New rank-based tests and estimators for Common Dynamic Factors *

Federico Carlini[†] Mirco Rubin[‡] Pierluigi Vallarino [§]

This version: February 28, 2024

Abstract

We propose a new rank-based test for the number of common dynamic factors q in a dynamic factor model for a large panel of observations. After estimating a VAR(1) model on r static factors extracted by principal component analysis, we estimate the number of common dynamic factors by testing the rank of the VAR residuals' covariance matrix. Our new rank 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 dynamic factors and to derive its asymptotic properties. We consider applications of our new tests and estimators on panels of macro-financial variables and individual stocks volatilities.

Keywords: Dynamic Factor Model, Test for the Number of Dynamic Factors, Dynamic Factors Estimator, Distribution of Eigenvalues, PCA

*We thank Matteo Barigozzi, Valentina Corradi, Patrick Gagliardini, Silvia Gonçalves, Manfred Deistler, Alex Horenstein, Oliver Linton, Marco Lippi, Serena Ng, Aleksey Onatski and Richard Smith 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 by the Institute for Advanced Studies in Vienna, the 2023 Asian Meeting of the Econometric Society, the 2023 Florence-Paris workshop on Statistics of Random Processes and Its Applications to Financial Econometrics and the seminar participants at the Universities of Cambridge, and Queen Mary in London.

[†]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

This paper proposes a new rank-based test for the number of common dynamic factors q in a dynamic factor models of the type:

$$y_t = \Lambda_0 h_t + \cdots + \Lambda_s h_{t-s} + \varepsilon_t, \quad (1.1)$$

$$h_t = \Phi_1 h_{t-1} + \cdots \Phi_p h_{t-p} + w_t, \quad (1.2)$$

where $y_t = [y_{1,t}, \ldots, y_{N,t}]'$ is an N-dimensional vector of observables, h_t is a q-dimensional vector of latent common dynamic factors which evolve according to the stationary VAR(p) in equation (1.2), ε_t is a vector of weakly correlated zero-mean innovations, w_t is a non-degenerate q-dimensional vector of mutually orthogonal common shocks, Λ_i for $i = 0, \ldots, s$ are $N \times q$ matrices of loadings and Φ_j with $j = 1, \ldots, p$ are autoregressive matrices such that the VAR(p) is stationary. Dynamic factor models have been extensively studied, starting from Stock and Watson (2002a) and Stock and Watson (2002b). When p is finite, the vectors $h_t, \ldots, h_{t-s}$ can be stacked into the $q(s+1)$-dimensional vector $f_t = [h'_t, h'_{t-1}, \ldots, h'_{t-s}]'$ and the model in equations (1.1)-(1.2) has a *static factor* representation:

$$y_t = \Lambda f_t + \varepsilon_t, \quad (1.3)$$

$$f_t = \Phi f_{t-1} + v_t, \quad (1.4)$$

where f_t follows a singular Vector Autorgression of order 1 with innovations $v_t = \begin{bmatrix} I_q & 0 & \ldots & 0 \end{bmatrix}' w_t$ which have a variance-covariance matrix with rank $q < r = q(s+1)$, i.e. matrix $V(v_t)$ is rank deficient. In this setting, estimation and inference for the number of dynamic factors q, and the factors w_t themselves has been studied by Bai and Ng (2007), Amengual and Watson (2007) and Breitung and Pigorsch (2013) and is based on the estimation of the r latent factors f_t by performing Principal Component Analysis (PCA) on the panel of observables y_t.

We contribute to the literature on dynamic factor models by establishing an estimator for the number of dynamic factors q based on a sequential testing procedure of the rank of $V(v_t)$. Our new procedure relies on testing the rank of the residuals' covariance matrix $V(v_t)$ obtained by estimating a VAR(1) model on the Principal Components (PC) estimate $\hat{f}_t$ of the r static factors f_t. In particular, our testing procedure is based on the asymptotic distribution of the sum of the smallest $r - q$ estimated eigenvalues of $V(v_t)$. These are not zero in any finite sample due to the estimation error of the static factors $\hat{f}_t$, which is reflected in the the estimation error of $V(v_t)$ and its eigenvalues. We propose two implementations of this test: a first one based on the plug-in estimation of the bias and the variance of the asymptotic distribution,and a second one relying on a residual (wild) bootstrap of the standardized test statistics. Having estimated q, we also develop an easy-to-implement estimator for the common dynamic factors. Given the eigenvectors associated to the largest q estimated eigenvalues of $V(v_t)$, we estimate dynamic factors by multiplying these eigenvectors with the principal component (PC) estimates of the static factors.

Consistent estimation procedures for the number q of dynamic factors based on Information Criteria (IC) have been derived by relying on the convergence rate of the PC estimator of the common factors by Amengual and Watson (2007), Bai and Ng (2007) and Breitung and Pigorsch (2013). Differently from these procedures, we consider a fully-fledged testing procedure for the number of dynamic factors $q \leq r$. Onatski (2009) tests and estimates the number of dynamic factors by exploiting the asymptotic distribution of ratios of eigen-gaps of the spectral matrix of the data y_t; this test is based on the spectrum of y_t instead of the variance-covariance matrix as it was developed for the *generalized dynamic factor model*. [1] Under the same setting, Hallin and Liska (2007) derive a consistent selection procedure for the number of common dynamic factors. Kapetanios (2010) derives an alternative to the test of Onatski (2009) based on the largest eigenvalues of the covariance of the data y_t. Importantly, our testing procedure allows for more general relative convergence rates of N and T compared those required for the asymptotic results in Hallin and Liska (2007), 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 positive semi-definite (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), all these results fail to hold when the matrix of interest is symmetric and semi-definite. Indeed, Donald, Fortuna, and Pipiras (2007) showed that when the rank of a (negative or positive) semi-definite matrix, say M_0, needs to be estimated using another (negative or positive) semi-

[1] Generalized dynamic factor models, introduced by Forni, Hallin, Lippi, and Reichlin (2000) are extension of the model (1.1) - (1.2) with an infinite number of factor lags in the RHS of equation (1.1) and/or equation (1.2). Hallin and Liska (2007) referred to all dynamic factor models that can expressed with a finite number of lags in equation (1.1) as "restricted dynamic factor model". Our testing methodology can be used in its current version to estimate q in DFM with any finite number of lags s and p in (1.1) - (1.2), but not in a generalized dynamic factor model with an infinite number of lags.

definite matrix, say $\hat{M}$, the asymptotic variance-covariance matrix of the estimator, say $W_0 = V(\hat{M})$, 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 successfully address the problem of testing the rank of a finite dimensional positive semi-definite (p.s.d.) matrix in a panel framework where both the time series and cross-sectional dimension diverge.[2]

Our solution consists in a sequential testing procedure based on the asymptotic distribution of the sum of the smallest $r-q$ estimated eigenvalues of $V(v_t)$. This distribution is derived by using a higher order expansion of the PC estimator of the factors, and applying perturbation methods to construct the asymptotic, for $N, T \to \infty$, expansion of the sum of the smallest $r-q$ estimated eigenvalues of $V(v_t)$. 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 of $N\sqrt{T}$, and features an asymptotic bias term of order $1/N$. The bias term is due to the measurement error in the eigenvalues of $V(v_t)$ caused by the estimation error of the static factors. Its strictly positive sign is compatible with the intuition that the factor estimation error drives the smallest eigenvalues away from zero while maintaining the positive definiteness of the estimator of the covariance matrix of the factor VAR innovations. Starting from this asymptotic distribution, we develop a consistent sequential testing procedure for determining the rank of a p.s.d. matrix. The approach 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 also develop a bootstrap implementation of the sequential testing procedure. Inspired by the work of Goncalves and Perron (2014), we consider a wild bootstrap scheme for the residuals of the factor model, when factors and loadings are estimated by PCA.

We study the empirical size and power of our asymptotic and bootstrap-based testing procedure in a Monte Carlo analysis. We find that the asymptotic test is over-sized unless N and T are particularly large. However, it exhibits a good empirical power even when controlling for the size distortion. The bootstrap scheme drastically refines the size of test while preserving much of its power. These results hold true across different data generating processes. We also look at the average accuracy of the estimator for q based either on the asymptotic or on the bootstrap sequential procedure. The latter provides reliable estimates for all data generating processes that we consider. Finally, we use the new

[2]Other rank-based testing procedures for the number r of static factors in panel data models woth a time observations T is fixed and diverging number N of cross-sectional units N have been recently developed by Fortin, Gagliardini, and Scaillet (2023a,b).

test and related estimator to study the factor structure of US macro-financial variables and of volatility measures on US stocks.

The rest of the paper is organized as follows: Section 2 formally introduces the modelling setting of the paper as well as the identification strategy for the number of common dynamic factors, and for the factors themselves. Section 3 introduces estimators of the static factors, the number q of dynamic factors, and of the dynamic factors themeselves. The large sample theory of these estimators is presented in Section 4. Section 5 introduces the bootstrap test and sequential estimation procedure. Section 6 shows results of different Monte Carlo studies, and Section 7 covers the empirical applications. Section 8 concludes the paper. Appendixes A and B provide the regularity conditions, and the proofs of the Proposition and Theorems, respectively. The Online Appendix (OA), provides the proofs of additional technical results (Section C), an alternative identification and testing strategies of the q common dynamic factors and their number based on the distribution of the largest $r-q$ canonical correlations between contemporaneous and lagged static factors (Section D), all statistical details of the bootstrap implementation of the tests based either on the smallest $r-q$ eigenvalues or on the largest $r-q$ canonical correlations (Section E), a detailed discussion of the alternative estimators of the number of dynamic factors proposed by the literature considered in our MC experiments (Section F), and additional details and results for the Monte Carlo experiments (Section G).[3]

[3] We use the following notation. We partition a generic r-dimensional vector x_t as $x_t = [x'_{Ht}, x'_{Lt}]'$, where the index H (resp. L) indicates that x_{Ht} (resp.x_{Lt}) its the upper (resp. bottom) q -dimensional (resp. $(r-q)$-dimensional) subvector. Moreover, we partition a generic (r,r) matrix A in four blocks as:

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

where A_{HH} its the upper-left (q,q) block, A_{HL} its the upper-right $(q,r-q)$ block, A_{LH} its the bottom-left $(r-q,q)$ block, and A_{LL} its the bottom-right $(r-q,r-q)$. I_r denotes the identity matrix of order r,while vec is the vectorization operator. The acronym p.d. (resp. p.s.d.) means positive definite (resp. positive semi-definite).

2 The model

We consider the following static factor representation of a dynamic factor model:

$$y_t = \breve{\Lambda}\breve{f}_t + \varepsilon_t \tag{2.5}$$

$$\breve{f}_t = \breve{\Phi}\breve{f}_{t-1} + \breve{v}_t, \tag{2.6}$$

$$\breve{v}_t = G\eta_t, \qquad \text{with} \qquad \eta_t \sim iid(0, I_q) \tag{2.7}$$

where $y_t = [y_{1t}, ..., y_{Nt}]'$ is the N-dimensional vector of observations for N individuals at time $t = 1, \ldots, T$, $\breve{\Lambda} = [\breve{\lambda}_1, ..., \breve{\lambda}_N]'$ is the $N \times r$ matrix of factor loadings, $\breve{f}_t$ is the r-dimensional vector of latent static factors with $1 < r \ll N, T$ and $\varepsilon_t = [\varepsilon_{1t}, ..., \varepsilon_{Nt}]'$ is an N-dimensional vector of weakly correlated error terms. Factors $\breve{f}_t$ follow a stationary VAR(1) process where $\breve{\Phi}$ is the $r \times r$ autoregressive matrix, and the r-dimensional innovations' vector $\breve{v}_t$ can be represented as linear combination of the q-dimensional vector of "dynamic factors shocks", or "primitive shocks", η_t which are orthogonal and independent over time, with $1 \leq q \leq r$. The $r \times q$ full-column rank matrix G represents the linear mapping linking the primitive shocks η_t and the static factors $\breve{t}_t$. Equation (2.7) implies

$$\breve{v}_t \sim iid(0, \breve{\Sigma}_v), \tag{2.8}$$

where $\breve{\Sigma}_v := E(\breve{v}_t\breve{v}_t') = GI_qG' = GG'$ is the covariance matrix of the primitive shocks $\breve{v}_t$.[4] Critically, if $q < r$ then the $r \times r$ matrix $\breve{\Sigma}_v$ has reduced rank q. Let σ_ℓ^2 be the ℓ-th largest eigenvalue of $\breve{\Sigma}_v$, with $\ell = 1, ..., r$. When $q < r$, the smallest $r - q$ eigenvalues of $\breve{\Sigma}_v$ are equal to zero, while its largest q eigenvalues are strictly positive, i.e. $\sigma_1^2 \geq \sigma_2^2 \geq ... \geq \sigma_q^2 > \sigma_{q+1}^2 = \sigma_{q+2}^2 = ... = \sigma_r^2 = 0$. The main result in our paper consists in the derivation of the asymptotic distribution of the smallest $r - q$ eigenvalues when the VAR(1) model in (2.6) is estimated on static factors $\breve{f}_t$ obtained as Principal Components from the large panel of observables y_t.

[4]The assumption $V(\eta_t) = I_q$ implicit in equation (2.7) is simply an identification condition for the dynamic factors, as different values of variance and (non-perfect) correlation among them can be obtained by appropriate values of the coefficients in the generic $r \times q$ full-column rank matrix G.

2.1 Identification of static and dynamic factors

Our identification strategy for the number of dynamic shocks is based on the number of zero eigenvalues of $\breve{\Sigma}_v$ in model (2.5) - (2.8), i.e. the innovation's covariance matrix of VAR(1) for the r static factors $\breve{f}_t$, which are unobservable. Similarly, the dynamic factors and shocks are then identified by using the eigenvectors of $\breve{\Sigma}_v$ associated to its largest q eigenvalues.

It is well known, that the static factors $\breve{f}_t$ are identified, up to a rotation (and change of sign) by performing PCA on the panel of observables y_t, if the following standard identification assumptions for linear latent factor models are made

$$E(\breve{f}_t) = 0 \qquad \text{and} \qquad V(\breve{f}_t) = E(\breve{f}_t \breve{f}_t') = I_r, \tag{2.9}$$

where the zero mean assumption of the factors can be made as we do not include any (vector of) intercepts in the r.h.s. of the model's equations (2.5) and (2.6). We refer to (2.9) as Assumption A.2 ii) in the list of regularity conditions in Appendix A.

Turning to the identification of the dynamic factors, when $\breve{\Sigma}_v$ has reduced rank $q < r$, there exists an equivalent way of expressing the Data Generating Process (DGP) in (2.5) - (2.8) which i) allows to identify the q common *dynamic* primitive shocks η_t, and ii) simplifies the derivation of the distribution of the test statistics for their number q. Let Σ_v be the $r \times r$ diagonal matrix collecting the the ordered eigenvalues σ_ℓ^2, with $\ell = 1, ..., r$, of $\breve{\Sigma}_v$:

$$\Sigma_v := diag(\sigma_1^2, ..., \sigma_q^2, 0, ..., 0), \tag{2.10}$$

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

$$\breve{\Sigma}_v W_v = W_v \Sigma_v\ , \qquad \text{with} \qquad W_v' W_v = W_v W_v' = I_r. \tag{2.11}$$

Let us define the rotated factors and their associated loadings as

$$f_t = \left[f_{H,t}',\ \ f_{L,t}'\right]' := W_v' \breve{f}_t,\ \ t = 1, ..., T \qquad \text{and} \qquad \Lambda = [\lambda_1', ..., \lambda_N']' := \breve{\Lambda} W_v \tag{2.12}$$

with $\lambda_i = W_v' \breve{\lambda}_i$ for $i = 1, ..., N$, respectively. Then, the DGP for the observable variables y_t in equation (2.5) can be equivalently written as:

$$y_t = \Lambda f_t + \varepsilon_t. \tag{2.13}$$

By defining $\Phi := W_v' \breve{\Phi} W_v$, the rotated innovation vector $v_t := W_v' \breve{v}_t = [(W_{v,q}' G \eta_t)' \ (W_{v,q}' G \eta_t)']'$, and by premultiplying both sides of equation (2.6) 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:

$$f_t = \Phi f_{t-1} + v_t, \qquad \text{with} \qquad v_t \sim iid(0, \Sigma_v) \tag{2.14}$$

and $\Sigma_v := V(v_t) = W_v' \breve{\Sigma}_v W_v$. By definition (2.10), the lower $(r-q)$ block $v_{t,L}$ of v_t is distributed as a degenerate multivariate random variable such that $E[v_{t,L}] = 0$ and $V(v_{t,L}) = \Sigma_{v,LL} = 0_{(r-q)\times(r-q)}$, which implies:

$$v_t = [v_{Ht}' \ , \ v_{Lt}']' = [v_{Ht}' \ , \ 0_{(r-q)\times 1}]' \qquad \forall t. \tag{2.15}$$

Equivalently, the rotated DGP for the rotated factors $f_t = [f_{Ht}' \ f_{Lt}']'$ in (2.14) can be re-written as:

$$\begin{bmatrix} f_{Ht} \\ f_{Lt} \end{bmatrix} = \begin{bmatrix} \Phi_{HH} & \Phi_{HL} \\ \Phi_{LH} & \Phi_{LL} \end{bmatrix} \begin{bmatrix} f_{Ht-1} \\ f_{Lt-1} \end{bmatrix} + \begin{bmatrix} v_{Ht} \\ 0 \end{bmatrix}, \qquad t = 1, ..., T. \tag{2.16}$$

In this rotated DGP, the non-degenerate q-dimensional 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 $q \times q$ matrix $W_{v,q}' G$. On the other hand, the *degenerate* $r-q$ factors collected in f_{Lt} have degenerate innovations $v_{Lt} = W_{v,r-q}' G \eta_t = 0$ for all dates t: this special feature of the innovations of f_{Lt} is key in deriving the asymptotic distribution of our test statistic for q. [5]

[5] The special form of the innovations in model (2.16) also implies that there exist $r-q$ 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. $\breve{f}_{t-1}$). In Appendix D we show that the dynamic factor space dimension q is also identifiable using canonical correlation analysis applied to $\breve{f}_t$ and $\breve{f}_{t-1}$. A similar argument was first used by Breitung and Pigorsch (2013): we formalize it in Appendix D, where we also show the analogies with our identification argument in this Section, and estimation and inference for the eigendecomposition of $\breve{\Sigma}_v$ proposed in the following sections.

3 Estimators

In Section 3.1 we show that the OLS estimator of $\breve{\Sigma}_v$ (or Σ_v) has $r - q$ zero eigenvalues when the true factor $\breve{f}_t$ (or f_t) are observed. In Section 3.2 we introduce estimators for the static and dynamic factors, and for parameters of their VAR model when the static factors are not observed, but instead estimated by PCA. In particular, we discuss the estimation of the eigenvalues and eigenvectors of $\breve{\Sigma}_v$. In Section 3.3 we discuss the sequential testing strategy characterizing our estimator of the number q of common dynamic factors, and the test statistics.

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 the observable factor is $\breve{f}_t$, and let $\tilde{\breve{v}}_t = \breve{f}_t - \tilde{\breve{\Phi}} \breve{f}_{t-1}$ be the VAR residuals estimated by 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 the factors f_t are observable, and $\tilde{v}_t = f_t - \tilde{\Phi} f_{t-1}$ be the VAR residuals estimated by using $\tilde{\Phi}$. In this case, the OLS estimator of Σ_v is:

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

Note that both estimators $\tilde{\breve{\Sigma}}_v$ and $\tilde{\Sigma}_v$ are unfeasible when the true factors are not observed.

3.2 Estimation when factors are unobserved

Let us first assume that the true number of static factors r is known, but the true factors $\breve{f}_t$ are unobservable and q is unknown. Our estimation procedure for q starts by estimating factors $\breve{f}_t$ by PCA. Let $\hat{\breve{F}} = [\hat{\breve{f}}_0, \hat{\breve{f}}_1, ..., \hat{\breve{f}}_T]'$ be the $(T+1, r)$ matrix of estimated Principal Components (PCs) extracted from the $(T+1, N)$ panel $Y = [y_0, y_1, ..., y_T]'$ associated with the largest r eigenvalues of matrix

$\frac{1}{N(T+1)}YY'$. That is $\hat{\breve{F}}$ satisfies the usual PCA eigenvalue-eigenvector equation:

$$\frac{1}{N(T+1)}YY'\hat{\breve{F}} = \hat{\breve{F}}\hat{V}, \tag{3.4}$$

where $\hat{V}$ is the (r, r) diagonal matrix containing the r largest sorted eigenvalues of matrix $YY'/(N(T+1))$, and the columns of matrix $\hat{\breve{F}}$ are the associated normalized eigenvectors such that $\frac{1}{T+1}\hat{\breve{F}}'\hat{\breve{F}} = \frac{1}{T+1}\sum_{t=0}^{T}\hat{\breve{f}}_t\hat{\breve{f}}_t' = I_r$. [6]

Let $\hat{\breve{\Phi}} = (\sum_{t=1}^{T}\hat{\breve{f}}_t\hat{\breve{f}}_{t-1}')(\sum_{t=1}^{T}\hat{\breve{f}}_{t-1}\hat{\breve{f}}_{t-1}')^{-1}$ be the OLS estimator of $\breve{\Phi}$ when the factor is $\breve{f}_t$ is estimated by PCA, and let $\hat{\breve{v}}_t = \hat{\breve{f}}_t - \hat{\breve{\Phi}}\hat{\breve{f}}_{t-1}$ be the VAR residuals estimated by using $\hat{\breve{\Phi}}$. In this case, the OLS estimator of $\breve{\Sigma}_v$ is:

$$\hat{\breve{\Sigma}}_v = \frac{1}{T}\sum_{t=1}^{T}\hat{\breve{v}}_t\hat{\breve{v}}_t'. \tag{3.5}$$

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 $\hat{\breve{\Sigma}}_v$:

$$\hat{\breve{\Sigma}}_v\hat{W}_v = \hat{W}_v\hat{\Sigma}_v\ , \tag{3.6}$$

where $\hat{\Sigma}_v := diag(\hat{\sigma}_1^2, ..., \hat{\sigma}_r^2)$ is the diagonal matrix collecting the sorted eigenvalues of $\hat{\breve{\Sigma}}_v$, and $\hat{W}_v'\hat{W}_v = \hat{W}_v\hat{W}_v' = I_r$. From the definition of rotated factors in (2.12), we can define the estimator $\hat{f}_t := \hat{W}_v'\hat{\breve{f}}_t$ of $f_t = W_v'\breve{f}_t$, and matrix $\hat{F} := [\hat{f}_0, \hat{f}_1, ..., \hat{f}_T]' = \hat{\breve{F}}\hat{W}_v$. Analogously, we can define the estimator $\hat{v}_t := \hat{W}_v'\hat{\breve{v}}_t$ 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 $\hat{f}_{H,t} = \hat{W}_{v,q}'\hat{\breve{f}}_t$, and the estimator of the q dynamic factors $v_{H,t}$ is $\hat{v}_{H,t} := \hat{W}_{v,q}'\hat{\breve{v}}_t$, for all $t = 1, ..., T$.*

Then, the loadings Λ of model (2.13) are estimated by the time-series regressions of y_{it} on the estimated factors $\hat{f}_t$. Therefore, the $N \times r$ matrix of estimated loadings $\hat{\Lambda} = [\hat{\lambda}_1, ..., \hat{\lambda}_N]'$ is computed as:

$$\hat{\Lambda} = Y'\hat{F}(\hat{F}'\hat{F})^{-1} = \frac{1}{T+1}Y'\hat{F}, \tag{3.7}$$

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

[6] 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 $\hat{\breve{F}}$ is computed as $\hat{\breve{F}} = \sqrt{T+1}\cdot\hat{F}^*$.

Let $\hat{\varepsilon}_t = y_t - \hat{\Lambda}_j \hat{f}_t$ be the N-dimensional vector of residuals of the regressions of y_{it} on $\hat{f}_t$, and let $\hat{\Xi} := [\hat{\varepsilon}_0, \hat{\varepsilon}_1, ..., \hat{\varepsilon}_T]' = Y - \hat{F}\hat{\Lambda}$ the (T, N) matrix collecting all residuals.

Let $\hat{\Phi}$ be the OLS estimator of Φ in the rotated factor VAR(1) in equation (2.14):

$$\hat{\Phi} = \left(\sum_{t=1}^{T} \hat{f}_t \hat{f}_{t-1}' \right) \left(\sum_{t=1}^{T} \hat{f}_{t-1} \hat{f}_{t-1}' \right)^{-1} = \begin{bmatrix} \hat{\Phi}_{HH} & \hat{\Phi}_{HL} \\ \hat{\Phi}_{LH} & \hat{\Phi}_{LL} \end{bmatrix}. \tag{3.8}$$

By using the definitions of $\hat{f}_t$, $\hat{W}_v$ and $\hat{v}_t$, it is easy to show that $\hat{v}_t = \hat{f}_t - \hat{\Phi}\hat{f}_{t-1}$ for all $t = 1, ..., T$ i.e. $\hat{v}_t$ are the VAR(1) residuals estimated on the rotated estimated factors $\hat{f}_t$. Then, it also follows that $(1/T)\sum_{t=1}^{T} \hat{v}_t \hat{v}_t'$, i.e. the consistent estimator of Σ_v, is equal to $\hat{\Sigma}_v$:

$$\frac{1}{T} \sum_{t=1}^{T} \hat{v}_t \hat{v}_t' = \hat{\Sigma}_v. \tag{3.9}$$

Equations (3.6) and (3.9) imply that the ℓ-th eigenvalue of $\breve{\hat{\Sigma}}_v$ is equal to the element in position (ℓ, ℓ) of the estimator $\frac{1}{T}\sum_{t=1}^{T} \hat{v}_t \hat{v}_t'$ of Σ_v (which is itself a diagonal matrix of ordered eigenvalues). Importantly, in Section 4 we show that all estimated eigenvalues $\hat{\sigma}_\ell^2$ are strictly positive w.p.a. 1, for all $\ell = 1, ..., r$, as the smallest $r - q$ eigenvalues are functions of the estimation error in the principal component estimator $\breve{\hat{f}}_t$.

3.3 Sequence of tests of hypotheses on the number of dynamic factors

From Section 2, the number of dynamic factors q coincides with the number of non-zero eigenvalues of matrix $\breve{\Sigma}_v$, which are contained in the diagonal matrix Σ_v. In order 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. [7]

[7] Note that the assumption $1 \leq q \leq r$ implies that there exists at least one factor in our model (2.5), and therefore we do not consider the degenerate case $H(0) = \left\{\sigma_1^2 = \sigma_2^2 = ... = \sigma_r^2 = 0\right\}$, which corresponds to the absence of any dynamic and static factor. This degenerate case is easy to detect empirically by applying the usual tests for the number of static factors mentioned below.

Table 1 – Hypotheses on the number of dynamic factors q

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

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

$$H_0 = H(q) \qquad \text{vs.} \qquad H_1 = \bigcup_{q<s\leq r} H(s), \qquad \text{for } \ q = 1, 2, ..., r-1 \ . \tag{3.10}$$

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

$$\hat{\xi}(q) = \sum_{\ell=1}^{r-q} \hat{\sigma}_{r-\ell+1}^2, \tag{3.11}$$

which corresponds to the sum of the $r-q$ smallest sample eigenvalues of $\hat{\breve{\Sigma}}_v$, that is $\hat{\xi}(q) = \hat{\sigma}_{q+1}^2 + \hat{\sigma}_{q+2}^2 + ... + \hat{\sigma}_r^2$. We reject the null $H_0 = H(q)$ when $\hat{\xi}(q)$ is positive and large, corresponding to the case that at least one of the $r-q$ smallest eigenvalues is significantly different from zero. Critical values of the test are obtained from the large sample distribution of the statistic when $N, T \to \infty$, which is derived in Section 4.2. The number of primitive shocks q is estimated by sequentially applying the tests for the null $H(q)$ starting from $q = 1$, progressively increasing q if the null $H(q)$ is rejected, and stopping the procedure for the smallest value of q for which the null hypothesis $H(q)$ is not rejected, as described in Section 4.3.

When the true number of static factors r is unknown, but a consistent estimators $\hat{r}$ is available, the asymptotic distribution and rate of convergence for the test statistic $\hat{\xi}(q)$ based on $\hat{r}$ is the same as those based on the true number of factors.[8] Therefore, the asymptotic distributions and rates of

[8] See, e.g., the discussion in Section 3.2 in AGGR and the reference therein.

convergence of the test statistics will be derived assuming that the true dimension r is known, as also done in Amengual and Watson (2007) and Bai and Ng (2007). [9] Examples of consistent estimators for the numbers of pervasive factors r are, e.g., those proposed by Bai and Ng (2002), Alessi, Barigozzi, and Capasso (2010), Onatski (2010), Ahn and Horenstein (2013) and Trapani (2018).

4 Large sample theory

In Section 4.1 we show that the OLS estimator of Σ_v has $r-q$ zero eigenvalues when the true factor $\breve{f}_t$ (or f_t) are observed, that is they have a degenerate finite sample distribution distribution, for any finite sample of dimension T, with mean and variance both equal to 0. This implies that when all factors are observed without error, that is they do not need to be estimated, testing for the number of dynamic factors is a degenerate problem. Then, in Section 4.2 we derive the large sample distribution of $\hat{\xi}(q)$, which is the sum of the smallest $r-q$ eigenvalues $\hat{\sigma}_\ell^2$ of $\hat{\breve{\Sigma}}_v$, and provide an implementation of the test based on consistent plug-in estimators of its asymptotic bias and variance. We also define a consistent selection procedure for the number of primitive shocks q.

4.1 Distribution of eigenvalue estimators when static factors are observed

We first study the eigenvalues of the OLS estimator matrices $\tilde{\breve{\Sigma}}_v$ (resp. $\tilde{\Sigma}_v$) of $\breve{\Sigma}_v$ (resp. Σ_v), obtained by estimating the VAR(1) model in equation (2.6) (resp. (2.14)) by OLS from the T-dimensional sample of true 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 of the true factors $\breve{f}_t$ (resp. f_t) generated by model* (2.6)-(2.8) *(resp. model* (2.14)*), and let the $r\times r$ matrix $\tilde{\breve{\Sigma}}_v$ (resp. $\tilde{\Sigma}_v$) be the OLS estimator of $\breve{\Sigma}_v$ (resp. Σ_v) defined in* (3.1) *(resp.* (3.3)*) based on the T observations of $\breve{f}_t$ (resp. f_t).*

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

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

[9] As 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.

for any $T \geq r^2$. (ii) The smallest $r - q$ (resp. largest q) sample eigenvalues of both matrices $\tilde{\Sigma}_v$ and $\tilde{\breve{\Sigma}}_v$ are equal to (resp. strictly larger than) zero, i.e. as $T \to \infty$

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

(iii) Assume that the largest q eigenvalues $\sigma_1^2, \ldots, \sigma_q^2$ 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 the conventional convergence rate $\sqrt{T}$, that is $\sqrt{T}(\tilde{\sigma}_\ell^2 - \sigma_\ell^2) \xrightarrow{d} N(0, V_{asy}(\tilde{\sigma}_\ell^2))$, as $T \to \infty$, where $V_{asy}(\tilde{\sigma}_\ell^2) = e_{q,\ell}'(e_{q,\ell}' \otimes I_q) \cdot V_1 \cdot (e_{q,\ell} \otimes I_q) e_{q,\ell}$, 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 if the factors $\breve{f}_t$ (or f_t) were observable, there would be no need for a test for the number of dynamic factors. Indeed, simply looking at the eigenvalues of $\tilde{\breve{\Sigma}}_v$ or $\tilde{\Sigma}_v$, computed from any (sufficiently large) sequence of observations og $\breve{f}_t$ (or f_t) would allow exact identification of the number of primitive shocks: 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 example of 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. Hence, if the true factors $\breve{f}_t$ were observed, testing for the number of common dynamic shocks - by testing for the number of zero eigenvalues of Σ_v by using those of $\tilde{\Sigma}_v$- results in a degenerate problem, as it would involve testing for the existence of deterministic relationships between the random vectors f_t and f_{t-1}.

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 $r - q$. Thus, this result refines the claim in Section 2 of Bai and Ng (2007) on the eigenvalues of $\tilde{\breve{\Sigma}}_v$: 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 the same $r - q$ smallest eigenvalues converge to 0 as $T \to \infty$. Moreover, as shown in Sections 3.2 and 4, 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, and converge to 0 only asymptotically when $N, T \to \infty$. Importantly, the smallest $r - q$ eigenvalues converge to zero at rate

$N\sqrt{T}$. On the other hand, the largest q eigenvalues of $\hat{\hat{\Sigma}}_v$ converge faster at rate $\min(\sqrt{N}, \sqrt{T})$, as in Bai and Ng (2007).

4.2 Distribution of eigenvalue estimators when static factors are estimated by PCA

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

$$\sqrt{T}/N = o(1), \qquad N/T^{5/2} = o(1), \tag{4.3}$$

which we refer to as Assumption A.1 in the list of regularity conditions in Appendix A. The conditions in (4.3) allow for a wide range of relative growth rates for the time-series and cross-sectional panel dimensions as long as N grows faster than $T^{1/2}$ and slower than $T^{5/2}$. To derive the large sample distribution of the test statistic for the number of common factors we deploy the refined asymptotic expansion for the estimated PCs derived by AGGR. This expansion extends results in Bai and Ng (2002), Stock and Watson (2002a), 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 $\hat{\breve{f}}_t$ is asymptotically equivalent (see details in Proposition B.1), up to negligible terms, to $\hat{\mathcal{H}}\left(\breve{f}_t + \frac{1}{\sqrt{N}}\breve{u}_t + \frac{1}{T}\breve{b}_t\right)$, where $\breve{u}_t = \left(\frac{1}{N}\sum_{i=1}^N \breve{\lambda}_i\breve{\lambda}_i'\right)^{-1} \frac{1}{\sqrt{N}}\sum_{i=1}^N \breve{\lambda}_i\varepsilon_{i,t}$, $\hat{\mathcal{H}}$ is a nonsingular stochastic factor rotation matrix, $\breve{b}_t = \left(\frac{1}{N}\sum_{i=1}^N \breve{\lambda}_i\breve{\lambda}_i'\right)^{-1}\left(\frac{1}{T}\sum_{t=1}^T \breve{f}_t\breve{f}_t'\right)^{-1}\eta_t^2\breve{f}_t$, and $\eta_t^2 = \operatorname*{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^N E[\varepsilon_{i,t}^2|\mathcal{F}_t]$ is the limit average error variance conditional on the sigma field $\mathcal{F}_t = \sigma(\breve{f}_s, s \le t)$ generated by current and past factor values $\breve{f}_t$. The zero-mean term $\breve{u}_t$ drives the randomness in factor estimates conditional on factor path. Vector $\breve{b}_t$ is measurable with respect to the factor path and induces a bias term at order T^{-1} in principal components estimates. [10] Let us also define $u_t := \left(\frac{1}{N}\sum_{i=1}^N \lambda_i\lambda_i'\right)^{-1}\frac{1}{\sqrt{N}}\sum_{i=1}^N \lambda_i\varepsilon_{i,t}$.

Let $\tilde{\Sigma}_{u,t}(h) = Cov(u_t, u_{t-h}|\mathcal{F}_t)$ be the conditional covariance between u_t and u_{t-h}, i.e.

$$\tilde{\Sigma}_{u,t}(h) = \left(\frac{1}{N}\sum_{i=1}^N \lambda_i\lambda_i'\right)^{-1}\frac{1}{N}\sum_{i=1}^N\sum_{\ell=1}^N \lambda_i\lambda_\ell' Cov(\varepsilon_{i,t}, \varepsilon_{\ell,t-h}|\mathcal{F}_t)\left(\frac{1}{N}\sum_{i=1}^N \lambda_i\lambda_i'\right)^{-1},$$

and $\tilde{\Sigma}_{u,t}(-h) = \tilde{\Sigma}_{u,t}(h)'$, for h= 0, 1, …. We set $\tilde{\Sigma}_{u,t} \equiv \tilde{\Sigma}_{u,t}(0)$., and define $\Sigma_{u,t}(h) = \operatorname*{plim}_{N\to\infty}\tilde{\Sigma}_{u,t}(h)$

[10] Vectors $\breve{u}_t$ and $\breve{b}_t$ depend on sample sizes but, for convenience, we omit the indices N, 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:*

$$N\sqrt{T}\Omega_{U,1}^{-1/2}\left[\hat{\xi}(q) - \frac{1}{N}tr\left\{B_U\right\}\right] \xrightarrow{d} N\left(0,1\right), \tag{4.4}$$

with $\Omega_{U,1} = 2\sum_{h=-\infty}^{\infty} E\left[tr\left\{\Sigma_{U,t}(h)\Sigma_{U,t}(h)'\right\}\right]$,

$$\begin{aligned}
\Sigma_{U,t}(h) &= \Sigma_{u,t,LL}(h) - \tilde{\Phi}_{LH}\Sigma_{u,t,LH}(h-1)' - \tilde{\Phi}_{LL}\Sigma_{u,t,LL}(h-1)' \\
&\quad -\Sigma_{u,t,LH}(h+1)\tilde{\Phi}_{LH}' + \tilde{\Phi}_{LH}\Sigma_{u,t,HH}(h)\tilde{\Phi}_{LH}' + \tilde{\Phi}_{LL}\Sigma_{u,t,LH}(h)\tilde{\Phi}_{LH}' \\
&\quad -\Sigma_{u,t,LL}(h+1)\tilde{\Phi}_{LL}' + \tilde{\Phi}_{LH}\Sigma_{u,t,HL}(h)\tilde{\Phi}_{LL}' + \tilde{\Phi}_{LL}\Sigma_{u,t,LL}(h)\tilde{\Phi}_{LL}' \ , \qquad h = ..., -1, 0, 1, ... \ , \\
B_U &= \tilde{\Sigma}_{u,t,LL}(0) - \tilde{\Phi}_{LH}\tilde{\Sigma}_{u,t,LH}(-1)' - \tilde{\Phi}_{LL}\tilde{\Sigma}_{u,t,LL}(-1)' \\
&\quad -\tilde{\Sigma}_{u,t,LH}(+1)\tilde{\Phi}_{LH}' + \tilde{\Phi}_{LH}\tilde{\Sigma}_{u,t,HH}(0)\tilde{\Phi}_{LH}' + \tilde{\Phi}_{LL}\tilde{\Sigma}_{u,t,LH}(0)\tilde{\Phi}_{LH}' \\
&\quad -\tilde{\Sigma}_{u,t,LL}(+1)\tilde{\Phi}_{LL}' + \tilde{\Phi}_{LH}\tilde{\Sigma}_{u,t,HL}(0)\tilde{\Phi}_{LL}' + \tilde{\Phi}_{LL}\tilde{\Sigma}_{u,t,LL}(0)\tilde{\Phi}_{LL}' \ .
\end{aligned}$$

Proof: See Appendix B.2.

Matrix $\Sigma_{U,t}(h)$ is the limit covariance matrix between the $(r-q)$-dimensional vector $u_{Lt}+\tilde{\Phi}_{LH}u_{Ht-1}+\tilde{\Phi}_{LL}u_{Lt-1}$ and the $(r-q)$-dimensional vector $u_{Lt-h} + \tilde{\Phi}_{LH}u_{Ht-h-1} + \tilde{\Phi}_{LL}u_{Lt-h-1}$. The estimation error of the PC estimator of the factors and its lagged value determines the asymptotic distribution of the statistic.

The asymptotic Gaussian distribution when testing a hypothesis for parameters at their boundary, i.e. eigenvalues of a positive definite matrix to be equal to zero, 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} generated by the sampling error in the first step estimates of the PCs.

To get a feasible distributional result for the statistic $\hat{\xi}(q)$, we need consistent estimators for the unknown scalars $tr\{B_U\}$ and $\Omega_{U,1}$, and matrices $\tilde{\Phi}_{LH}$ and $\tilde{\Phi}_{LH}$ in Theorem 1. The natural estimators of the latter two matrices are the corresponding blocks of the feasible estimator $\hat{\Phi}$ provided in equation (3.8), namely $\hat{\Phi}_{LH}$ and $\hat{\Phi}_{LL}$. To estimate the first two scalars, at this stage and as in AGGR, 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 homoschedastic martingale difference sequence for each individual i, conditional on the factor path, i.e.

$$\begin{aligned} Cov(\varepsilon_{i,t}, \varepsilon_{j,t-h}|\mathcal{F}_t) &= 0, \qquad \text{if} \qquad i \neq j, \\ E[\varepsilon_{i,t}|\{\varepsilon_{i,t-h}\}_{h\geq 1}, \mathcal{F}_t] &= 0, \qquad E[\varepsilon_{i,t}^2|\{\varepsilon_{i,t-h}\}_{h\geq 1}, \mathcal{F}_t] = \gamma_{ii} \quad \text{(say)}, \end{aligned} \tag{4.5}$$

for all i, t, h (see Assumption A.9). These assumptions imply $\tilde{\Sigma}_{u,t} = \Sigma_{u,t}(h) = 0$ for all $h \neq 0$, and therefore:

$$B_U = \tilde{\Sigma}_{u,LL}(0) + \tilde{\Phi}_{LH}\tilde{\Sigma}_{u,HH}(0)\tilde{\Phi}'_{LH} + \tilde{\Phi}_{LL}\tilde{\Sigma}_{u,LH}(0)\tilde{\Phi}'_{LH} + \tilde{\Phi}_{LH}\tilde{\Sigma}_{u,HL}(0)\tilde{\Phi}'_{LL} + \tilde{\Phi}_{LL}\tilde{\Sigma}_{u,LL}(0)\tilde{\Phi}'_{LL}\ , \tag{4.6}$$

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

and $\Sigma_{U,t}(h) = 0$, for all $h \neq 0, 1$, implying $\Omega_{U,1} = 2tr\left\{\Sigma_U(0)\Sigma_U(0)' + \Sigma_U(1)\Sigma_U(1)' + \Sigma_U(-1)\Sigma_U(-1)'\right\}$. Importantly, we have that $\tilde{\Sigma}_{u,j\ell}(h) \equiv \tilde{\Sigma}_{u,t,j\ell}(h)$ for all dates t, that is these matrices do not depend on time, for $j, \ell = L, H$. Also matrices $\tilde{\Sigma}_U(h) \equiv \tilde{\Sigma}_{U,t}(h)$ and $\tilde{\Sigma}_u \equiv \tilde{\Sigma}_{u,t}(0)$ do not depend on time. The same holds for matrices $\Sigma_U(0) \equiv \Sigma_{U,t}(0)$.[11]

In Theorem 2 below, we replace matrices $\tilde{\Sigma}_U$ and Σ_U by consistent estimators. We show that the estimation error for $\frac{1}{N}tr\{\hat{B}_U\}$ in the bias adjustment is of order $o_p\left(\frac{1}{N\sqrt{T}}\right)$, implying that the asymptotic distribution of the feasible statistic is unchanged compared to infeasible one of Theorem 1.

THEOREM 2. *Let*

$$\hat{\Sigma}_u = \left(\frac{1}{N}\hat{\Lambda}'\hat{\Lambda}\right)^{-1}\left(\frac{1}{N}\hat{\Lambda}'\hat{\Gamma}\hat{\Lambda}\right)\left(\frac{1}{N}\hat{\Lambda}'\hat{\Lambda}\right)^{-1} = \begin{bmatrix} \hat{\Sigma}_{u,HH} & \hat{\Sigma}_{u,HL} \\ \hat{\Sigma}_{u,LH} & \hat{\Sigma}_{u,LL} \end{bmatrix}, \tag{4.7}$$

where $\hat{\Lambda}$ are the loadings estimators defined in equation (3.7), $\hat{\Gamma} = diag(\hat{\gamma}_{ii},\ i = 1, ..., N)$ with

[11] If the errors are weakly correlated across series and/or time, consistent estimation of $\tilde{\Sigma}_U$ and $\Omega_{U,1}$ requires thresholding of estimated cross-sectional covariances and/or HAC-type estimators. If the errors are conditionally heteroskedastic, we need consistent estimators of $\Omega_{U,2}$ and $\tilde{\Sigma}_B$ as well.

$\hat{\gamma}_{ii} = \frac{1}{T}\sum_{t=1}^{T}\hat{\varepsilon}_{i,t}^{2}$, *and* $\hat{\varepsilon}_{i,t} = y_{i,t} - \hat{\lambda}_i'\hat{f}_t$. *Let also* $\hat{\Phi}$ *be the estimator of* Φ *defined in* (3.8). *Define also:*

$$
\begin{aligned}
\hat{B}_U &= \hat{\Sigma}_{u,LL} + \hat{\Phi}_{LH}\hat{\Sigma}_{u,HH}\hat{\Phi}'_{LH} + \hat{\Phi}_{LL}\hat{\Sigma}_{u,LH}\hat{\Phi}'_{LH} + \hat{\Phi}_{LH}\hat{\Sigma}_{u,HL}\hat{\Phi}'_{LL} + \hat{\Phi}_{LL}\hat{\Sigma}_{u,LL}\hat{\Phi}'_{LL}\ , \\
\hat{\Sigma}_U(0) &= \hat{\Sigma}_{u,LL} + \hat{\Phi}_{LH}\hat{\Sigma}_{u,HH}\hat{\Phi}'_{LH} + \hat{\Phi}_{LL}\hat{\Sigma}_{u,LH}\hat{\Phi}'_{LH} + \hat{\Phi}_{LH}\hat{\Sigma}_{u,HL}\hat{\Phi}'_{LL} + \hat{\Phi}_{LL}\hat{\Sigma}_{u,LL}\hat{\Phi}'_{LL}\ , \\
\hat{\Sigma}_U(1) &= -\hat{\Phi}_{LH}\hat{\Sigma}'_{u,LH} - \hat{\Phi}_{LL}\hat{\Sigma}'_{u,LL}, \\
\hat{\Sigma}_U(-1) &= -\hat{\Sigma}_{u,LH}\hat{\Phi}'_{LH} - \hat{\Sigma}_{u,LL}\hat{\Phi}'_{LL}, \\
\hat{\Omega}_{U,1} &= 2tr\left\{\hat{\Sigma}_U(0)\hat{\Sigma}_U(0)' + \hat{\Sigma}_U(1)\hat{\Sigma}_U(1)' + \hat{\Sigma}_U(-1)\hat{\Sigma}_U(-1)'\right\},
\end{aligned}
$$

the test statistic:

$$
\tilde{\xi}(q) \ := \ N\sqrt{T}\left(\hat{\Omega}_{U,1}\right)^{-1/2}\left[\hat{\xi}(q) - \frac{1}{N}tr\left\{\hat{B}_U\right\}\right], \tag{4.8}
$$

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

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 a test of the null hypothesis at asymptotic 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* at significance level α as

$$
AR_\alpha = \left\{x \in \mathbb{R}\ :\ 0 \leq x \leq \frac{1}{N\sqrt{T}}\sqrt{\hat{\Omega}_{U,1}}z_{1-\alpha} + \frac{1}{N}\hat{B}_U\right\}, \tag{4.9}
$$

so that we cannot reject the null of q common dynamic factors as long as $\hat{\xi}(q) \in AR_\alpha$. From Theorem 2 *(ii)*, the test is consistent. To the best of our knowledge no existing test of hypothesis for the rank of matrices allows to estimate the rank of the symmetric matrix $\breve{\Sigma}_v$ when both N and T diverge.

4.3 Sequential tests for the number of common dynamic factors q

One way to estimate the number primitive shocks consists in testing sequentially the null hypothesis $H_0 = H(k)$, against the alternative $H_1 = \bigcup_{k<\ell\leq r} H(\ell)$, using the test statistic $\tilde{\xi}(k)$ defined in Theorem 2 for any generic number k of primitive shocks. 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. $\tilde{\xi}(k) \leq z_{1-\alpha}$. Otherwise, set $\hat{q}_{naive} = r$ if the test rejects the null $H(k)$ for all $k = 1, ..., r-1$. This "naive" procedure is not a consistent estimator of the number of common factors. Indeed, asymptotically a non-zero probability α of underestimating q exists coming from the type I error of the test of $H(q_0)$ 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 the true number of common factors $q > 0$ is obtained 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 (proved in OA Appendix C.2) defines a consistent inference procedure for the number primitive shocks.

PROPOSITION 2. *Let $\alpha_{N,T}$ be a sequence of real scalars defined in the interval $(0, 1)$ 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 $z_{1-\alpha_{N,T}} > 0$. Define also $z_{\alpha_{N,T}} = -z_{1-\alpha_{N,T}} < 0$. Consider the estimator $\hat{q}$ for the number primitive shocks q defined as:*

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

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 $q \in [1, r]$.

Condition *(i)* ensures 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 of each step of the procedure. The conditions in

Proposition 2 are satisfied e.g. for $\alpha_{N,T}$ such that:

$$z_{1-\alpha_{N,T}} = c(N\sqrt{T})^{\gamma}, \tag{4.10}$$

for constants $c > 0$ and $0 < \gamma < 1$. From the above definition it follows that $z_{\alpha_{N,T}} = -c(N\sqrt{T})^{\gamma}$.

The estimator for the rank q of matrix Σ_v proposed by Bai and Ng (2007) is based on (functions of) the smallest eigenvalues of the residual covariance matrix $\hat{\hat{\Sigma}}_v$. Their sequential estimation procedure is based on the rate of convergence of (functions of the sum of) the eigenvalues of $\hat{\hat{\Sigma}}_v$, but, differently from our, it is not based on a testing procedure.

5 Bootstrap

Starting from the seminal work of Gonçalves and Perron (2014), we propose a residual-based wild bootstrap implementation of the test for the number of dynamic factors q. That is, we develop a bootstrap counterpart of the test in Theorem 2. This new approach relies on PCA estimation of the pervasive factors given N_b bootstrapped panels of observations $\left\{y_1^{(b)}, \dots, y_T^{(b)}\right\}$ with $b = 1, \dots, N_b$. Differently from the recent work of Cavaliere, Gonçalves, Nielsen, and Zanelli (2023), we do not use the bootstrap to overcome the issue that the bias and/or the variances of our test statistics cannot be consistently estimated. Indeed, we derived consistent estimators of these quantities under the assumptions of Theorem 2. What we do is to 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 (4.8). 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 $T = 100$.

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 dynamic factors. In particular, this new testing procedure relies on a wild bootstrap resampling scheme and can be implemented as a three-step procedure.

- ***Step (1)***: Estimate the r static factors f_t by using the PCA estimator $\hat{f}_t$ defined in Section 3.2. Estimate a VAR(1) model on the factors $\hat{f}_t$, 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 (3.7)) 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 the 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 similar to that of Gonçalves and Perron (2014). In particular, for any $t = 1, \ldots, T$ define:

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

    where $\eta_{\varepsilon,it}$ is a zero-mean and unit-variance "external" random variable that is i.i.d. across all individuals and dates.[12] Starting from the variables in equations (5.1) and (5.2), construct the following bootstrap analogous of the DGP in equations (2.13)-(2.16) 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 $f_0^{(b)}$ which is a bootstrapped value of $\hat{f}_0$.

  - ***Step (2.b)***: As detailed in Section E.1 of the Online Appendix, use bootstrapped data $y_t^{(b)}$ 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

[12] We assume that $\eta_{\varepsilon,it} \sim iiN(0,1)$ is Gaussian in our Monte Carlo analysis and empirical applications. Experiments with other distributions such as the Rademacher distribution led to very similar results.

the variance covariance matrix of $\hat{v}_t^{H_0(q)}$:

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

which is the bootstrap counterpart of $\hat{\xi}(q)$ in equation (3.11). Then, consider the bootstrap analogous of $\hat{\Sigma}_u$ in equation (4.7):

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

where $\hat{\Gamma}^{(b)} = diag(\hat{\gamma}_{ii}^{(b)},\ i = 1, ..., N)$ with $\hat{\gamma}_{ii}^{(b)} = \frac{1}{T}\sum_{t=1}^{T} \hat{\varepsilon}_{i,t}^{(b)2}$, for $\hat{\varepsilon}_{i,t}^{(b)}$ the estimator of $\varepsilon_t^{(b)}$ based on the b-th bootstrap sample (expressions for all quantities can be found in Section E.1 of the Online Appendix). Using $\hat{\Sigma}_u^{(b)}$ to derive bootstrap equivalents of matrices $\hat{B}_U$ and $\hat{\Omega}_{U,1}$ in Theorem 2, one obtains a bootstrap-based version of the feasible test statistic in equation (4.8)

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

– ***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 dynamic factors. Using these values, one can evaluate the cumulative distribution function of $\tilde{\xi}(q)$ under the bootstrap probability measure at any $c^* \in \mathbb{R}$:

$$\hat{F}^B_{\tilde{\xi}}(c^*; q) := \frac{1}{N_b}\sum_{b=1}^{N_b} \mathbb{1}\left\{\tilde{\xi}^{(b)}(q) \le c^*\right\}, \tag{5.8}$$

where $\mathbb{1}\left\{\tilde{\xi}^{(b)}(q) \le c^*\right\} = 1$ if $\tilde{\xi}^{(b)}(q) \le c^*$, and 0 otherwise. For any $\alpha \in (0, 1)$, the α-percentile of the bootstrapped distributions of $\tilde{\xi}(q)$ is

$$\hat{p}^B_\alpha(q) \ := \ \inf\left\{p \ : \ \hat{F}^B_{\tilde{\xi}}(p; q) \ge \alpha\right\}, \tag{5.9}$$

from which we can construct the bootstrap-based *acceptance rejection* as

$$AR_\alpha^B = \left\{ x \in \mathbb{R} \,:\, 0 \leq x \leq \frac{1}{N\sqrt{T}} \sqrt{\hat{\Omega}_{U,1}} \hat{p}_{1-\alpha}^B(q) + \frac{1}{N} \hat{B}_U \right\}, \tag{5.10}$$

for $\hat{\Omega}_{U,1}$ and $\hat{B}_U$ as in Theorem 2, and which implies not rejecting the null of q common dynamic factors when $\hat{\xi}(q) \in AR_\alpha^B$.

- ***Step (3)***: Define the bootstrap-based estimator of the number of dynamic factors q as done in Proposition 2, this time replacing the (size-adjusted) critical values $z_{\alpha_{N,T}}$ with the bootstrapped percentiles $\hat{p}_\alpha^B(r)$. Hence, the bootstrap-based estimator of q is

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

which selects the right number of common dynamic factors with probability approaching $1 - \alpha$.

We remark that the bootstrap DGP in equations (5.3)-(5.4) satisfies by construction the null hypothesis of q dynamic factors, 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, being the lower $(r-q)$ block of the VAR innovations equal to zero under the bootstrap DGP for all dates t.

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 finite sample size and power of both the plug-in and bootstrap versions of the test of q dynamic factors based on $\tilde{\xi}(q)$, and iii) comparing the estimator of q based on Proposition 2 and Section 5 with some of 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 $\{\varepsilon_1, \ldots, \varepsilon_T\}$ with $\varepsilon_t = [\varepsilon_{1,t}, ..., \varepsilon_{i,t}, ...\varepsilon_{N,t}]'$ are i.i.d. draws from a Gaussian random variable with zero mean and covariance matrix $\Sigma_\epsilon = \{\beta^{|i-j|}\}_{ij}$, for $i, j = 1, ..., N$. The scalar $\beta \in [0, 1)$ induces cross-sectional dependence among the idiosyncratic innovations. We consider the case $\beta = 0$ in the main analysis, i.e. zero cross-sectional dependence, and resample these innovations in each MC simulation. For each individual i, the loadings are drawn from N independent Gaussian distributions as follows: $\breve{\lambda}_i = \lambda^* \cdot \tilde{\lambda}_i$, where $\tilde{\lambda}_i \sim i.i.N\,(\,0,\ I_r\,)$ for $i = 1, ..., N$ and $\lambda^* > 0$ controls the signal-to-noise ratio of the factors. In this analysis, we let $\lambda^* = 1$.

The r-dimensional vector $\breve{f}_t$ follows the stationary VAR(1) process:

$$\breve{f}_t \;=\; \breve{\Phi} \breve{f}_{t-1} + \breve{v}_t, \qquad \text{and} \qquad \breve{v}_t \;=\; G\eta_t,$$

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). That is, 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. We also 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. Having generated these matrices, 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.

We consider a data generating process characterised by $r = 7$ static factors and $q_0 = 5$ dynamic ones.[13] Such design is consistent with the number of static factors often documented in macroeconomic studies (see Onatski, 2010, among others). The autoregressive matrix of the VAR(1) process is given by

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

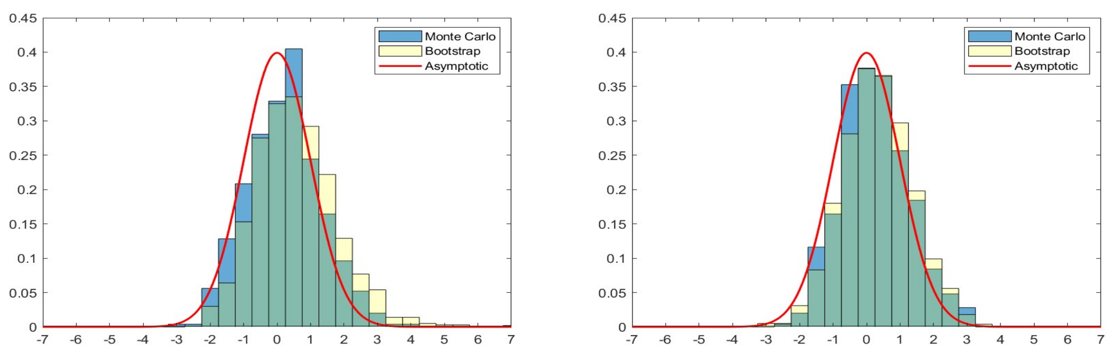
[13] The notation q_0 highlights that this is the true number of dynamic factors which we estimate with $\hat{q}$ as in Proposition 2 and Section 5.

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 data generating processes are reported in Section G.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 of $q = q_0$ dynamic factors. Histograms are based on data coming from the DGP of Section 6.1 and overlapped with the the density of the asymptotic $N(0, 1)$ distribution (red solid lines). The empirical distribution is slightly shifted to the right with respect to the asymptotic one when $(N, T) = (100, 100)$. A leftward shift makes the two much more similar when $(N, T) = (400, 600)$, become Table 13, in the OA (Section G.2), reports summary statistics of the small sample distribution of $\tilde{\xi}(q_0)$. 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. This suggests that for the smallest sample sized a test based on its percentiles will perform better than one based on the asymptotic Gaussian distribution.

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 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.

Table 2 presents the empirical size and power of the one-sided test for the null hypothesis of

Table 2 – Empirical size and power of the plug-in and of the bootstrap versions of the test of the number of dynamic factors q

		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

This table reports the empirical size and power of the one-sided test for the null hypothesis of q common dynamic factors. 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 the DGP of Section 6.1 with $r = 7$ and $q_0 = 5$. The empirical size is assessed at significance levels $\alpha \in \{0.01, 0.05, 0.1\}$. For the plug-in version of the test, the null hypothesis of q is rejected when simulated data return a value of the test statistic larger than the $(1-\alpha)$-quantile of the asymptotic distribution of $\tilde{\xi}(q)$. The rejection region for the bootstrap test is based on the same percentile of the bootstrap distribution obtained from $N_b = 499$ bootstrap iterations. For both tests, empirical 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.

q dynamic factors 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). The actual size is assessed as the empirical rejection frequency of the null hypothesis $H_0 = H(5)$ given that $q_0 = 5$, i.e. the null holds. For the plug-in test, we consider significance levels $\alpha \in \{0.01, 0.05, 0.1\}$ and reject $H(5)$ 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 $H_0 = H(4)$, i.e. either three or four dynamic factors, against the alternatives $q > 3$ and $q > 4$, respectively. Powers for other two null hypotheses ($H(1)$ and $H(2)$) always equal to one and have been omitted to save space. We report powers for a test performed at 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 macroeconomic indicators of the US economy. The bootstrap test corrects this over-rejection of the null, especially when $N < 200$. The asymptotic test has power one irrespectively of $H(q)$ and of the combination of N and T.[14] The bootstrap implementation also returns good power results. Table 10 in the Online Appendix shows actual size of the asymptotic test becomes very close to the nominal one when N and T are very large.

6.3 Estimation of the number of primitive shocks

We compare our estimators of q based on different implementations of the sequential testing procedures against alternative estimators already proposed in the literature. In particular, we focus on: $\hat{q}_3$ and $\hat{q}_4$ introduced by Bai and Ng (2007); $\hat{q}_{aw,A}$ and $\hat{q}_{aw,B}$ developed by Amengual and Watson (2007) and $\hat{q}_{bp}$ proposed by Breitung and Pigorsch (2013). As for our approach, all these estimators were developed or dynamic factor models estimated from their static factors representation, i.e. model (2.5) - (2.6). Appendix F summarizes these alternative estimators. Section G.3 of the Online Appendix shows that our estimators perform well also with respect to those of Hallin and Liska (2007). We defer this comparison to the Online Appendix because the estimator of Hallin and Liska (2007) was designed for selecting the number of dynamic factors in generalized dynamic factors models, i.e. in a setting different from the one of all estimators considered in this section.

We consider the DGP of Section 6.1 and report results both for the plug-in and for the bootstrap versions of the test. For the asymptotic test, we consider both the standard sequential procedure and the one based on the adjusted critical values as in Proposition 2. The former selects the right number of factors with probability approaching $1-\alpha$ while the latter is a consistent selection procedure. Similarly to the standard asymptotic test, also the bootstrap sequential procedure selects the right number of dynamic factors with probability approaching $1 - \alpha$. In what follows, we fix $\alpha = 0.05$. Constants c and γ in Equation (4.10) are set to 0.95 and 0.1, respectively. As before, we always consider $M = 2000$ Monte Carlo iterations.

Table 3 reports the average estimated number of factors using the five approaches. The third and fourth columns report the estimators $\hat{q}_3$ and $\hat{q}_4$ of Bai and Ng (2007). Both estimators consistently underestimate the number of common dynamic factors, though their performances improve when N and T increase. For all sample sizes, all estimators that we propose improve upon those of Bai and Ng (2007). Estimators $\hat{q}_{aw,A}$ and $\hat{q}_{aw,B}$ by Amengual and Watson (2007) underestimates q_0 when N and

[14] The same conclusion holds when we consider the size-adjusted power of the test. These further results are available upon request.

T are small but their performance improves when both sizes diverge. Note that our bootstrap-based estimator delivers better results than those of Amengual and Watson (2007) when $N < 400$ and T is at most 200 (performances of the bootstraped-based estimator and of the one based on adjusted critical values are very similar to those of $\hat{q}_{aw,A}$ and $\hat{q}_{aw,B}$ in all other cases). Results on $\hat{q}_{bp}$ by Breitung and Pigorsch (2013) point to a consistent underestimation of q_0. Notably, all our estimators improve upon $\hat{q}_{bp}$ for any sample size. Our asymptotic sequential procedure based on the quantiles of the standard Gaussian (eight column, labelled by N(0,1)) always overestimates the number of dynamic. This is a by-product of the test being oversized for the considered sample sizes. Results significantly improved when we adjust the critical value of the test to allow for a consistent selection procedure (ninth column, labelled by $z_{\alpha_{N,t}}$). The bootstrap-based estimator (tenth column, labelled by Boot) delivers the best results when $N = 100$ and its performance is rather stable across different sample sizes.

Table 3 – Comparison of estimators of q

N	T	$\hat{q}_3$	$\hat{q}_4$	$\hat{q}_{aw,A}$	$\hat{q}_{aw,B}$	$\hat{q}_{bp}$	N(0,1)	$z_{\alpha_{N,T}}$	Boot
100	100	4.48	4.49	4.86	4.88	4.23	5.17	5.12	5.01
100	200	4.49	4.50	4.92	4.93	4.39	5.30	5.20	5.03
200	100	4.45	4.46	4.90	4.92	4.34	5.07	5.04	5.06
200	200	4.63	4.63	4.94	4.95	4.55	5.12	5.06	5.05
200	300	4.64	4.64	4.96	4.96	4.63	5.17	5.06	5.05
400	100	4.44	4.45	4.92	4.93	4.40	5.04	5.01	5.07
400	200	4.62	4.62	4.96	4.97	4.63	5.06	5.02	5.06
400	300	4.70	4.70	4.97	4.98	4.71	5.08	5.02	5.06
400	600	4.74	4.74	4.99	4.99	4.79	5.12	5.03	5.05

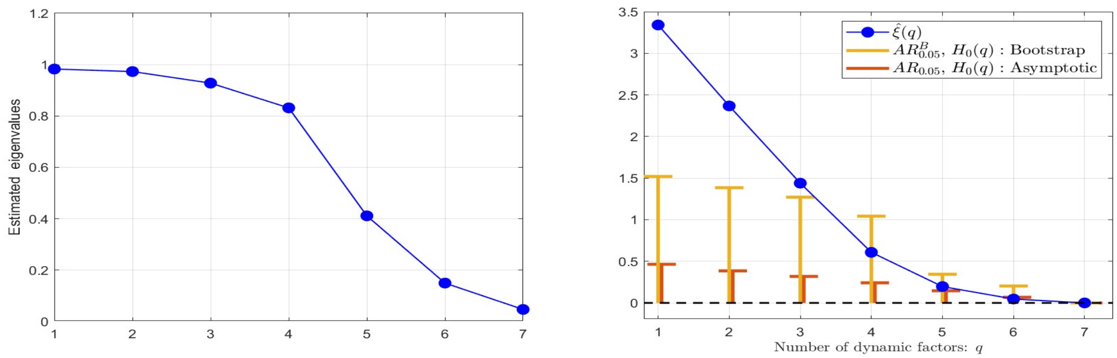
This table reports the average estimated number of dynamic factors q under the DGP of Section 6.1, i.e. $r = 7$ and $q_0 = 5$. The third and the fourth columns present results for estimators $\hat{q}_3$ and $\hat{q}_4$ introduced by Bai and Ng (2007). The fifth and sixth columns consider $\hat{q}_{aw,A}$ by $\hat{q}_{aw,B}$ 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 F of the Online Appendix. 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 column is based on the bootstrap version of the sequential testing procedure that we perform at the 5% significance level. The whole table is based on $M = 2000$ MC simulations.

7 Common factors in volatility and macro-financial panels

7.1 Common dynamic factors in US macro-financial data

We consider a macro-financial application where we test the number of common dynamic factors 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).

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



(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 the number of dynamic factors q. In the right panel, vertical bars denote the acceptance region when testing the null hypothesis of q dynamic factors, i.e. $H_0(q)$, at the 5% significance level. Yellow bars pertain to the bootstrap-based test while orange ones come from the asymptotic version of the test.

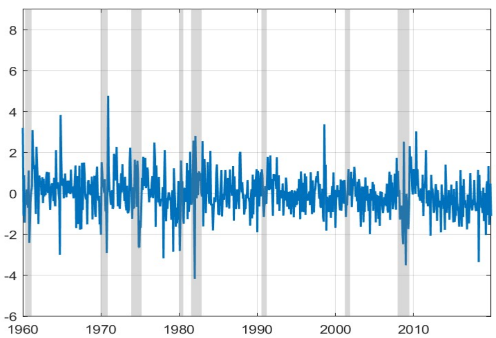
Both the information criteria IC_{p1} and IC_{p2} of Bai and Ng (2002), and their modifications by Alessi, Barigozzi, and Capasso (2010), suggest the presence of $r = 7$ static factors. We use this value as starting point for our sequential testing procedure. Following results of the Monte Carlo analysis, we use the bootstrap version of our estimator for the number of common dynamic factors. For coherence with the Monte Carlo analysis, we run the testing procedure at the 5% level of significance. Running the procedure with $N_b = 999$ bootstrap samples returns $q = 4$ dynamic factors. The same result holds when the procedure is run at the 1% level of significance and if $N_b = 499$ and $N_b = 1499$ bootstrap samples are considered. Thus, we conclude that the US macro-financial system can be fully characterised by four primitive shocks.[15] The remaining static factors are just linear combinations of past values of themselves and of the dynamic ones.

Figure 2a shows the estimated eigenvalues when $\hat{r} = 7$. Estimates for the first four eigenvalues

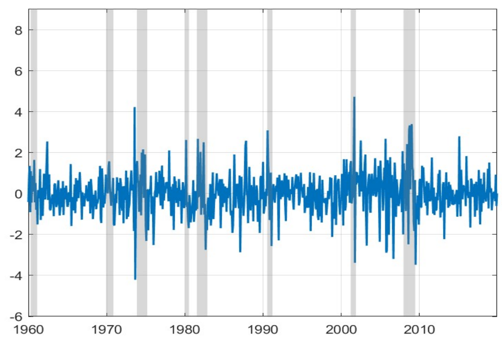
[15]Note that Bai and Ng (2007) reach same conclusions in terms of (IC-based) $\hat{r}$ and $\hat{q}$ when analysing the monthly dataset of Stock and Watson (2005).

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 being 0.41. Our test signals that we cannot reject the hypothesis that this eigenvalue (together with the 2 smaller ones) is zero at least at the 5% significance level. Remaining eigenvalues are estimated at 0.15 and 0.05. The blue solid line in Figure 2b 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 dynamic factors, i.e. $H_0(q)$, at the 5% significance level. Orange ones denote the plug-in version (see equation (4.9)) while yellow bars are for the bootstrap-based implementation of the test(see equation (5.10)). In line with results from the Monte Carlo analysis, acceptance regions for the plug-in implementation are much larger than those for the bootstrap-based version so that the former estimates a larger number of dynamic factors (five instead of four).

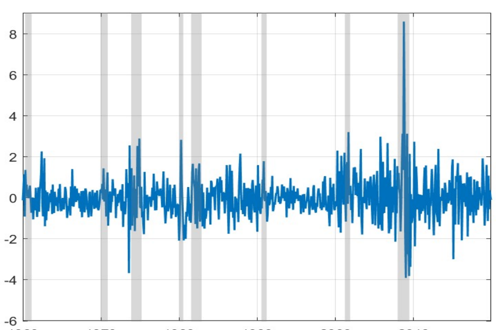
Figure 3 – Estimated dynamic factors between January 1960 and December 2019.



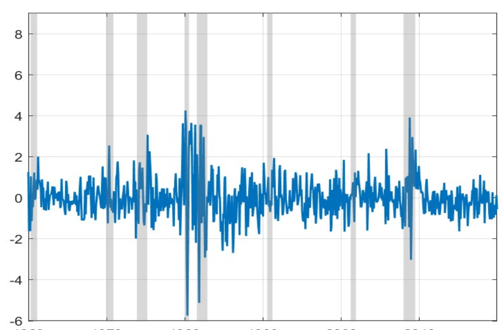
(a) First dynamic factor



(b) Second dynamic factor



(c) Third dynamic factor



(d) Fourth dynamic factor

Monthly values of the estimated dynamic factors $\hat{f}_{H,t}$ between January 1960 and December 2019. Grey shaded areas denote official NBER recession dates. Dynamic factors are estimated as in Definition 1.

Figure 3 plots the four estimated dynamic factors between January 1960 and December 2019, while

Table 4 – Macro-financial variables exhibiting the highest absolute correlation with the estimated dynamic factors from a panel of US macro-financial variables

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

This table reports the ten macro-financial variables characterised by the highest absolute correlation with each of the four estimated dynamic factors. For each variable, the value of the estimated correlation coefficient is also reported.

Table 4 reports the ten observable macro-financial variables that exhibit the highest absolute correlation with the estimated factors. This is done separately for each of the four estimated processes. The first factor is positively correlated with time series that characterise the output of the US economy.[16] Hence, this is a cyclical factor that we can view as a proxy of the state of the 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.

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

7.2 Common dynamic factors in volatility measures

We now study the common dynamic factors in a panel of volatility measures for the constituents of the S&P 500 index.[17] We 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. We set $Y_{i,t} = \log\left(\sigma_{i,t}^2\right)$ in what follows.

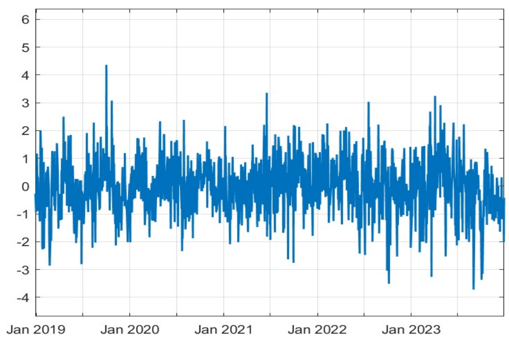
The same information criteria of Section 7.1 suggest the presence of $r = 7$ static factors, while our bootstrap-based procedure estimates $q = 4$ dynamic factors.[18] Figure 4 plots the four estimated dynamic factors for the sample of interest. To interpret these time series, table 5 reports the ten stocks which exhibit the highest absolute correlation with the estimated factors. This is done separately for each of the four estimated processes. We also report linear correlation coefficients between the stocks and the factor of interest. The first dynamic factors is significantly negatively related to providers of electricity and natural gas.[19] 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.

[17] We consider the S&P 500 composition of December 2023.

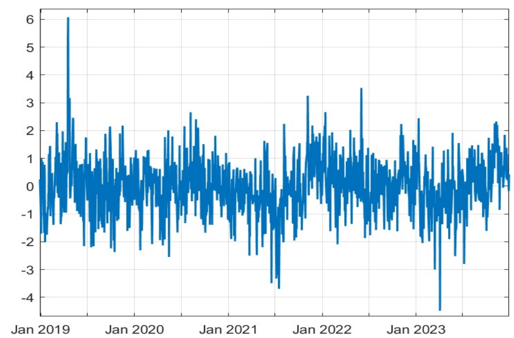
[18] As in the previous application, we consider $N_b = 999$ bootstrap samples and check for the robustness of our findings with respect to this value. The bootstrap test is run at the 5% level of significance and we check robustness with to the respect to the 1% level of significance.

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

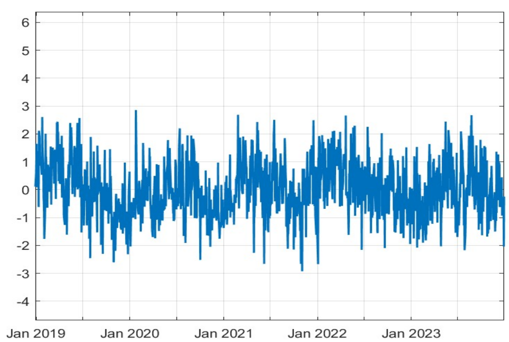
Figure 4 – Estimated dynamic factors between December 28, 2018 and December 28, 2023.



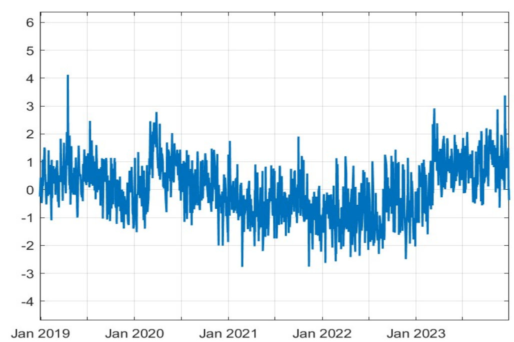
(a) First dynamic factor



(b) Second dynamic factor



(c) Third dynamic factor



(d) Fourth dynamic factor

Daily values of the estimated dynamic factors $\hat{f}_{H,t}$ between December 28, 2018 and December 28, 2023. Dynamic factors are estimated as in Definition 1.

Table 5 – Stocks exhibiting the highest absolute correlation with the estimated dynamic factors from a panel of volatility measures

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

This table reports the ten stocks that exhibit the highest absolute correlation with each of the four estimated dynamic factors. For each variable, the value of the estimated correlation coefficient is also reported.

8 Conclusions

We present new tests and estimators for the number of common dynamic factors in large panel data. The starting point of our testing procedure is the static factor representation of a dynamic factor model where r static factors evolve as a VAR(1) whose innovations have rank-deficient covariance matrix Σ_v. In particular, its rank $q \leq r$ coincides with the number of common dynamic factors in the data. Hence, we test the number of dynamic factors by testing the rank of Σ_v. In doing it, we are the first to provide a way to test for the rank of a finite dimensional positive semi-definite matrix in a panel context where both the cross-sectional and the time dimension go to infinity. Despite the well known problems with such a type of test, we manage to construct a test statistic whose distribution under the null of q dynamic factors is Gaussian. 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 dynamic factors based on both implementations. Monte Carlo results suggest that the bootstrap-based test and estimator perform well for sample sizes similar to those encountered in financial and macro-financial applications.

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.

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

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

BAI, J., AND S. NG (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.

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 (2023): "Bootstrap inference in the presence of bias," *Journal of the American Statistical Association*, (just-accepted), 1–26.

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.

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.

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

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

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

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

STOCK, J. H., AND M. W. WATSON (2002b): “Macroeconomic forecasting using diffusion indexes,” *Journal of Business and Economic Statistics*, 20, 147–162.

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 $\overset{d}{\to}$ convergence in distribution. For a sigma-field $\mathcal{F}$, we denote by $Z_n \overset{d}{\to} Z$ ($\mathcal{F}$-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 $Z_n \overset{d}{\to} N(0, \Omega)$ ($\mathcal{F}$-stably) we mean $Z_n \overset{d}{\to} \Omega^{1/2}\varepsilon$ ($\mathcal{F}$-stably), where $\varepsilon \sim N(0, I)$ is independent of $\mathcal{F}$.

A Assumptions

We make the following assumptions:

Assumption A.1. *We have $N, T \to \infty$ such that the conditions in (4.3) hold, that is:*

$$\sqrt{T}/N = o(1), \qquad N/T^{5/2} = o(1).$$

Assumption A.2. *i) The factor process $\breve{f}_t = \breve{\Phi}\breve{f}_{t-1} + \breve{v}_t$ in* (2.6) *is stationary, that is all the eigenvalues of the autoregressive matrix $\breve{\Phi}$ have modulus (strictly) smaller then one, ii) it satisfies the normalization restrictions in (2.9), that is:*

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

and the VAR innovations $\breve{v}_{H,t}$, are such that $\breve{v}_{H,t} \sim iid(0, I_q)$, with $E[\|\breve{v}_{H,t}\|^4] \leq M$, for a constant $M < \infty$. Therefore, we also have iii) the “rotated” factor process $f_t = \Phi f_{t-1} + v_t$ in (2.14) *is stationary, that is all the eigenvalues of the autoregressive matrix Φ have modulus (strictly) smaller then one, iv) it satisfies analogous normalization restrictions:*

$$E(f_t) = 0 \qquad \textit{and} \qquad V(f_t) = E(f_t f_t') = I_r,$$

and the VAR innovations $v_{H,t}$, are such that $v_{H,t} \sim iid(0, I_q)$, with $E[\|v_{H,t}\|^4] \leq M$, 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. *Moreover, 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 $E[\|f_t\|^{2r\vee 8}] \leq M$, 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:*

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

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

$$\Omega_t(h) = \begin{bmatrix} \Omega_t(0) & \Omega_t(h) \\ & \Omega_{t-h}(0) \end{bmatrix},$$

for $\Omega_t(h) = \operatorname*{plim}_{N\to\infty} \frac{1}{N} \sum_{i=1}^N \sum_{\ell=1}^N \lambda_i \lambda_\ell' cov(\varepsilon_{i,t}, \varepsilon_{\ell,t-h}|\mathcal{F}_t)$, for any k, 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. *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 $N \geq 1$.* [20] *Moreover,*
b) $\|E(\xi_t\xi_t'|\mathcal{F}_t) - E(\xi_t\xi_t'|F_t, ..., F_{t-m})\|_2 = O(m^{-\psi})$, as $m \to \infty$, uniformly in $N \geq 1$, and
c) $\|\eta_t^2 - E(\eta_t^2|\mathcal{V}_{t-m}^{t+m})\|_8 = O(m^{-\psi})$, as $m \to \infty$, uniformly in $N \geq 1$, where $\mathcal{V}_{t-m}^{t+m} = \sigma(V_s, t-m \leq s \leq t+m)$ and $\psi > 1$.

Assumption A.7. *For $j = 1, 2$: 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 $\zeta_t = (\eta_t^2 f_t', \kappa_t f_t', \xi_t', \alpha_t')'$.

Assumption A.8. *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) $P[\|\frac{1}{T} \sum_{t=1}^T z_t\| \geq \delta] \leq c_1 T \exp(-c_2 \delta^2 T^\eta) + c_3 T \delta^{-1} \exp(-c_4 T^{\bar{\eta}})$, 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 dynamic factor process (VAR) , and the first- and second-order moments of the static factor vector. All the remaining assumptions are the same as in AGGR, and we refer to their Appendix A for a detailed discussion of each of them e their relationship with analogous assumptions made in the literature.

[20] That is, $\alpha(h) = O(h^{-\phi})$ for some $\phi > \frac{r}{r-2}$, where $\alpha(h) = \sup_{N_1, N_2 \geq 1} \sup_{t\geq 1} \sup_{A\in\mathcal{V}_{-\infty}^t, B\in\mathcal{V}_{t+h}^\infty} |P(A\cap B) - P(A)P(B)|$, where $\mathcal{V}_{t-m}^{t+m} = \sigma(V_s, t-m \leq s \leq t+m)$, and similarly for V_t^*.

Assumption A.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 $\Omega_t(0)$ 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 $h \neq 0$. We note that, Assumption A.9 simplifies substantially the proof of Theorems 2 and **??**, but is not needed in the proofs of Theorem 1.

B Proofs

Section B.1 presents the proof of Proposition 1, Section B.2 presents the proof of Theorem 1, and Section B.3 presents the proof of Theorem 2. To save space, the proofs of all technical Lemmas are provided In Section C.1 in the Online Appendix.

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 $V_{12} := E(f_{t-1}f_t') = V_{21}'$. The stationarity of the factor process for $\breve{f}_t$ from Assumption A.2 i) implies that also the factor process of f_t in (2.14) is stationary, and that $V_{22} = V_{11}$, irrespectively of the normalization in (2.9). Moreover, as W is an orthonormal (r, r) eigenvector matrix, from the normalization in (2.9) 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 variance-covariance matrices $\Sigma_v = V(v_t)$ and its OLS estimator, $\tilde{\Sigma}_v = \sum_{t=1}^{T}\tilde{v}_t\tilde{v}_t'/T$, and in particular show that their smallest $r-q$ eigenvalues are all equal to zero. Then, in Section B.1.2 we show the $r-q$ smallest eigenvalues of the sample variance-covariance matrix $\breve{\tilde{\Sigma}}_v$ obtained by OLS estimation of the VAR(1) for the factors $\breve{f}_t = W_v f_t$ in (2.6) on the observed (without error) static factors are equal to 0, which proves part (i) of Proposition 1. Then, in Section B.1.3 we show that the largest q eigenvalues converge, as $T \to \infty$ to the largest non-zero q eigenvalues of $\tilde{\Sigma}_v$, which, together with the result in part (i), proves part (ii) of Proposition 1. Finally, we derive their asymptotic distribution of the largest q eigenvalues of $\breve{\tilde{\Sigma}}_v$, 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 matrices E_H and E_L span the space $\mathbb{R}^r$. Then, given the special form of $\Sigma_v = E[v_tv_t'] = diag(\sigma_1^2, ..., \sigma_q^2, 0, ..., 0)$, the eigenvectors associated with the smallest $r-q$ zero eigenvalues of Σ_v are spanned by the columns of matrix E_L. Analogously, the eigenvectors associated with the largest q non-zero eigenvalues of Σ_v are spanned by the columns of matrix E_H If we make the additional assumption that all the q largest eigenvalues of Σ_v are distinct, that is $\sigma_1^2 > \sigma_2^2 > ...... > \sigma_q^2 > 0$, then the orthonormal eigenvectors associated with the largest q non-zero eigenvalues of Σ_v are given exactly by the columns of matrix E_H.

Let us now characterize matrix $\tilde{\Sigma}_v$ and its eigenvalues, which are denoted by $\tilde{\sigma}_1, ..., \tilde{\sigma}_r$. Let us 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 Section 3.1, and the one of the OLS residuals $\tilde{v}_t = f_t - \tilde{\Phi} f_{t-1}$ obtained by estimating the VAR(1) model in equation (2.14) by OLS from the T-dimensional sample of true factors f_t, we get:

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

where the third equality follows from the definitions in (B.2) and the definition of the OLS estimator $\tilde{\Phi}$ of Φ in equation (3.2), which can be rewritten as $\tilde{\Phi} = \tilde{V}_{21}\tilde{V}_{11}^{-1}$. 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 $q \times q$ matrix $\tilde{\Sigma}_{v,HH}$ is:

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

Proof: see Online Appendix C.1.

From Lemma B.1 it follows immediately that the $r-q$ smallest eigenvalues of matrix $\tilde{\Sigma}_v$ are $\tilde{\sigma}_{q+1} = ... = \tilde{\sigma}_r = 0$. This is our first non-trivial result, which shows that the smallest $r-q$ sample eigenvalues of the covariance matrix $\tilde{\Sigma}_v$ of the the VAR(1) innovations $\tilde{v}_t$ obtained by OLS estimation of the sample of observed factors f_t (assuming they are observable without errors), are exactly equal to the smallest (null) $r-q$ eigenvalues of the population covariance matrix $\Sigma_v = V(v_t)$ of v_t, for any finite sample of dimension $T \geq r^2$. From (B.5), it immediately follows that the orthonormal eigenvectors associated to the $r-q$ zero 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 $r \times (r-q)$ orthogonal matrix, i.e. $A'A = AA' = I_q$, from which it follows: $\tilde{\Sigma}_v\tilde{W}_{v,r-q} = \tilde{W}_{v,r-q} \cdot 0_{(r-q,r-q)}$, and $\tilde{W}_{v,r-q}'\tilde{W}_{v,r-q} = I_{r-q}$.

Moreover, the $q \times q$ matrix $\tilde{\Sigma}_{v,HH}$ in (B.6) is the sample variance-covariance matrix of the residuals obtained by estimating a multivariate regression of the non-zero VAR innovation vector $v_{H,t}$ on the realizations of all the r (lagged) factors f_{t-1}.[21] This can be easily shown by noting that the estimated matrix of regression

[21] Equivalently, $\tilde{\Sigma}_{v,HH}$ is the sample variance matrix of the orthogonal projection of $v_{H,t}$ on f_{t-1}.

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 $v_t - \tilde{V}_{vf,12} \tilde{V}_{11}^{-1} f_{t-1}$, from which it immediately follows that the OLS residuals' covariance matrix is:

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

As $\tilde{\Sigma}_{v,HH}$ is a sample variance-covariance matrix, it is positive semi-definite, and therefore all its eigenvalues are non-negative. Therefore the largest q eigenvalues of matrix $\tilde{\Sigma}_v$, denoted as $\tilde{\sigma}_1^2, ..., \tilde{\sigma}_q^2$, are such that $\tilde{\sigma}_1^2 \geq ... \geq \tilde{\sigma}_q^2 \geq 0$, which implies

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

From (B.5), it immediately follows 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 $q \times q$ orthogonal matrix, i.e. $B'B = BB' = I_q$, from which it follows: $\tilde{\Sigma}_v \tilde{W}_{v,q} = \tilde{W}_{v,q} \tilde{\Sigma}_{v,HH}^{eig}$, and $\tilde{W}'_{v,q} \tilde{W}_{v,q} = I_q$, and $\tilde{\Sigma}_{v,HH}^{eig} := diag(\tilde{\sigma}_1^2, ..., \tilde{\sigma}_q^2)$

Let $\tilde{\Sigma}_v^{eig}$ be the $r \times r$ diagonal matrix collecting the the ordered eigenvalues $\tilde{\sigma}_\ell^2$, with $\ell = 1, ..., r$, of $\tilde{\Sigma}_v$:

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

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

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

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

Let us define

$$\tilde{\breve{V}}_{11} \ := \ \frac{1}{T}\sum_{t=1}^{T} \breve{f}_{t-1}\breve{f}'_{t-1} , \qquad \tilde{\breve{V}}_{22} \ := \ \frac{1}{T}\sum_{t=1}^{T} \breve{f}_t \breve{f}'_t , \qquad \tilde{\breve{V}}_{12} \ := \ \frac{1}{T}\sum_{t=1}^{T} \breve{f}_{t-1}\breve{f}'_t = \tilde{\breve{V}}'_{12} \tag{B.10}$$

From the definition of $\tilde{\breve{\Sigma}}_v$ in Section 3.1, and the one of the OLS residuals $\tilde{\breve{v}}_t = \breve{f}_t - \tilde{\breve{\Phi}} \breve{f}_{t-1}$ obtained by estimating the VAR(1) model in equation (2.14) from the T-dimensional sample of true factors $\breve{f}_t$, we get:

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

where the third equation follows from (B.10) and the expression of the OLS estimator of $\breve{\Phi}$, which can be written as $\tilde{\breve{\Phi}} = \tilde{\breve{V}}_{21} \tilde{\breve{V}}_{11}^{-1}$. 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 also implies:

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

and

$$\tilde{\breve{\Sigma}}_v \ = \ W_v \breve{V}_{22} W'_v - W_v \tilde{V}_{21} W'_v W_v \tilde{V}_{11}^{-1} W'_v W_v \tilde{V}_{12} W'_v \ = \ W_v \tilde{\Sigma}_v W'_v. \tag{B.13}$$

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 $\tilde{W}_v' W_v' W_v \tilde{W}_v = \tilde{W}_v' I_r \tilde{W}_v = I_r$, it also implies

$$\tilde{\tilde{\Sigma}}_v (W_v \tilde{W}_v) = (W_v \tilde{W}_v) \tilde{\Sigma}_v^{eig}, \tag{B.14}$$

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 concludes the proof of part (i) of Proposition 1.

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

From the assumption on the DGP of the innovations $v_{H,t}$ made in (2.14), that is $v_t \sim iid(0, \Sigma_v)$, which together with Assumption A.2 implies the stationarity of the VAR(1) process for f_t, and from Assumption A.2:

$$\sqrt{T} \cdot vec\left(\frac{1}{T} \sum_{t=1}^{T} v_{H,t} v_{H,t}' - \Sigma_{v,HH} \right) \xrightarrow{d} N(0, V_1), \qquad \text{where } V_1 := E[vec(v_t v_t' - \Sigma_v) \cdot vec(v_t v_t' - \Sigma_v)'], \tag{B.15}$$

as $T \to \infty$, or equivalently $\sum_{t=1}^{T} v_{H,t} v_{H,t}'/T - \Sigma_{v,HH} = O_p(1/\sqrt{T})$, by a standard application of the Central Limit Theorem (CLT). Under the same assumptions it is easy to show that

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

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

$$\tilde{\Sigma}_v = \begin{bmatrix} \Sigma_{v,HH} + O_p(1/\sqrt{T}) & 0_{(q,r-q)} \\ 0_{(r-q,q)} & 0_{(r-q,r-q)} \end{bmatrix}. \tag{B.16}$$

We know from the Subsection B.1.1 that the eigenspace associated with the smallest eigenvalue of $\tilde{\Sigma}_v$ (equal to 0) has dimension $r-q$ and is spanned by the columns of matrix E_L. Therefore, from (B.16), (B.8) and (B.9) we can write the following expansions for the eigenvalue matrix $\tilde{\Sigma}_{v,HH}^{eig} = diag(\tilde{\sigma}_1^2, ..., \tilde{\sigma}_q^2)$ collecting the largest q eigenvalues of $\tilde{\Sigma}_v$, and the associated eigenvector matrix $\tilde{W}_{v,q}$:

$$\tilde{W}_{v,q} = E_H \mathcal{U}_{v,q}, \qquad \tilde{\Sigma}_{v,HH}^{eig} = \Sigma_{v,HH} + M_{v,q} = diag(\tilde{\sigma}_1^2, ..., \tilde{\sigma}_q^2) + M_{v,q}, \tag{B.17}$$

where E_L is defined in equation (B.1), the stochastic $(r-q) \times (r-q)$ 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, $M_{v,q}$ converge in probability to a null matrix as $T \to \infty$ at rate $O_p(1/\sqrt{T})$, that is $M_{v,q} = 0_{(r-q),(r-q)} + O_p(1/\sqrt{T})$, which implies:

$$\tilde{\sigma}_\ell^2 = \sigma_\ell^2 + O_p(1/\sqrt{T}), \qquad \ell = 1, ..., q, \tag{B.18}$$

and $\mathcal{U}_{v,q}$ converge in probability to a nonsingular 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

$$\tilde{\sigma}_1^2 \geq ... \geq \tilde{\sigma}_q^2 > \tilde{\sigma}_{q+1}^2 = ... \tilde{\sigma}_r^2 = 0 \qquad w.p.a.\ 1, \tag{B.19}$$

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 make the assumption that all the q non-zero eigenvalues of Σ_v are distinct, namely:

$$\sigma_1^2 > ... > \sigma_q^2 > 0. \tag{B.20}$$

This assumption implies that the orthonormal eigenvectors associated with the largest q non-zero eigenvalues of Σ_v are given exactly by the 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

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

we get

$$\tilde{\Phi} = \tilde{V}_{21}\tilde{V}_{11}^{-1} = \Phi + \tilde{V}_{11}^{-1}\left[\begin{array}{c} \frac{1}{T}\sum_{t=1}^{T} v_{H,t}f_{t-1}' \\ 0_{(r-q,1)} \end{array}\right]. \tag{B.21}$$

By using Assumption (A.2), which allow the application of the CLT for the term $\frac{1}{T}\sum_{t=1}^{T} v_{H,t}f_{t-1}$, we get $\tilde{\Phi}_{HL} = \Phi_{HL} + O_p(1/\sqrt{T})$ and $\tilde{\Phi}_{HL} = \Phi_{HL} + O_p(1/\sqrt{T})$, which implies

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

Therefore, $\frac{1}{T}\sum_{t=1}^{T} \tilde{v}_{H,t}\tilde{v}_{H,t}' = \frac{1}{T}\sum_{t=1}^{T} v_{H,t}v_{H,t}' + O_p(1/T)$, which together with (B.15) implies:

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

and

$$\sqrt{T}\cdot vec\left(\frac{1}{T}\sum_{t=1}^{T} \tilde{v}_{H,t}\tilde{v}_{H,t}' - \Sigma_{v,HH}\right) \xrightarrow{d} N(0, V_1), \tag{B.22}$$

where V_1 is defined in (B.15).

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.22), 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

$$\begin{aligned}\sqrt{T}(\tilde{\sigma}_\ell^2 - \sigma_\ell^2) &= \sqrt{T}e_{q,\ell}'\left(\frac{1}{T}\sum_{t=1}^T \tilde{v}_{H,t}\tilde{v}_{H,t}' - \Sigma_{v,HH}\right)e_{q,\ell} + o_p(1)\\ &= e_{q,\ell}'(e_{q,\ell}'\otimes I_m)\sqrt{T}\cdot vec\left(\frac{1}{T}\sum_{t=1}^T \tilde{v}_{H,t}\tilde{v}_{H,t}' - \Sigma_{v,HH}\right) + o_p(1), \qquad \text{(B.23)}\end{aligned}$$

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.22) and (B.23) imply

$$\sqrt{T}(\tilde{\sigma}_\ell^2 - \sigma_\ell^2) \xrightarrow{d} N(0, V_2), \qquad \text{where} \quad V_2 := e_{q,\ell}'(e_{q,\ell}'\otimes I_q)\cdot V_1\cdot(e_{q,\ell}\otimes I_m)e_{q,\ell}. \qquad \text{(B.24)}$$

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 the estimates of the pervasive factors estimated by PCA as in AGGR (Subsection B.2.1). This result yields an asymptotic expansion for the VAR residual matrix $\hat{\Sigma}_v$ defined in equation (3.9) (Subsection B.2.2), and in turn it is used to obtain the asymptotic expansions of the eigenvalues and eigenvectors of matrix $\hat{\Sigma}_v$ by perturbation methods (Subsections B.1.1 and B.2.3). This yields the asymptotic expansions of the canonical correlations 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), similarly to AGGR.

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:*

$$\hat{\breve{f}}_t = \hat{\mathcal{H}}(\breve{f}_t + \breve{\psi}_t), \qquad \breve{\psi}_t := \frac{1}{\sqrt{N}}\breve{u}_t + \frac{1}{T}\breve{b}_t + \frac{1}{\sqrt{NT}}\breve{d}_t + \breve{\vartheta}_t, \qquad \text{(B.25)}$$

for $t = 1, \ldots, T$, *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* $N, T \to \infty$, *and the matrix* $\hat{\mathcal{H}}$ *converges in probability to a nonstochastic orthogonal* (k, k) *matrix.*

This Proposition corresponds to Proposition 3 in AGGR and the proof is analogous as the one in their paper, and therefore 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 1 follows immediately from Proposition B.1.

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

$$\hat{f}_t = \hat{\mathcal{H}} W_v (f_t + \psi_t), \qquad \psi_t := \frac{1}{\sqrt{N}} u_t + \frac{1}{T} b_t + \frac{1}{\sqrt{NT}} d_t + \vartheta_t, \tag{B.26}$$

for $t = 1, \ldots, T$, *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 the matrix* $\hat{\mathcal{H}}$ *is defined in Proposition B.1.*

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

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

$$\hat{\hat{V}}_{11} = \frac{1}{T} \sum_{t=1}^T \hat{\hat{f}}_{t-1} \hat{\hat{f}}_{t-1}' = \frac{1}{T} \sum_{t=1}^T \hat{\mathcal{H}} W_v (f_{t-1} + \psi_{t-1}) (f_{t-1} + \psi_{t-1})' W_v' \hat{\mathcal{H}}' = \hat{\mathcal{H}} W_v (\tilde{V}_{11} + \hat{X}_{11}) W_v \hat{\mathcal{H}}', \tag{B.27}$$

$$\hat{\hat{V}}_{22} = \frac{1}{T} \sum_{t=1}^T \hat{\hat{f}}_t \hat{\hat{f}}_t' = \frac{1}{T} \sum_{t=1}^T \hat{\mathcal{H}} W_v (f_t + \psi_t) (f_{t1} + \psi_t) W_v \hat{\mathcal{H}} = \hat{\mathcal{H}} W_v (\tilde{V}_{22} + \hat{X}_{22}) W_v' \hat{\mathcal{H}}', \tag{B.28}$$

$$\begin{aligned} \hat{\hat{V}}_{12} &= \frac{1}{T} \sum_{t=1}^T \hat{\hat{f}}_{t-1} \hat{\hat{f}}_t' = \frac{1}{T} \sum_{t=1}^T \hat{\mathcal{H}} W_v (f_{t-1} + \psi_{t-1}) (f_t + \psi_t)' W_v' \hat{\mathcal{H}}' = \hat{\mathcal{H}} W_v (\tilde{V}_{12} + \hat{X}_{12}) W_v' \hat{\mathcal{H}}', \\ \hat{\hat{V}}_{21} &= \hat{\hat{V}}_{12}' \end{aligned} \tag{B.29}$$

where:

$$\tilde{V}_{11} = \frac{1}{T} \sum_{t=1}^T f_{t-1} f_{t-1}' , \qquad \tilde{V}_{22} = \frac{1}{T} \sum_{t=1}^T f_t f_t' , \qquad \tilde{V}_{12} = \frac{1}{T} \sum_{t=1}^T f_{t-1} f_t' , \quad \tilde{V}_{21} = \tilde{V}_{12}' \tag{B.30}$$

and

$$\hat{X}_{11} = \frac{1}{T} \sum_{t=1}^T (f_{t-1} \psi_{t-1}' + \psi_{t-1} f_{t-1}') + \frac{1}{T} \sum_{t=1}^T \psi_{t-1} \psi_{t-1}', \tag{B.31}$$

$$\hat{X}_{22} = \frac{1}{T} \sum_{t=1}^T (f_t \psi_t' + \psi_t f_t') + \frac{1}{T} \sum_{t=1}^T \psi_t \psi_t', \tag{B.32}$$

$$\hat{X}_{12} = \frac{1}{T} \sum_{t=1}^T (f_{t-1} \psi_t' + \psi_{t-1} f_t') + \frac{1}{T} \sum_{t=1}^T \psi_{t-1} \psi_t', \qquad \hat{X}_{21} = \hat{X}_{12}'. \tag{B.33}$$

From the definition of matrix $\hat{\breve{\Sigma}}_v = \frac{1}{T}\sum_{t=1}^{T} \hat{\breve{v}}_t\hat{\breve{v}}_t'$ and $\hat{\breve{\Phi}} := \hat{\breve{V}}_{21}\hat{\breve{V}}_{11}^{-1}$ from definitions (B.27)-(B.27) we get:

$$\hat{\breve{\Sigma}}_v \;=\; \frac{1}{T}\sum_{t=1}^{T}\hat{\breve{v}}_t\hat{\breve{v}}_t' \;=\; \frac{1}{T}\sum_{t=1}^{T}(\hat{\breve{f}}_t - \hat{\breve{\Phi}}\hat{\breve{f}}_{t-1})(\hat{\breve{f}}_t - \hat{\breve{\Phi}}\hat{\breve{f}}_{t-1})' = \hat{\breve{V}}_{22} - \hat{\breve{V}}_{21}\hat{\breve{V}}_{11}^{-1}\hat{\breve{V}}_{12}. \tag{B.34}$$

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

$$\hat{\breve{\Sigma}}_v \;=\; \hat{\mathcal{H}}W_v\hat{\Sigma}_vW_v'\hat{\mathcal{H}}', \tag{B.35}$$

where:

$$\hat{\Sigma}_v \;:=\; \tilde{V}_{22} + \hat{X}_{22} - (\tilde{V}_{21} + \hat{X}_{21})\tilde{V}_{11}^{-1}\left(I_r + \hat{X}_{11}\tilde{V}_{11}^{-1}\right)^{-1}(\tilde{V}_{12} + \hat{X}_{12}). \tag{B.36}$$

By using the definition of ψ_t in Corollary 1, in the next Lemma we derive an upper bound for terms $\hat{X}_{j,k}$, $j, k = 1, 2$.

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* $\hat{\breve{\Sigma}}_v$ *is:*

$$\hat{\breve{\Sigma}}_v = \hat{\mathcal{H}}W_v(\tilde{\Sigma}_v + \hat{\Psi})W_v'\hat{\mathcal{H}}' + O_p(\delta_{N,T}^2), \tag{B.37}$$

where $\tilde{\Sigma}_v \;=\; \tilde{V}_{22} - \tilde{V}_{21}\tilde{V}_{11}^{-1}\tilde{V}_{12}$ *and* $\hat{\Psi} = \hat{\Psi}^{*\,(I)} + \hat{\Psi}^{*\,(II)}$,

$$\hat{\Psi}^{*\,(I)} \;=\; \hat{X}_{22} - \tilde{\Phi}\hat{X}_{12} - \hat{X}_{21}\tilde{\Phi}' + \tilde{\Phi}\hat{X}_{11}\tilde{\Phi}', \tag{B.38}$$

and $\tilde{\Phi} \;=\; \tilde{V}_{21}\tilde{V}_{11}^{-1}$.

Equation (B.37) represents matrix $\hat{\breve{\Sigma}}_v$ as (a function of) the sum of the sample VAR errors matrix $\tilde{\breve{\Sigma}} = W_v'\tilde{\Sigma}_uW_v$ (from (B.13)) computed with the true factor values $\breve{f}_t = W_vf_t$, the estimation error term $W_v'\hat{\Psi}W_v$ that consists of first-order and second-order components $W_v\hat{\Psi}^{*\,(I)}W_v'$ and $W_v\hat{\Psi}^{*\,(II)}W_v'$, and the third-order remainder term $O_p(\delta_{N,T}^3)$.

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

The estimators of the smallest $r - q$ zero eigenvalues of $\hat{\breve{\Sigma}}_v$ are $\hat{\sigma}_\ell^2$, for $\ell = q + 1, ..., r$, that is the $r - q$ smallest eigenvalues of matrix $\hat{\breve{\Sigma}}_v$. We now derive their asymptotic expansion under the null hypothesis $H(q)$ using perturbations arguments applied to equation (B.37). Let $\hat{W}_1^*$ be a $(r, r - q)$ matrix whose columns are eigenvectors of matrix $\hat{\breve{\Sigma}}_v$ associated with the eigenvalues $\hat{\sigma}_\ell^2$, with $\ell = q + 1, ..., r$. We have:

$$\hat{\breve{\Sigma}}_v\hat{W}_1^* \;=\; \hat{W}_1^*\hat{\Lambda}_v, \tag{B.39}$$

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 $\hat{\breve{\Sigma}}_v$. We know from Subsection B.1.1 that the eigenspace associated with the eigenvalue zero 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 eigenvalue zero of $W_v\tilde{\Sigma}_v W_v'$ has also dimension $r-q$ and is spanned by the columns of matrix $W_v E_L$. Since the columns of E_L and E_H span $\mathbb{R}^r$, from (B.37) we can write the following expansions:

$$\hat{W}_1^* = (\hat{\mathcal{H}}')^{-1} W_v [E_L\, \hat{\mathcal{U}} + E_H \hat{\alpha}], \qquad \hat{\Lambda}_v = 0_{(r-q,r-q)} + \hat{M}, \tag{B.40}$$

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 $q \times (r-q)$ 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.37) and (B.40) into the eigenvalue-eigenvector equation (B.39), using the characterization of matrix $\tilde{\Sigma}_v$ obtained in Lemma B.1, and keeping terms up to order $O_p(\delta_{N,T}^2)$, 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:*

$$\hat{\Lambda}_v = \hat{\mathcal{U}}^{-1}\hat{\Psi}_{LL}\hat{\mathcal{U}} + O_p(\delta_{N,T}^2) \qquad (OLD, WRONG!!!!!!!!!!!!!), \tag{B.41}$$

$$\hat{\Lambda}_v = \hat{\mathcal{U}}^{-1}\left[E_L' W_v' (\hat{\mathcal{H}}^{-1})(\hat{\mathcal{H}}')^{-1} W_v E_L\right]^{-1} \hat{\Psi}_{LL}\hat{\mathcal{U}} + O_p(\delta_{N,T}^2) \qquad (NEW), \tag{B.42}$$

$$\hat{W}_1^* = (\hat{\mathcal{H}}')^{-1} W_v \left(E_L - E_H \tilde{\Sigma}_{v,HH}^{-1}\hat{\Psi}_{HL}\right)\hat{\mathcal{U}} + O_p(\delta_{N,T}^2), \tag{B.43}$$

where $\hat{\Psi}_{LL}$, and $\hat{\Psi}_{LH}$ denote the lower-right $(r-q, r-q)$ block and the lower-left $(r-q, q)$ block, respectively, of matrix $\hat{\Psi}$ defined in Lemma B.3.

Note that the approximation in (B.41) 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 $\sum_{\ell=1}^{r-q} \hat{\sigma}_{r-\ell+1}^2$

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

$$\sum_{\ell=1}^{r-q} \hat{\sigma}_{r-\ell+1}^2 = tr\left\{\hat{\Psi}_{LL}\right\} + O_p(\delta_{N,T}^2), \tag{B.44}$$

by the commutative property of the trace and including third-order terms in $O_p(\delta_{N,T}^3)$. In the following Lemma B.5 we provide the asymptotic expansions of the terms within the trace operator in the r.h.s. of (B.44) by plugging the expressions of $\hat{\Psi}_{LL}$ and its components from Lemma B.3, and noting that :

$$\tilde{\Phi} = \left(\sum_{t=1}^{T} f_t f_{t-1}'\right)\left(\sum_{t=1}^{T} f_{t-1} f_{t-1}'\right)^{-1} = \tilde{V}_{21}\tilde{V}_{11}^{-1} = \begin{bmatrix} \tilde{\Phi}_{HH} & \tilde{\Phi}_{HL} \\ \tilde{\Phi}_{LH} & \tilde{\Phi}_{LL} \end{bmatrix}. \tag{B.45}$$

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

$$\begin{aligned}\sum_{\ell=1}^{r-q}\hat{\sigma}^2_{r-\ell+1} &= \frac{1}{N}tr\left\{\frac{1}{T}\sum_{t=1}^{T}E[(u_{Lt}-\tilde{\Phi}_{LH}u_{Ht-1}-\tilde{\Phi}_{LL}u_{Lt-1})(u_{Lt}-\tilde{\Phi}_{LH}u_{Ht-1}-\tilde{\Phi}_{LL}u_{Lt-1})'|\mathcal{F}_t]\right\} \\ &+\frac{1}{N\sqrt{T}}tr\Bigg\{\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\Big[(u_{Lt}-\tilde{\Phi}_{LH}u_{Ht-1}-\tilde{\Phi}_{LL}u_{Lt-1})(u_{Lt}-\tilde{\Phi}_{LH}u_{Ht-1}-\tilde{\Phi}_{LL}u_{Lt-1})' \\ &-E[(u_{Lt}-\tilde{\Phi}_{LH}u_{Ht-1}-\tilde{\Phi}_{LL}u_{Lt-1})(u_{Lt}-\tilde{\Phi}_{LH}u_{Ht-1}-\tilde{\Phi}_{LL}u_{Lt-1})'|\mathcal{F}_t]\Big]\Bigg\} \\ &+o_p\left(\epsilon_{N,T}\right),\end{aligned} \tag{B.46}$$

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

Let us define the process

$$\begin{aligned}U_t &:= u_{Lt}-\tilde{\Phi}_{LH}u_{Ht-1}-\tilde{\Phi}_{LL}u_{Lt-1} \\ &= \breve{\Phi}(r-q,q)\begin{bmatrix}u_t\\ u_{t-1}\end{bmatrix} = \breve{\Phi}(r-q,q)\begin{bmatrix}u_{Ht}\\ u_{Lt}\\ u_{Ht-1}\\ u_{Lt-1}\end{bmatrix},\end{aligned} \tag{B.47}$$

where matrix $\breve{\Phi}(r-q,q)$ is defined as:

$$\breve{\Phi}(r-q,q) := [\ 0_{(r-q,q)} \ \vdots\ I_{r-q}\ \vdots\ -\tilde{\Phi}_{LH}\ \vdots\ -\tilde{\Phi}_{LL}\ ]. \tag{B.48}$$

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

$$\sum_{\ell=1}^{r-q}\hat{\sigma}^2_{r-\ell+1}-\frac{1}{N}tr\left\{B_U\right\} = \frac{1}{N\sqrt{T}}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[U_t'U_t-E(U_t'U_t|\mathcal{F}_t)\right]\right)+o_p\left(\epsilon_{N,T}\right), \tag{B.49}$$

where:

$$B_U := \frac{1}{T}\sum_{t=1}^{T}E(U_tU_t'|\mathcal{F}_t), \tag{B.50}$$

with

$$E(U_tU_t'|\mathcal{F}_t) = E\left[(u_{Lt}-\tilde{\Phi}_{LH}u_{Ht-1}-\tilde{\Phi}_{LL}u_{Lt-1})(u_{Lt}-\tilde{\Phi}_{LH}u_{Ht-1}-\tilde{\Phi}_{LL}u_{Lt-1})'|\mathcal{F}_t\right]. \tag{B.51}$$

Under our set of assumptions, term $\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[U_t'U_t-E(U_t'U_t|\mathcal{F}_t)\right]$ is $O_p(1)$. In fact, in the next subsection we show that this term is asymptotically Gaussian distributed. The remainder term $o_p\left(\epsilon_{N,T}\right)$ in the r.h.s. of (B.49) is negligible with respect to the first term in the r.h.s.

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

From the asymptotic expansion (B.49) 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 dynamic factors. First, we apply a CLT for weakly dependent triangular array data to 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.47).

i) CLT for Near-Epoch Dependent (NED) processes

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)-c) 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) $\sup_{t\geq1,N\geq1}E\left[\|\mathcal{Z}_{N,t}\|^r\right]<\infty$*, 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* $N_1,N_2\geq 1$, [22]

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

$$\Omega_U=\sum_{h=-\infty}^{\infty}\Gamma(h),\qquad \Gamma(h):=\lim_{N\to\infty}Cov\left(\mathcal{Z}_{N,t},\mathcal{Z}_{N,t-h}\right). \tag{B.52}$$

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

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\mathcal{Z}_{N,t}\xrightarrow{d}N\left(0,\Omega_U\right), \tag{B.53}$$

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

$$\Gamma(h)=\lim_{N\to\infty}E\left[Cov\left(\mathcal{Z}_{N,t},\mathcal{Z}_{N,t-h}|\mathcal{F}_t\right)\right]. \tag{B.54}$$

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

$$\begin{pmatrix}U_t\\U_{t-h}\end{pmatrix}\xrightarrow{d}\begin{pmatrix}U_t^{\infty}\\U_{t-h}^{\infty}\end{pmatrix}\sim N\left(0_{(2r,1)},\begin{bmatrix}\Sigma_{U,t}(0) & \Sigma_{U,t}(h)\\ \Sigma_{U,t}(h)' & \Sigma_{U,t}(0)\end{bmatrix}\right),\quad (\mathcal{F}_t\text{-stably}), \tag{B.55}$$

[22] 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 $t\geq1$ and $N\geq1$, where $\xi(m)=O(m^{-\psi})$ for some $\psi>1$.

where

$$\begin{aligned}
\Sigma_{U,t}(h) &= Cov(U_t^\infty, U_{t-h}^\infty|\mathcal{F}_t) = E[U_t^\infty U_{t-h}^{\infty\prime}|\mathcal{F}_t] \\
&= \breve{\Phi}(r-q,q) \cdot \left(\underset{N\to\infty}{\text{plim}} \begin{bmatrix} E[u_t u_t|\mathcal{F}_t] & E[u_t u_{t-1}|\mathcal{F}_t] \\ E[u_{t-1} u_t|\mathcal{F}_t] & E[u_{t-1} u_{t-1}|\mathcal{F}_t] \end{bmatrix} \right) \cdot \breve{\Phi}(r-q,q)' \\
&= \breve{\Phi}(r-q,q) \cdot \begin{bmatrix} \tilde{\Sigma}_{u,t}(h) & \tilde{\Sigma}_{u,t}(h-1) \\ \tilde{\Sigma}_{u,t}(h-1)' & \tilde{\Sigma}_{u,t}(h) \end{bmatrix} \cdot \breve{\Phi}(r-q,q)' \\
&= [\ 0_{(r-q,q)} \ \vdots\ I_{r-q} \ \vdots\ -\tilde{\Phi}_{LH} \ \vdots\ -\tilde{\Phi}_{LL}\ ] \times \\
&\quad \times \begin{bmatrix} \Sigma_{u,t,HH}(h) & \Sigma_{u,t,HL}(h) & \Sigma_{u,t,HH}(h-1) & \Sigma_{u,t,HL}(h-1) \\ \Sigma_{u,t,LH}(h) & \Sigma_{u,t,LL}(h) & \Sigma_{u,t,LH}(h-1) & \Sigma_{u,t,LL}(h-1) \\ \Sigma_{u,t,HH}(h-1)' & \Sigma_{u,t,LH}(h-1)' & \Sigma_{u,t,HH}(h) & \Sigma_{u,t,HL}(h) \\ \Sigma_{u,t,HL}(h-1)' & \Sigma_{u,t,LL}(h-1)' & \Sigma_{u,t,LH}(h) & \Sigma_{u,t,LL}(h) \end{bmatrix} \cdot \begin{bmatrix} 0_{(r-q,q)} \\ I_{r-q} \\ -\tilde{\Phi}'_{LH} \\ -\tilde{\Phi}'_{LL} \end{bmatrix} \\
&= \Sigma_{u,t,LL}(h) - \tilde{\Phi}_{LH}\Sigma_{u,t,LH}(h-1)' - \tilde{\Phi}_{LL}\Sigma_{u,t,LL}(h-1)' \\
&\quad -\Sigma_{u,t,LH}(h-1)\tilde{\Phi}'_{LH} + \tilde{\Phi}_{LH}\Sigma_{u,t,HH}(h)\tilde{\Phi}'_{LH} + \tilde{\Phi}_{LL}\Sigma_{u,t,LH}(h)\tilde{\Phi}'_{LH} \\
&\quad -\Sigma_{u,t,LL}(h-1)\tilde{\Phi}'_{LL} + \tilde{\Phi}_{LH}\Sigma_{u,t,HL}(h)\tilde{\Phi}'_{LL} + \tilde{\Phi}_{LL}\Sigma_{u,t,LL}(h)\tilde{\Phi}'_{LL} \ , \qquad \text{(B.56)}
\end{aligned}$$

and $\Sigma_{U,t}(0) = V(U_t^\infty|\mathcal{F}_t) = E[U_t^\infty U_t^{\infty\prime}|\mathcal{F}_t]$. Using analogous arguments,we can also compute explicitly B_U:

$$\begin{aligned}
B_U &= \frac{1}{T}\sum_{t=1}^{T} E(U_t U_t'|\mathcal{F}_t) = \frac{1}{T}\sum_{t=1}^{T} \tilde{\Sigma}_{U,t}(0) \\
&= \frac{1}{T}\sum_{t=1}^{T} \Big\{ \tilde{\Sigma}_{u,t,LL}(0) - \tilde{\Phi}_{LH}\tilde{\Sigma}_{u,t,LH}(-1)' - \tilde{\Phi}_{LL}\tilde{\Sigma}_{u,t,LL}(-1)' \\
&\quad -\tilde{\Sigma}_{u,t,LH}(-1)\tilde{\Phi}'_{LH} + \tilde{\Phi}_{LH}\tilde{\Sigma}_{u,t,HH}(0)\tilde{\Phi}'_{LH} + \tilde{\Phi}_{LL}\tilde{\Sigma}_{u,t,LH}(0)\tilde{\Phi}'_{LH} \\
&\quad -\tilde{\Sigma}_{u,t,LL}(-1)\tilde{\Phi}'_{LL} + \tilde{\Phi}_{LH}\tilde{\Sigma}_{u,t,HL}(0)\tilde{\Phi}'_{LL} + \tilde{\Phi}_{LL}\tilde{\Sigma}_{u,t,LL}(0)\tilde{\Phi}'_{LL} \Big\} \ . \qquad \text{(B.57)}
\end{aligned}$$

We use the Lebesgue Lemma to interchange the limes for $N \to \infty$ and the outer expectation in the r.h.s. of (B.54), and the fact that convergence in distribution plus uniform integrability imply convergence of the expectation for a sequence of random variables (see Theorem 25.12 in Billingsley (1995)) to show the next lemma.

LEMMA B.7. *Under Assumptions A.3 and A.5 b), we have:*

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

Lemma B.7 allows to deploy the joint asymptotic Gaussian distribution of $(U_t^{\infty\prime}, U_{t-h}^{\infty\prime})'$ to compute the limit autocovariance $\Gamma(h)$. To compute the upper-left block of matrix $\Gamma(h)$, we use Theorem 12 p. 284 in Magnus and Neudecker (2007) and Theorem 10.21 in Schott (2005) which provide the covariance between two quadratic forms of Gaussian vectors. We get $Cov(U_t^{\infty\prime} U_t^\infty, U_{t-h}^{\infty\prime} U_{t-h}^\infty|\mathcal{F}_t) = 2tr\left\{\Sigma_{U,t}(h)\Sigma_{U,t}(h)'\right\}$. Therefore, from (B.52) and Lemma B.7 we get:

$$\Omega_U = \Omega_{U,1} = \sum_{h=-\infty}^{\infty} 2tr\left\{E\left[\Sigma_{U,t}(h)\Sigma_{U,t}(h)'\right]\right\} \ . \qquad \text{(B.58)}$$

ii) Asymptotic Gaussian distribution of the test statistic

Let us define vector $D_{N,T} = \frac{1}{N\sqrt{T}}$. From equations (B.49) and (B.58), and by using:

$$(D_{N,T}\Omega_{U,1}D_{N,T})^{1/2} = \left(\frac{1}{(N\sqrt{T})^2}\Omega_{U,1}\right)^{1/2} = \frac{1}{N\sqrt{T}}\Omega_{U,1}^{1/2},$$

and $N\sqrt{T}\Omega_{U,1}^{-1/2} = O\left(N\sqrt{T}\right) = O(\epsilon_{N,T}^{-1})$, under the hypothesis of q common factors in each group the statistics $\hat{\xi}(q) = \sum_{\ell=1}^{r-q}\hat{\sigma}_{r-\ell+1}$ is such that:

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

From equation (B.53), the r.h.s. converges in distribution to a standard normal distribution, which yields Theorem 1.

B.3 Proof of Theorem 2

[TO BE COMPLETED]

ONLINE APPENDIX

"New Tests and Estimators for Common Dynamic Factors".

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

This version: February 28, 2024

[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).

C Proofs of Lemmas

C.1 Proof of Lemma B.1

By substituting the expression for f_t provided in (2.14) into the definitions of $\tilde{V}_{22}$ and $\tilde{V}_{12}$ in (B.2) we get:

$$\begin{aligned}
\tilde{V}_{22} &= \frac{1}{T}\sum_{t=1}^{T} f_t f_t' = \frac{1}{T}\sum_{t=1}^{T}(\Phi f_{t-1} + v_t)(\Phi f_{t-1} + v_t)' \\
&= \Phi\tilde{V}_{11}\Phi' + \Phi\left(\frac{1}{T}\sum_{t=1}^{T} f_{t-1}v_t'\right) + \left(\frac{1}{T}\sum_{t=1}^{T} v_t f_{t-1}'\right)\Phi' + \frac{1}{T}\sum_{t=1}^{T} v_t v_t', \qquad \text{(C.1)} \\
\tilde{V}_{12} &= \frac{1}{T}\sum_{t=1}^{T} f_{t-1} f_t' = \frac{1}{T}\sum_{t=1}^{T} f_{t-1}(\Phi f_{t-1} + v_t)' = \tilde{V}_{11}\Phi' + \frac{1}{T}\sum_{t=1}^{T} f_{t-1}v_t'. \qquad \text{(C.2)}
\end{aligned}$$

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:

$$\begin{aligned}
\tilde{\Sigma}_v &= \tilde{V}_{22} - \tilde{V}_{21}\tilde{V}_{11}^{-1}\tilde{V}_{12} \\
&= \Phi\tilde{V}_{11}\Phi' + \Phi\left(\frac{1}{T}\sum_{t=1}^{T} f_{t-1}v_t'\right) + \left(\frac{1}{T}\sum_{t=1}^{T} v_t f_{t-1}'\right)\Phi' + \frac{1}{T}\sum_{t=1}^{T} v_t v_t' \\
&\quad - \left(\Phi\tilde{V}_{11} + \frac{1}{T}\sum_{t=1}^{T} v_t f_{t-1}'\right)\tilde{V}_{11}^{-1}\left(\tilde{V}_{11}\Phi' + \frac{1}{T}\sum_{t=1}^{T} f_{t-1}v_t'\right) \\
&= \frac{1}{T}\sum_{t=1}^{T} v_t v_t' - \left(\frac{1}{T}\sum_{t=1}^{T} v_t f_{t-1}'\right)\tilde{V}_{11}^{-1}\left(\frac{1}{T}\sum_{t=1}^{T} f_{t-1}v_t'\right). \qquad \text{(C.3)}
\end{aligned}$$

By by substituting the definition $v_t = [v_{Ht}'\ ,\ 0_{q\times 1}]'$ from (2.15) into (C.3) it follows that

$$\begin{aligned}
\tilde{\Sigma}_v &= \begin{bmatrix} \frac{1}{T}\sum_{t=1}^{T} v_{H,t}v_{H,t}' & 0_{q\times(r-q)} \\ 0_{(r-q)\times q} & 0_{(r-q)\times(r-q)} \end{bmatrix} - \begin{bmatrix} \frac{1}{T}\sum_{t=1}^{T} v_{H,t}f_{t-1}' \\ 0_{(r-q)\times r} \end{bmatrix}\tilde{V}_{11}^{-1}\begin{bmatrix} \frac{1}{T}\sum_{t=1}^{T} f_{t-1}v_{H,t}' & 0_{r\times(r-q)} \end{bmatrix} \\
&= \begin{bmatrix} \frac{1}{T}\sum_{t=1}^{T} v_{H,t}v_{H,t}' & 0_{q\times(r-q)} \\ 0_{(r-q)\times q} & 0_{(r-q)\times(r-q)} \end{bmatrix} - \begin{bmatrix} \frac{1}{T}\sum_{t=1}^{T} v_{H,t}f_{t-1}'\tilde{V}_{11}^{-1}\left(\frac{1}{T}\sum_{t=1}^{T} f_{t-1}v_{H,t}'\right) & 0_{q\times(r-q)} \\ 0_{(r-q)\times q} & 0_{(r-q)\times(r-q)} \end{bmatrix}, \qquad \text{(C.4)}
\end{aligned}$$

which concludes the proof.

C.2 Proof of Lemma B.2

[TO BE COMPLETED]

C.3 Proof of Lemma B.3

[TO BE COMPLETED]

C.4 Proof of Lemma B.4

[TO BE COMPLETED]

C.5 Proof of Lemma B.5

[TO BE COMPLETED]

C.6 Proof of Lemma B.6

[TO BE COMPLETED]

C.7 Proof of Lemma B.7

[TO BE COMPLETED]

D Alternative identification and testing of dynamic factors based on canonical correlation analysis

E 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).

E.1 Test based on the smallest eigenvalues

Let $\hat{\hat{F}}^{(b)} = [\hat{\hat{f}}_0^{(b)}, \hat{\hat{f}}_1^{(b)}, ..., \hat{\hat{f}}_T^{(b)}]'$ 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, $\hat{\hat{F}}^{(b)}$ satisfies the usual PCA eigenvalue-eigenvector equation:

$$\frac{1}{N(T+1)} Y^{(b)} Y^{(b)\prime} \hat{\hat{F}}^{(b)} = \hat{\hat{F}}^{(b)} \hat{V}^{(b)},$$

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 $\frac{1}{T+1}\hat{\breve{F}}^{(b)\prime}\hat{\breve{F}}^{(b)\prime} = \frac{1}{T+1}\sum_{t=0}^{T}\hat{\breve{f}}_t^{(b)}\hat{\breve{f}}_t^{(b)\prime} = I_r$. [4]

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 $\hat{F}^{(b)} := [\hat{f}_1^{(b)}, ..., \hat{f}_T^{(b)}]' = \hat{W}_v^{(b)\prime}\hat{\breve{F}}^{(b)}$. The (N,r) matrix of estimated loadings $\hat{\Lambda}^{(b)} = [\hat{\lambda}_1^{(b)}, ..., \hat{\lambda}_N^{(b)}]'$ is computed as:

$$\hat{\Lambda}^{(b)} \ = \ Y^{(b)\prime}\hat{F}^{(b)}(\hat{F}^{(b)\prime}\hat{F}^{(b)})^{-1}$$

Let $\hat{\varepsilon}_t^{(b)} = y_t^{(b)} - \hat{\Lambda}_j^{(b)}\hat{f}_t^{(b)}$, $\hat{\Xi}^{(b)} := [\hat{\varepsilon}_0^{(b)}, \hat{\varepsilon}_1^{(b)}, ..., \hat{\varepsilon}_T^{(b)}]' = Y^{(b)} - \hat{F}^{(b)}\hat{\Lambda}^{(b)}$, and 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 $\Sigma_v^{(b)}$:

$$\hat{\Sigma}_v^{(b)} = \frac{1}{T}\sum_{t=1}^{T}\hat{v}_t^{(b)}\hat{v}_t^{(b)\prime}, \tag{E.1}$$

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 $\hat{\sigma}_{q+1}^{2(b)}, \ldots, \hat{\sigma}_r^{2(b)}$ is what we use when constructing $\hat{\xi}^{(b)}(q)$ in equation (5.5).

[4] 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 $\hat{\breve{F}}^{(b)} = \sqrt{T+1}\cdot\hat{F}^{(b)*}$.

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.2 Test based on the largest canonical correlations

F Estimators of q proposed in the literature

F.1 Estimators of q 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:

$$\hat{q}_{bn,3} = \min(k \in K_3), \qquad \hat{q}_{bn,4} = \min(k \in K_4) \;. \tag{G.1}$$

Bai and Ng (2007) set either $m_3 = m_4 = 1$, or $m_3 = 1.25$ and $m_4 = 2.25$. The former combination is to be preferred when working with covariance matrices while the latter shall be adopted when the focus is on correlation matrices.

F.2 Estimators of q 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:

$$
\begin{aligned}
\hat{Z}_t^A &:= Y_t - \sum_{i=1}^{p} \hat{\breve{\Lambda}} \hat{\breve{\Phi}}_i \hat{\breve{F}}_{t-i}, \\
\hat{Z}_t^B &:= Y_t - \sum_{i=1}^{p} \hat{\breve{\Pi}} \hat{\breve{F}}_{t-i},
\end{aligned}
$$

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

$$
\begin{aligned}
\hat{q}_{aw,A} &= \operatorname*{arg\,min}_{0 \le k \le r} \left\{ \ln[\hat{\sigma}^2_{\hat{Z}^A} - R(k, \hat{Z}^A)] + k \times \frac{ln[s_{NT}] \cdot (N+T)}{NT} \right\}, && \text{(G.2)} \\
\hat{q}_{aw,B} &= \operatorname*{arg\,min}_{0 \le k \le r} \left\{ \ln[\hat{\sigma}^2_{\hat{Z}^B} - R(k, \hat{Z}^B)] + k \times \frac{ln[s_{NT}] \cdot (N+T)}{NT} \right\}, && \text{(G.3)}
\end{aligned}
$$

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

$$
R(k, \hat{Z}^A) := \sum_{\ell=1}^{k} \omega_\ell^A, \qquad \text{(G.4)}
$$

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 the quantities based on panel $\hat{Z}^B$. Because our MC analysis is comparable to ours, we follow Amengual and Watson (2007) and set $p = 2$ in Section 6. Finally, note that $\hat{q}_{aw,A}$ and $\hat{q}_{aw,B}$ correspond to the "IPC2" estimator of the number of r static factors of Bai and Ng (2002) applied to the panels $\hat{Z}^A$ and $\hat{Z}^B$.

F.3 Estimator of q 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:

$$
\tilde{S}_{00} := \sum_{t=m+1}^{T} \hat{\breve{F}}_t \hat{\breve{F}}'_t \qquad \tilde{S}_{01} := \sum_{t=m+1}^{T} \hat{\breve{F}}_t \hat{G}'_{t-1}, \qquad \tilde{S}_{11} := \sum_{t=m+1}^{T} \hat{G}_{t-1} \hat{G}'_{t-1}. \qquad \text{(G.5)}
$$

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 following quantity:

$$\hat{\xi}_{bp}(k) = \tilde{C}_{NT}^{2-\delta} \cdot \sum_{\ell=1}^{r-k}(1-\hat{\rho}_{bp,\ell}), \qquad k=1,...,r-1$$
$$\hat{\xi}_{bp}(r) = 0\,,$$

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:

$$\hat{q}_{bp} = \min\Big(k \ : \ \hat{\xi}^{*}_{bp}(k) < 4.5\Big)\,, \tag{G.6}$$

where:

$$\hat{\xi}^{*}_{bp}(k) = \tilde{C}_{NT}^{4/3} \cdot \sum_{\ell=1}^{r-k}(1-\hat{\rho}_{bp,\ell}), \qquad k=1,...,r-1 \tag{G.7}$$
$$\hat{\xi}_{bp}(r) = 0\,. \tag{G.8}$$

G Monte Carlo: additional results

G.1 Alternative data generating processes

This section repeats the Monte Carlo analysis of Section 6 but using alternative data generating processes. In particular, we work under the same setting of Section 6.1 but consider different values of (r, q_0) and of the auto-regressive matrices. The first alternative DGP, that we call Design 1, is based on $r = 5$ static factors, $q_0 = 3$ dynamic ones and autoregressive matrix

$$\Phi = diag(0.2, 0.375, 0.55, 0.725, 0.9).$$

The second DGP of this section relies on $r = 9$ static factors, $q_0 = 8$ dynamic ones and autoregressive matrix

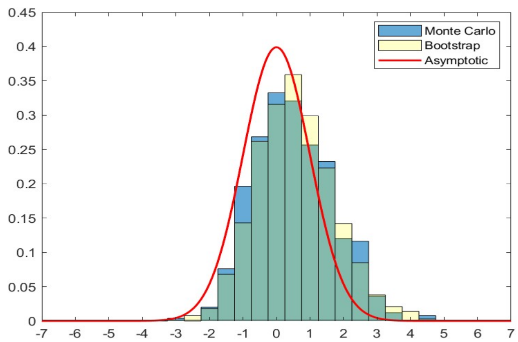
$$\Phi = diag(0.2, 0.2875, 0.375, 0.4625, 0.55, 0.6375, 0.725, 0.8125, 0.9).$$

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.

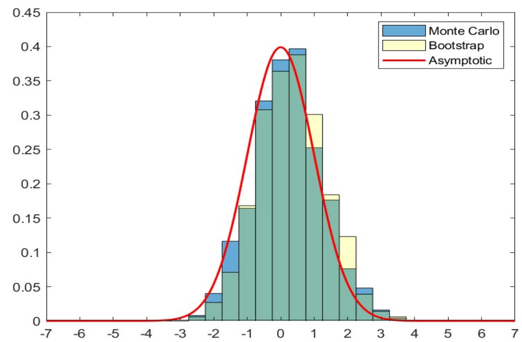
For both DGPs, the empirical distribution is poorly approximated by the asymptotic one when $(N, T) = (100, 100)$. The approximation improves substantially when we move to $(N, T) = (1000, 600)$.

Blue histograms in Figure 1 display the empirical distribution of $\tilde{\xi}(q)$ under the null hypothesis of $q = q_0$ dynamic factors. 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 consider. 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 G.2.

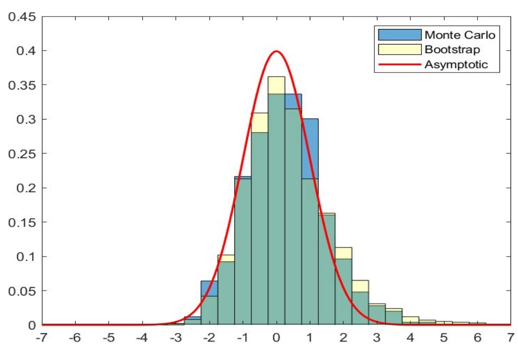
Figure 5 – Small sample and bootstrapped distribution of the test statistic $\tilde{\xi}(q_0)$.



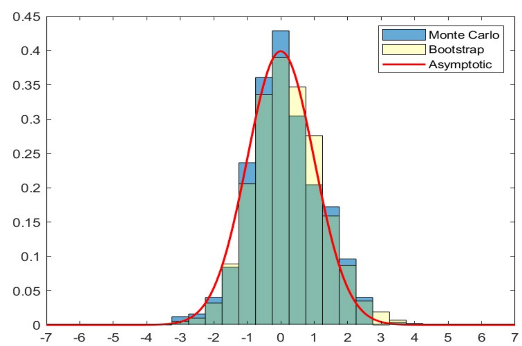
(a) $\tilde{\xi}(3)$, $N = 100$, and $T = 100$; Design 1



(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 the empirical size and power of the two testing procedures. As far as power is concerned, we test the null hypotheses $H_0 = H(2)$ and $H_0 = H(3)$ for Design 1, and $H_0 = H(6)$ and $H_0 = H(7)$. In all cases, the alternative hypothesis is given by $q > k$ for k the number of factors that is being tested. For both tables, the left panel pertains to the asymptotic test while the right one studies the Bootstrap test. The DGP notwithstanding, the asymptotic tests is always oversized and has unit power.[5] Adopting a bootstrap procedure always improves the size of the test at the cost of some loss of power under Design 1.

[5]The asymptotic tests consistently returns unit power also when controlling for the size distortion.

Table 6 – Empirical size and power of the plug-in and of the bootstrap versions of the test of the number of dynamic factors q; 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

This table reports the empirical size and power of the one-sided test for the null hypothesis of q common dynamic factors. 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 so that $r = 5$ and $q_0 = 3$. The empirical size is assessed at significance levels $\alpha \in \{0.01, 0.05, 0.1\}$. For the plug-in version of the test, the null hypothesis of q is rejected when simulated data return a value of the test statistic larger than the $(1-\alpha)$-quantile of the asymptotic distribution of $\tilde{\xi}(q)$. The rejection region for the bootstrap test is based on the same percentile of the bootstrap distribution obtained from $N_b = 499$ bootstrap iterations. For both tests, empirical 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 dynamic factors q; 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

This table reports the empirical size and power of the one-sided test for the null hypothesis of q common dynamic factors. 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 $q_0 = 8$. The empirical size is assessed at significance levels $\alpha \in \{0.01, 0.05, 0.1\}$. For the plug-in version of the test, the null hypothesis of q is rejected when simulated data return a value of the test statistic larger than the $(1-\alpha)$-quantile of the asymptotic distribution of $\tilde{\xi}(q)$. The rejection region for the bootstrap test is based on the same percentile of the bootstrap distribution obtained from $N_b = 499$ bootstrap iterations. For both tests, empirical 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 rpeort results for the asymptotic test. Actual sizes are now very close to nominal ones while powers are unaltered.

Table 8 – Empirical size and power of the plug-in version of the feasible test of the number of dynamic factors q 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

This table reports the empirical size and power of the one-sided test for the null hypothesis of q common dynamic factors. 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 $q_0 = 3$. The empirical size is assessed at significance levels $\alpha \in \{0.01, 0.05, 0.1\}$. The null hypothesis of q is rejected when simulated data return a value of the test statistic larger than the $(1-\alpha)$-quantile of the asymptotic distribution of $\tilde{\xi}(q)$. Empirical 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 size and power of the plug-in version of the feasible test of the number of dynamic factors 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

This table reports the empirical size and power of the one-sided test for the null hypothesis of q common dynamic factors. 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 $q_0 = 8$. The empirical size is assessed at significance levels $\alpha \in \{0.01, 0.05, 0.1\}$. The null hypothesis of q is rejected when simulated data return a value of the test statistic larger than the $(1-\alpha)$-quantile of the asymptotic distribution of $\tilde{\xi}(q)$. Empirical 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 size and power of the plug-in version of the feasible test of the number of dynamic factors q when N and T are large

		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

This table reports the empirical size and power of the one-sided test for the null hypothesis of q common dynamic factors. 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 $q_0 = 5$. The empirical size is assessed at significance levels $\alpha \in \{0.01, 0.05, 0.1\}$. The null hypothesis of q is rejected when simulated data return a value of the test statistic larger than the $(1-\alpha)$-quantile of the asymptotic distribution of $\tilde{\xi}(q)$. Empirical 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 tends to underestimate q_0 under Design 2. The estimators of Bai and Ng (2007) and Breitung and Pigorsch (2013) perform on par with our bootstrap based estimator under Design 1 but are substantially outperformed when we consider a richer factor structure, i.e. for Design 2. In this second case, both the bootstrap based procedure and the (adjusted) asymptotic one always outperform all the other estimators.

Table 11 – Comparison of estimators of q under Design 1

N	T	$\hat{q}_3$	$\hat{q}_4$	$\hat{q}_{aw,A}$	$\hat{q}_{aw,B}$	$\hat{q}_{bp}$	N(0,1)	$z_{\alpha_{N,T}}$	Boot
100	100	2.94	2.95	2.98	2.99	2.82	3.19	3.13	2.93
100	200	2.94	2.95	2.99	2.99	2.88	3.27	3.19	2.93
200	100	2.94	2.94	2.99	2.99	2.86	3.09	3.05	2.97
200	200	2.96	2.96	2.99	2.99	2.93	3.13	3.06	2.98
200	300	2.96	2.96	2.99	3.00	2.94	3.16	3.06	2.98
400	100	2.94	2.94	2.99	2.99	2.88	3.07	3.03	3.02
400	200	2.96	2.96	3.00	3.00	2.94	3.07	3.02	3.02
400	300	2.97	2.97	3.00	3.00	2.96	3.09	3.02	3.01
400	600	2.98	2.98	3.00	3.00	2.97	3.12	3.03	3.02

This table reports the average estimated number of dynamic factors 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 $\hat{q}_4$ introduced by Bai and Ng (2007). The fifth and sixth columns consider $\hat{q}_{aw,A}$ by $\hat{q}_{aw,B}$ 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 F of the Online Appendix. 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 column is based on the bootstrap version of the sequential testing procedure that we perform at the 5% significance level. The whole table is based on $M = 2000$ MC simulations.

Table 12 – Comparison of estimators of q under Design 2

N	T	$\hat{q}_3$	$\hat{q}_4$	$\hat{q}_{aw,A}$	$\hat{q}_{aw,B}$	$\hat{q}_{bp}$	N(0,1)	$z_{\alpha_{N,T}}$	Boot
100	100	6.44	6.58	7.50	7.58	6.53	8.13	8.09	8.05
100	200	6.51	6.63	7.71	7.74	6.77	8.25	8.18	8.05
200	100	6.42	6.54	7.62	7.70	6.68	8.06	8.03	8.07
200	200	6.84	6.90	7.77	7.79	6.95	8.10	8.04	8.07
200	300	6.85	6.90	7.84	7.84	7.06	8.11	8.05	8.06
400	100	6.41	6.52	7.69	7.75	6.76	8.03	8.01	8.07
400	200	6.82	6.88	7.84	7.85	7.07	8.06	8.02	8.06
400	300	7.00	7.04	7.88	7.88	7.20	8.06	8.02	8.06
400	600	7.12	7.14	7.92	7.92	7.37	8.09	8.02	8.06

This table reports the average estimated number of dynamic factors 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 $\hat{q}_4$ introduced by Bai and Ng (2007). The fifth and sixth columns consider $\hat{q}_{aw,A}$ by $\hat{q}_{aw,B}$ 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 F of the Online Appendix. 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 column is based on the bootstrap version of the sequential testing procedure that we perform at the 5% significance level. The whole table is based on $M = 2000$ MC simulations.

G.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 G.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 G.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 G.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.

G.3 Comparison with the estimators of HL

In this section, we compare our estimators for the number of common dynamic factors with those of Hallin and Liska (2007). Because the approach of Hallin and Liska (2007) is based on frequency domain analysis and is developed within the context of generalized dynamic factor models, we dot no present it in details as done for the estimators in Section F.

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).[6] 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, i.e. five static factors and three dynamic ones. The first six columns contain results for the estimators of Hallin and Liska (2007), where HL_{11} labels the one based on information criterion IC_1 and penalty term p_1; other columns are similarly labelled. The bootstrap-based estimator delivers 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

[6] We are grateful to Alexey Onatskiy for sharing codes of his paper.

the consistent selection procedure also outperforms all estimators of Hallin and Liska (2007) for most combinations of N and T.

Table 14 – Comparison of estimators of q based on Hallin and Liska (2007) under Design 1

N	T	$HL_{1,1}$	$HL_{1,2}$	$HL_{1,3}$	$HL_{2,1}$	$HL_{2,2}$	$HL_{2,3}$	N(0,1)	$z_{\alpha_{N,T}}$	Boot
100	100	3.80	3.81	3.92	3.94	3.86	3.87	3.19	3.13	2.93
100	200	3.34	3.18	3.41	3.98	3.97	3.98	3.27	3.19	2.93
200	100	3.84	3.81	3.92	3.96	3.91	3.91	3.09	3.05	2.97
200	200	3.20	3.04	3.28	3.98	3.94	3.98	3.13	3.06	2.98
200	300	3.30	3.16	3.46	3.99	3.98	3.99	3.16	3.06	2.98
400	100	3.83	3.83	3.92	3.96	3.91	3.92	3.07	3.03	3.02
400	200	3.18	3.00	3.27	3.97	3.92	3.98	3.07	3.02	3.02
400	300	3.31	3.18	3.42	3.98	3.97	3.99	3.09	3.02	3.01
400	600	2.90	2.73	3.03	3.97	3.95	3.98	3.12	3.03	3.02

Results for the data generating process of the main body are in Table 15. One of our estimators outperforms those of Hallin and Liska (2007) for all sample sizes but $(N, T) = (200, 300)$, in which case estimator $HL_{1,2}$ delivers the best results. Identical results hold for Design 2 of the previous sections, 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	$HL_{1,1}$	$HL_{1,2}$	$HL_{1,3}$	$HL_{2,1}$	$HL_{2,2}$	$HL_{2,3}$	N(0,1)	$z_{\alpha_{N,T}}$	Boot
100	100	5.67	5.62	5.82	5.76	5.61	5.62	5.17	5.12	5.01
100	200	5.29	5.62	5.37	5.96	5.95	5.96	5.30	5.20	5.03
200	100	5.71	5.67	5.85	5.84	5.73	5.74	5.07	5.04	5.06
200	200	5.14	4.95	5.23	5.97	5.93	5.97	5.12	5.06	5.05
200	300	5.22	5.03	5.38	5.96	5.95	5.96	5.17	5.06	5.05
400	100	5.76	5.68	5.86	5.86	5.75	5.75	5.04	5.01	5.07
400	200	5.09	4.94	5.20	5.97	5.93	5.97	5.06	5.02	5.06
400	300	5.26	5.03	5.40	5.96	5.95	5.96	5.08	5.02	5.06
400	600	4.82	4.50	4.92	5.97	5.95	5.97	5.12	5.03	5.05

Table 16 – Comparison of estimators of q based on Hallin and Liska (2007) under Design 2

N	T	$HL_{1,1}$	$HL_{1,2}$	$HL_{1,3}$	$HL_{2,1}$	$HL_{2,2}$	$HL_{2,3}$	N(0,1)	$z_{\alpha_{N,T}}$	Boot
100	100	7.65	7.50	7.73	7.22	7.14	7.26	8.13	8.09	8.05
100	200	7.52	7.34	7.52	7.89	7.86	7.90	8.25	8.18	8.05
200	100	7.66	7.54	7.75	7.38	7.41	7.48	8.06	8.03	8.07
200	200	7.33	7.22	7.41	7.94	7.89	7.94	8.10	8.04	8.07
200	300	7.23	7.09	7.41	7.97	7.97	7.97	8.11	8.05	8.06
400	100	7.68	7.57	7.78	7.48	7.47	7.54	8.03	8.01	8.07
400	200	7.30	7.20	7.38	7.94	7.87	7.95	8.06	8.02	8.06
400	300	7.25	7.05	7.39	7.97	7.96	7.97	8.06	8.02	8.06
400	600	6.77	6.52	6.92	7.97	7.96	7.98	8.09	8.02	8.06