \tag{E.8}$$

F Monte Carlo: additional results

F.1 Alternative data generating processes

This section repeats the Monte Carlo analysis of Section 6 but using alternative DGPs. In particular, we work under the same setting of Section 6.1 but consider different values of (r,q_0) and of the autoregressive matrices. The first alternative DGP, that we call Design 1, sets $r=5$, $q_0=3$, and considers the autoregressive matrix

$$\Phi=diag(0.2,0.375,0.55,0.725,0.9).$$

The second DGP of this section relies on $r=9$, $q_0=8$ and autoregressive matrix

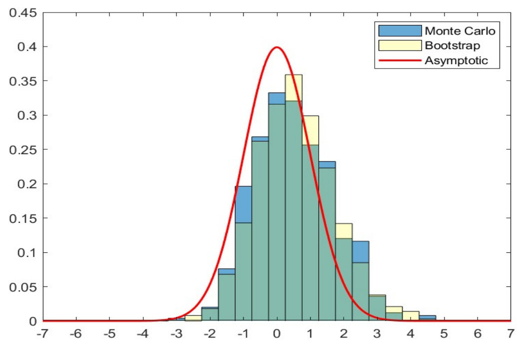
$$\Phi=diag(0.2,0.2875,0.375,0.4625,0.55,0.6375,0.725,0.8125,0.9).$$

Design 1 is very similar to that of Amengual and Watson (2007) and Bai and Ng (2007), while the second one extends it to allow for a richer factor space.

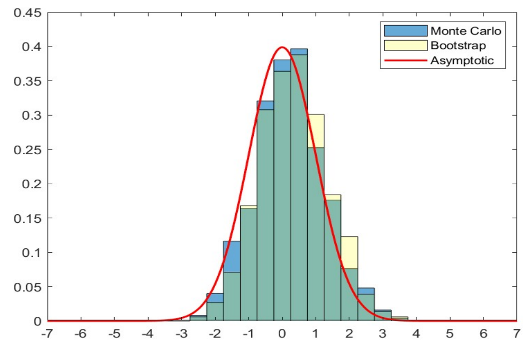
Blue histograms in Figure 1 display the empirical distribution of $\tilde{\xi}(q)$ under the null hypothesis of $q=q_0$ common shocks. Histograms are based on data simulated from Design 1 (first row) and Design 2 (second row). Red solid lines denote the probability density function of the asymptotic $N(0,1)$ distribution. Under Design 1, the empirical distribution is a bit far from the asymptotic one when $(N,T)=(100,100)$. The difference is smaller yet still present when Design 2 is considered.

Results for both designs improve when $(N,T) = (400, 600)$, in which case the empirical distribution becomes quite similar to a standard Gaussian one. The DGP notwithstanding, the distribution based on $N_b = 499$ bootstrap replicates for the first Monte Carlo sample (yellow histogram) provides a more accurate approximation to the empirical one of $\tilde{\xi}(q_0)$. As for the DGP in the main body, summary statistics for $\tilde{\xi}(q_0)$ are reported in Table 13 of Section F.2.

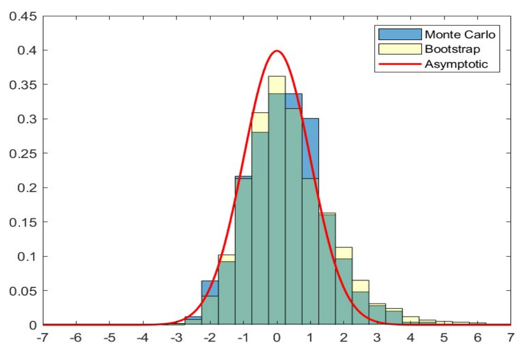
Figure 5 – Small sample and bootstrapped distribution of the test statistic $\tilde{\xi}(q_0)$.



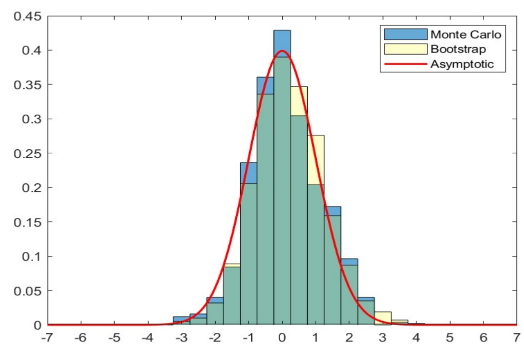
(a) $\tilde{\xi}(3)$, $N = 100$, and $T = 100$; Design 1



(b) $\tilde{\xi}(3)$, $N = 400$, and $T = 600$; Design 1



(c) $\tilde{\xi}(8)$, $N = 100$, and $T = 100$; Design 2



(d) $\tilde{\xi}(8)$, $N = 400$, and $T = 600$; Design 2

Empirical distribution of the test statistic $\tilde{\xi}(q_0)$ for $(N,T) = (100, 100)$ and $(N,T) = (400, 600)$. The first row refers to Design 1, while the second one is based on Design 2. Red solid lines correspond to the asymptotic distribution $N(0, 1)$ of the re-centered and re-scaled statistic.

Tables 6 and 7 exhibit empirical sizes and powers. As far as powers are concerned, we test the null hypotheses $H_0 = H(1)$ and $H_0 = H(2)$ for Design 1, and $H_0 = H(6)$ and $H_0 = H(7)$ for Design 2. The alternative hypothesis is always given by $q > k$ for k the number of factors under the null. Left panels pertain to the asymptotic test, while right ones study the bootstrap procedure. The DGP notwithstanding, the asymptotic test is always oversized and has unit power.[6] Adopting a bootstrap procedure always improves the size of the test at the cost of some loss of power under Design 1.

[6]The asymptotic test consistently returns unit power also when controlling for size distortion.

Table 6 – Empirical sizes and powers of the plug-in and of the bootstrap versions of the test of the number of common shocks; Design 1

		Plug-in: Th. 2					*Bootstrap: Th. 2*				
		size			*power*		*size*			*power*	
N	T	1%	5%	10%	$H(1)$	$H(2)$	1%	5%	10%	$H(1)$	$H(2)$
100	100	0.07	0.16	0.24	1.00	1.00	0.02	0.07	0.14	1.00	0.84
100	200	0.09	0.22	0.31	1.00	1.00	0.01	0.06	0.12	1.00	0.85
200	100	0.03	0.09	0.14	1.00	1.00	0.01	0.07	0.13	1.00	0.90
200	200	0.04	0.11	0.19	1.00	1.00	0.01	0.06	0.11	1.00	0.90
200	300	0.04	0.13	0.21	1.00	1.00	0.01	0.06	0.11	1.00	0.90
400	100	0.02	0.06	0.11	1.00	1.00	0.02	0.07	0.12	1.00	0.93
400	200	0.02	0.06	0.12	1.00	1.00	0.01	0.06	0.12	1.00	0.93
400	300	0.02	0.08	0.13	1.00	1.00	0.01	0.06	0.11	1.00	0.93
400	600	0.03	0.10	0.17	1.00	1.00	0.01	0.05	0.11	1.00	0.93

Empirical sizes and powers of the one-sided test for the null hypothesis of q common shocks. Results in the left panel are based on the plug-in version of the feasible test statistic in Theorem 2. Those in the right panel pertain to the bootstrap counterpart of this test. Simulated data come from Design 1 with $r = 5$ and $q_0 = 3$. Empirical sizes are assessed at significance levels $\alpha \in \{0.01, 0.05, 0.1\}$, while powers represent the empirical rejection frequency of the null hypotheses $H_0 = H(1)$ and $H_0 = H(2)$ under the alternatives $q > 1$ and $q > 2$, respectively. These powers are assessed at the 5% significance level. Results are based on $M = 2000$ MC simulations.

Table 7 – Empirical size and power of the plug-in and of the bootstrap versions of the test of the number of common shocks; Design 2

		Plug-in: Th. 2					*Bootstrap: Th. 2*				
		size			*power*		*size*			*power*	
N	T	1%	5%	10%	$H(6)$	$H(7)$	1%	5%	10%	$H(6)$	$H(7)$
100	100	0.08	0.15	0.22	1.00	0.98	0.04	0.11	0.17	0.99	0.96
100	200	0.12	0.25	0.34	1.00	1.00	0.03	0.08	0.14	0.99	0.98
200	100	0.02	0.07	0.11	1.00	0.99	0.03	0.08	0.15	1.00	0.99
200	200	0.03	0.10	0.16	1.00	1.00	0.02	0.07	0.13	1.00	1.00
200	400	0.04	0.11	0.19	1.00	1.00	0.01	0.07	0.13	1.00	1.00
400	100	0.01	0.03	0.07	1.00	1.00	0.02	0.07	0.14	1.00	1.00
400	200	0.02	0.06	0.10	1.00	1.00	0.01	0.06	0.12	1.00	1.00
400	300	0.02	0.06	0.11	1.00	1.00	0.01	0.06	0.11	1.00	1.00
400	600	0.03	0.09	0.15	1.00	1.00	0.02	0.06	0.11	1.00	1.00

Empirical sizes and powers of the one-sided test for the null hypothesis of q common shocks. Results in the left panel are based on the plug-in version of the feasible test statistic in Theorem 2. Those in the right panel pertain to the bootstrap counterpart of this test. Simulated data come from Design 2 so that $r = 9$ and $q_0 = 8$. Empirical sizes is assessed at significance levels $\alpha \in \{0.01, 0.05, 0.1\}$, while powers represent the empirical rejection frequency of the null hypotheses $H_0 = H(6)$ and $H_0 = H(7)$ under the alternatives $q > 6$ and $q > 7$, respectively. These powers are assessed at the 5% significance level. Results are based on $M = 2000$ MC simulations.

Tables 8, 9 and 10 report empirical sizes and powers when N and T are large for the DGPs of this section and of the main body. Given that bootstrap inference is usually employed when sample sizes are small, we only consider the asymptotic test. Actual sizes are now very close to nominal ones while powers are unaltered.

Table 8 – Empirical sizes and powers of the plug-in version of the feasible test of the number of common shocks when N and T are large; Design 1

		Plug-in: Th. 2				
		size			*power*	
N	T	1%	5%	10%	$H(1)$	$H(2)$
1000	100	0.01	0.05	0.09	1.00	1.00
1000	200	0.01	0.05	0.09	1.00	1.00
1000	300	0.02	0.05	0.10	1.00	1.00
1000	600	0.02	0.07	0.12	1.00	1.00
1000	1000	0.02	0.07	0.13	1.00	1.00
1000	2000	0.02	0.08	0.15	1.00	1.00
2000	100	0.01	0.05	0.09	1.00	1.00
2000	200	0.01	0.05	0.08	1.00	1.00
2000	300	0.01	0.05	0.09	1.00	1.00
2000	600	0.01	0.05	0.09	1.00	1.00
2000	1000	0.01	0.05	0.10	1.00	1.00
2000	2000	0.01	0.06	0.11	1.00	1.00

Empirical sizes and powers of the one-sided test for the null hypothesis of q common shocks. Results are based on the plug-in version of the feasible test statistic in Theorem 2. Simulated data come from Design 1 so that $r = 5$ and $q_0 = 3$. Empirical sizes are assessed at significance levels $\alpha \in \{0.01, 0.05, 0.1\}$, while powers represent the empirical rejection frequency of the null hypotheses $H_0 = H(1)$ and $H_0 = H(2)$ under the alternatives $q > 1$ and $q > 2$, respectively. These powers are assessed for a test performed at the 5% significance level. All empirical probabilities are based on $M = 2000$ MC simulations.

Table 9 – Empirical sizes and powers of the plug-in version of the feasible test of the number of shocks q when N and T are large; Design 2

		Plug-in: Th. 2				
		size			*power*	
N	T	1%	5%	10%	$H(6)$	$H(7)$
1000	100	0.01	0.03	0.04	1.00	1.00
1000	200	0.01	0.03	0.06	1.00	1.00
1000	300	0.01	0.03	0.07	1.00	1.00
1000	600	0.01	0.05	0.09	1.00	1.00
1000	1000	0.01	0.06	0.11	1.00	1.00
1000	2000	0.02	0.08	0.14	1.00	1.00
2000	100	0.01	0.02	0.05	1.00	1.00
2000	200	0.01	0.03	0.06	1.00	1.00
2000	300	0.01	0.03	0.07	1.00	1.00
2000	600	0.01	0.05	0.09	1.00	1.00
2000	1000	0.01	0.04	0.10	1.00	1.00
2000	2000	0.01	0.06	0.10	1.00	1.00

Empirical sizes and powers of the one-sided test for the null hypothesis of q common shocks. Results are based on the plug-in version of the feasible test statistic in Theorem 2. Simulated data come from Design 2 so that $r = 9$ and $q_0 = 8$. Empirical sizes are assessed at significance levels $\alpha \in \{0.01, 0.05, 0.1\}$, while powers represent the empirical rejection frequency of the null hypotheses $H_0 = H(6)$ and $H_0 = H(7)$ under the alternatives $q > 6$ and $q > 7$, respectively. These powers are assessed for a test performed at the 5% significance level. All empirical probabilities are based on $M = 2000$ MC simulations.

Table 10 – Empirical sizes and powers of the plug-in version of the feasible test of the number of shocks when N and T are large; DGP of the main body

		Plug-in: Th. 2				
		size			*power*	
N	T	1%	5%	10%	$H(3)$	$H(4)$
1000	100	0.01	0.03	0.07	1.00	1.00
1000	200	0.01	0.04	0.07	1.00	1.00
1000	300	0.01	0.04	0.09	1.00	1.00
1000	600	0.01	0.05	0.10	1.00	1.00
1000	1000	0.01	0.07	0.13	1.00	1.00
1000	2000	0.03	0.09	0.16	1.00	1.00
2000	100	0.01	0.03	0.05	1.00	1.00
2000	200	0.01	0.03	0.06	1.00	1.00
2000	300	0.01	0.03	0.07	1.00	1.00
2000	600	0.01	0.04	0.09	1.00	1.00
2000	1000	0.01	0.05	0.10	1.00	1.00
2000	2000	0.01	0.06	0.11	1.00	1.00

Empirical sizes and powers of the one-sided test for the null hypothesis of q common shocks. Results are based on the plug-in version of the feasible test statistic in Theorem 2. Simulated data come from the DGP of Section 6.1 with $r = 7$ and $q_0 = 5$. Empirical sizes is assessed at significance levels $\alpha \in \{0.01, 0.05, 0.1\}$., while powers represent the empirical rejection frequency of the null hypotheses $H_0 = H(3)$ and $H_0 = H(4)$ under the alternatives $q > 3$ and $q > 4$, respectively. These powers are assessed for a test performed at the 5% significance level. All empirical probabilities are based on $M = 2000$ MC simulations.

Finally, Tables 11 and 12 repeat the comparison with some alternative estimators already proposed in the literature. The former deals with Design 1 while the latter with Design 2. The estimators of Amengual and Watson (2007) perform very well for Design 1 but tend to underestimate q_0 under Design 2. The approaches of Bai and Ng (2007) and Breitung and Pigorsch (2013) perform on par with our unadjusted bootstrap based estimator (p_α^B) under Design 1; interestingly, the size-adjustment improves the asymptotic-based estimator but hinder performances of the bootstrap one for this data generating process. In the case of Design 2, both bootstrap based procedures, and the (adjusted) asymptotic estimator always outperform all competitors.

Table 11 – Comparison of estimators of q under Design 1

N	T	$\hat{q}_3$	$\hat{q}_4$	$\hat{q}_{aw,A}$	$\hat{q}_{aw,B}$	$\hat{q}_{bp}$	z_α	$z_{\alpha_{N,T}}$	p_α^B	$p_{\alpha_{N,T}^*}^B$
100	100	2.94	2.95	2.98	2.99	2.82	3.19	3.13	2.93	2.88
100	200	2.94	2.95	2.99	2.99	2.88	3.27	3.19	2.93	2.89
200	100	2.94	2.94	2.99	2.99	2.86	3.09	3.05	2.97	2.93
200	200	2.96	2.96	2.99	2.99	2.93	3.13	3.06	2.98	2.94
200	300	2.96	2.96	2.99	3.00	2.94	3.16	3.06	2.98	2.94
400	100	2.94	2.94	2.99	2.99	2.88	3.07	3.03	3.02	2.97
400	200	2.96	2.96	3.00	3.00	2.94	3.07	3.02	3.02	2.98
400	300	2.97	2.97	3.00	3.00	2.96	3.09	3.02	3.01	2.96
400	400	2.98	2.98	3.00	3.00	2.97	3.12	3.03	3.02	2.97

Average estimated number of shocks q under Design 1 so that $r = 5$ and $q_0 = 3$. The third and the fourth columns present results for estimators $\hat{q}_3$ and $\hat{q}_4$ of Bai and Ng (2007). The fifth and sixth columns consider $\hat{q}_{aw,A}$ and $\hat{q}_{aw,B}$ by Amengual and Watson (2007), while the seventh one is based on $\hat{q}_{bp}$ of Breitung and Pigorsch (2013). The eighth and ninth columns show results for our estimator $\hat{q}$ based on the asymptotic sequential testing procedure. The former is based on the 95% quantile of the asymptotic $N(0,1)$ distribution while the latter considers quantiles adjusted for a consistent selection procedure. The last two columns are based on the bootstrap version of the test performed at the 5% significance level: the first one considers unadjusted bootstrap percentiles (i.e. p_α^B) and the second adjusted ones (i.e. $p_{\alpha_{N,T}^*}^B$). We adjust the significance level α using equations (4.9) and (5.9), where we always set $c = 0.95$ and $\gamma = 0.1$. Results are based on $M = 2000$ MC simulations.

Table 12 – Comparison of estimators of q under Design 2

N	T	$\hat{q}_3$	$\hat{q}_4$	$\hat{q}_{aw,A}$	$\hat{q}_{aw,B}$	$\hat{q}_{bp}$	z_α	$z_{\alpha_{N,T}}$	p_α^B	$p_{\alpha_{N,T}^*}^B$
100	100	6.44	6.58	7.50	7.58	6.53	8.13	8.09	8.05	8.01
100	200	6.51	6.63	7.71	7.74	6.77	8.25	8.18	8.05	8.01
200	100	6.42	6.54	7.62	7.70	6.68	8.06	8.03	8.07	8.03
200	200	6.84	6.90	7.77	7.79	6.95	8.10	8.04	8.07	8.03
200	300	6.85	6.90	7.84	7.84	7.06	8.11	8.05	8.06	8.03
400	100	6.41	6.52	7.69	7.75	6.76	8.03	8.01	8.07	8.03
400	200	6.82	6.88	7.84	7.85	7.07	8.06	8.02	8.06	8.03
400	300	7.00	7.04	7.88	7.88	7.20	8.06	8.02	8.06	8.03
400	400	7.12	7.14	7.92	7.92	7.37	8.09	8.02	8.06	8.02

Average estimated number of shocks q under Design 2, i.e. $r = 9$ and $q_0 = 8$. The third and the fourth columns present results for estimators $\hat{q}_3$ and $\hat{q}_4$ of Bai and Ng (2007). The fifth and sixth columns consider $\hat{q}_{aw,A}$ and $\hat{q}_{aw,B}$ by Amengual and Watson (2007), while the seventh one is based on $\hat{q}_{bp}$ of Breitung and Pigorsch (2013). Details on these estimators can be found in Section E. The eighth and ninth columns show results for our estimator $\hat{q}$ based on the asymptotic sequential testing procedure. The former is based on the 95% quantile of the asymptotic $N(0,1)$ distribution while the latter considers quantiles adjusted for a consistent selection procedure. The last two columns are based on the bootstrap version of the test performed at the 5% significance level: the first one considers unadjusted bootstrap percentiles (i.e. p_α^B) and the second adjusted ones (i.e. $p_{\alpha_{N,T}^*}^B$). We adjust the significance level α using equations (4.9) and (5.9), where we always set $c = 0.95$ and $\gamma = 0.1$. Results are based on $M = 2000$ MC simulations.

F.2 Summary statistics for the empirical distribution of $\tilde{\xi}(q)$

Table 13 reports the mean, median, standard deviation and interquartile range for the simulated distribution of the test statistic $\tilde{\xi}(q)$ when $q = q_0$, i.e. the null hypothesis holds. The central panel pertains to the Design of Section 6.1 while the left (right) one is based on Design 1 (2) of Section F.1.

Table 13 – Summary statistics for the empirical distribution of the test statistic $\tilde{\xi}(q_0)$ in Theorem 2.

		$r=5, q_0=3$				$r=7, q_0=5$				$r=9, q_0=8$			
N	T	m.	med.	std.	iqr	m.	med.	std.	iqr	m.	med.	std.	iqr
100	100	0.50	0.44	1.16	1.54	0.52	0.42	1.26	1.62	0.43	0.27	1.52	1.73
100	200	0.81	0.73	1.13	1.47	0.97	0.93	1.19	1.61	0.86	0.77	1.33	1.71
200	100	0.11	0.03	1.08	1.42	0.03	-0.08	1.13	1.48	-0.11	-0.24	1.24	1.51
200	200	0.32	0.27	1.06	1.42	0.30	0.25	1.09	1.51	0.20	0.14	1.11	1.42
200	300	0.43	0.38	1.05	1.38	0.47	0.44	1.10	1.46	0.35	0.31	1.09	1.43
400	100	-0.01	-0.07	1.04	1.34	-0.17	-0.24	1.09	1.45	-0.29	-0.37	1.12	1.34
400	200	0.11	0.08	1.01	1.34	0.02	-0.05	1.05	1.42	-0.08	-0.10	1.05	1.37
400	300	0.17	0.14	1.01	1.34	0.10	0.06	1.04	1.42	0.03	0.00	1.03	1.39
400	600	0.31	0.26	1.02	1.35	0.30	0.29	1.04	1.44	0.24	0.22	1.02	1.37
1000	100	-0.11	-0.13	1.01	1.34	-0.30	-0.34	1.04	1.41	-0.46	-0.54	0.99	1.30
1000	200	-0.04	-0.03	1.02	1.37	-0.19	-0.22	1.01	1.31	-0.30	-0.35	1.00	1.33
1000	300	-0.01	-0.03	1.01	1.36	-0.12	-0.13	1.01	1.36	-0.18	-0.22	1.00	1.36
1000	600	0.08	0.05	1.01	1.33	0.01	-0.01	1.02	1.43	-0.03	-0.07	0.99	1.35
1000	1000	0.16	0.16	1.02	1.40	0.12	0.10	1.00	1.33	0.07	0.04	1.00	1.39
1000	2000	0.25	0.25	1.01	1.33	0.30	0.30	1.00	1.32	0.20	0.19	1.01	1.35
2000	100	-0.16	-0.21	1.03	1.39	-0.36	-0.41	0.98	1.27	-0.49	-0.56	0.99	1.33
2000	200	-0.10	-0.11	1.01	1.36	-0.28	-0.32	0.98	1.28	-0.34	-0.38	1.00	1.31
2000	300	-0.08	-0.13	1.01	1.32	-0.21	-0.23	0.99	1.31	-0.24	-0.29	1.00	1.35
2000	600	-0.01	-0.06	0.99	1.28	-0.11	-0.11	1.01	1.34	-0.08	-0.10	1.01	1.35
2000	1000	0.02	0.03	0.99	1.27	-0.03	-0.06	0.99	1.37	-0.02	-0.05	0.98	1.29
2000	2000	0.10	0.13	0.99	1.32	0.09	0.07	0.99	1.37	0.04	0.04	0.97	1.32

This table reports the mean (*m.*), median (*med.*), standard deviation (*std.*) and interquartile range (*iqr.*) of the empirical distribution of the statistic $\tilde{\xi}(q)$ in Theorem 2. The first four columns pertain to Design 1 in Section F.1 ($r = 5, q_0 = 3$), the second four columns refer to the Design of Section 6.1 ($r = 7$, $q_0 = 5$) and the last four ones are based on Design 2 of Section F.1 ($r = 9$, $q_0 = 8$). Empirical distributions are obtained for different sample sizes (N, T) and using $M = 2000$ MC simulations. The asymptotic distribution of the statistics is always $N(0, 1)$ and has interquartile range of approximately 1.35.

F.3 Comparison with the estimators of Hallin and Liska

In this section, we compare our estimators for the number of common shocks with those of Hallin and Liska (2007). Because the latter are based on frequency domain analysis within the context of generalized dynamic factor models, we do not present them in details.

We combine their information criteria IC_1 and IC_2 with their penalty terms p_1, p_2 and p_3, thus ending up with six different estimators. Their implementation always follows the same steps and modelling choices of Onatski (2009).[7] Comparisons are done for the data generating process of the main body, as well as for Design 1 and Design 2 of the previous sections.

Table 14 presents results for Design 1. The first six columns contain results for the estimators of Hallin and Liska (2007), where HL_{11} labels that based on information criterion IC_1 and penalty term p_1, and similarly for the remaining columns. The unadjusted bootstrap-based estimator is the most accurate one for all sample sizes but $(N, T) = (400, 200)$, in which case there is a slight outperformance from estimator HL_{12}. The latter is also the best performer among all estimators of Hallin and Liska (2007) for most sample sizes. The asymptotic estimator based on the consistent selection procedure also improves upon all estimators of Hallin and Liska (2007) for most combinations of N and T.

[7] We are grateful to Professor Alexey Onatskiy for sharing codes of his paper.

Table 14 – Comparison of estimators of q based on Hallin and Liska (2007) under Design 1

N	T	$HL_{1,1}$	$HL_{1,2}$	$HL_{1,3}$	$HL_{2,1}$	$HL_{2,2}$	$HL_{2,3}$	z_{α}	$z_{\alpha_{N,T}}$	p_{α}^{B}	$p_{\alpha_{N,T}^{*}}^{B}$
100	100	3.80	3.81	3.92	3.94	3.86	3.87	3.19	3.13	2.93	2.88
100	200	3.34	3.18	3.41	3.98	3.97	3.98	3.27	3.19	2.93	2.89
200	100	3.84	3.81	3.92	3.96	3.91	3.91	3.09	3.05	2.97	2.93
200	200	3.20	3.04	3.28	3.98	3.94	3.98	3.13	3.06	2.98	2.94
200	300	3.30	3.16	3.46	3.99	3.98	3.99	3.16	3.06	2.98	2.94
400	100	3.83	3.83	3.92	3.96	3.91	3.92	3.07	3.03	3.02	2.97
400	200	3.18	3.00	3.27	3.97	3.92	3.98	3.07	3.02	3.02	2.94
400	300	3.31	3.18	3.42	3.98	3.97	3.99	3.09	3.02	3.01	2.96
400	600	2.90	2.73	3.03	3.97	3.95	3.98	3.12	3.03	3.02	2.97

Average estimated number of shocks q under Design 1, i.e. $r = 5$ and $q_0 = 3$. Columns $HL_{1,1}$ to $H_{2,3}$ report six possible implementations of the consistent estimator of Hallin and Liska (2007). The ninth and the tenth columns show results for our estimator $\hat{q}$ using the asymptotic sequential testing procedure based either on the 95% quantile of the asymptotic $N(0,1)$ distribution (z_{α}) or on its adjusted counterpart ($z_{\alpha_{N,T}}$). The last two columns are based on the bootstrap version of the test performed at the 5% significance level: the first one considers unadjusted bootstrap percentiles (i.e. p_{α}^{B}) and the second adjusted ones (i.e. $p_{\alpha_{N,T}^{*}}^{B}$). We adjust the significance level α using equations (4.9) and (5.9), where we always set $c = 0.95$ and $\gamma = 0.1$. Results are based on $M = 2000$ MC simulations.

Results for the data generating process of the main body are in Table 15. At least one of our estimators always outperforms those of Hallin and Liska (2007) for all sample sized. Identical results hold for Design 2, as can be seen from Table 16.

Table 15 – Comparison of estimators of q based on Hallin and Liska (2007) under the DGP of main body

N	T	$HL_{1,1}$	$HL_{1,2}$	$HL_{1,3}$	$HL_{2,1}$	$HL_{2,2}$	$HL_{2,3}$	z_{α}	$z_{\alpha_{N,T}}$	p_{α}^{B}	$p_{\alpha_{N,T}^{*}}^{B}$
100	100	5.67	5.62	5.82	5.76	5.61	5.62	5.17	5.12	5.01	4.97
100	200	5.29	5.62	5.37	5.96	5.95	5.96	5.30	5.20	5.03	4.97
200	100	5.71	5.67	5.85	5.84	5.73	5.74	5.07	5.04	5.06	5.01
200	200	5.14	4.95	5.23	5.97	5.93	5.97	5.12	5.06	5.05	5.01
200	300	5.22	5.03	5.38	5.96	5.95	5.96	5.17	5.06	5.05	5.01
400	100	5.76	5.68	5.86	5.86	5.75	5.75	5.04	5.01	5.07	5.03
400	200	5.09	4.94	5.20	5.97	5.93	5.97	5.06	5.02	5.06	5.02
400	300	5.26	5.03	5.40	5.96	5.95	5.96	5.08	5.02	5.06	5.02
400	600	4.82	4.50	4.92	5.97	5.95	5.97	5.12	5.03	5.05	5.02

Average estimated number of shocks q under the DGP of the main body, i.e. $r = 7$ and $q_0 = 5$. Columns $HL_{1,1}$ to $H_{2,3}$ report six possible implementations of the consistent estimator of Hallin and Liska (2007). The ninth and the tenth columns show results for our estimator $\hat{q}$ using the asymptotic sequential testing procedure based either on the 95% quantile of the asymptotic $N(0,1)$ distribution (z_{α}) or on its adjusted counterpart ($z_{\alpha_{N,T}}$). The last two columns are based on the bootstrap version of the test performed at the 5% significance level: the first one considers unadjusted bootstrap percentiles (i.e. p_{α}^{B}) and the second adjusted ones (i.e. $p_{\alpha_{N,T}^{*}}^{B}$). We adjust the significance level α using equations (4.9) and (5.9), where we always set $c = 0.95$ and $\gamma = 0.1$. Results are based on $M = 2000$ MC simulations.

Table 16 – Comparison of estimators of q based on Hallin and Liska (2007) under Design 2

N	T	$HL_{1,1}$	$HL_{1,2}$	$HL_{1,3}$	$HL_{2,1}$	$HL_{2,2}$	$HL_{2,3}$	z_{α}	$z_{\alpha_{N,T}}$	p^B_{α}	$p^B_{\alpha^*_{N,T}}$
100	100	7.65	7.50	7.73	7.22	7.14	7.26	8.13	8.09	8.05	8.01
100	200	7.52	7.34	7.52	7.89	7.86	7.90	8.25	8.18	8.05	8.01
200	100	7.66	7.54	7.75	7.38	7.41	7.48	8.06	8.03	8.07	8.03
200	200	7.33	7.22	7.41	7.94	7.89	7.94	8.10	8.04	8.07	8.03
200	300	7.23	7.09	7.41	7.97	7.97	7.97	8.11	8.05	8.06	8.03
400	100	7.68	7.57	7.78	7.48	7.47	7.54	8.03	8.01	8.07	8.03
400	200	7.30	7.20	7.38	7.94	7.87	7.95	8.06	8.02	8.06	8.03
400	300	7.25	7.05	7.39	7.97	7.96	7.97	8.06	8.02	8.06	8.03
400	600	6.77	6.52	6.92	7.97	7.96	7.98	8.09	8.02	8.06	8.02

Average estimated number of shocks q under the DGP of the main body, i.e. $r = 9$ and $q_0 = 8$. Columns $HL_{1,1}$ to $H_{2,3}$ report six possible implementations of the consistent estimator of Hallin and Liska (2007). The ninth and the tenth columns show results for our estimator $\hat{q}$ using the asymptotic sequential testing procedure based either on the 95% quantile of the asymptotic $N(0,1)$ distribution (z_{α}) or on its adjusted counterpart $(z_{\alpha_{N,T}})$. The last two columns are based on the bootstrap version of the test performed at the 5% significance level: the first one considers unadjusted bootstrap percentiles (i.e. p^B_{α}) and the second adjusted ones (i.e. $p^B_{\alpha^*_{N,T}}$). We adjust the significance level α using equations (4.9) and (5.9), where we always set $c = 0.95$ and $\gamma = 0.1$. Results are based on $M = 2000$ MC simulations.