Queen's Economics Department Working Paper No. 1392

Comovements in the Real Activity of Developed and Emerging Economies: A Test of Global versus Specific International Factors

Antoine A. Djogbenou
Queen's University

Department of Economics
Queen's University
94 University Avenue
Kingston, Ontario, Canada
K7L 3N6

4-2018

Comovements in the Real Activity of Developed and Emerging Economies: A Test of Global versus Specific International Factors *

Antoine A. Djogbenou[†]
Queen's University
antoined@econ.queensu.ca

April 9, 2018

Abstract

Although globalization has shaped the world economy in recent decades, emerging economies have experienced impressive growth compared to developed economies, suggesting a decoupling between developed and emerging business cycles. Using observed developed and emerging economy activity variables, we investigate whether the latter assertion can be supported by observed data. Based on a two-level factor model, we assume these activity variables can be decomposed into a global component, emerging or developed common component and idiosyncratic national shocks. We propose a statistical test for the null hypothesis of a one-level specification, where it is irrelevant to distinguish between emerging and developed latent factors against the two-level alternative. This paper provides a theoretical justification and simulations that document the testing procedure. An application of the test to a panel of developed and emerging countries leads to strong statistical evidence of decoupling.

Keywords: test statistic, latent factors, decoupling, emerging economies, developed economies.
JEL classification: C12, C55, F44, O47.

1 Introduction

The empirical and theoretical econometric analysis of high-dimensional factor models has been a heavily researched area since the seminal paper of Stock and Watson (2002). These models allow the reduction of a large set of macroeconomic and financial variables into a very small number of indexes, which are useful to span various information related to economic agents. Factor models generally assume a one-level structure, where the comovements into a large panel of variables can be summarized into a few latent factors affecting all variables. In particular, each variable in the large panel can decomposed in an idiosyncratic error component and a common component. See, e.g., Stock and Watson (2002), Bai and Ng (2002) and Bai (2003) for details. In many prediction and

*I am grateful to Sílvia Gonçalves, James G. MacKinnon, Morten Ø. Nielsen, Benoit Perron and participants to Queen's Economic Department Quantitative Workshop and the New York Camp Econometrics XIII for useful discussions and comments. I thank Queen's Economic Department for financial support.

[†]Address: Department of Economics, 94 University Avenue, Queen's University, Kingston, Ontario K7L 3N6, Canada. Email: antoined@econ.queensu.ca.

policy exercises, empirical researchers have found it useful to extract factors from a large set of series. Among others, Ludvigson and Ng (2007) have investigated the risk-return relation in the United States equity market based on extracted factors from a large panel of macroeconomic indicators. More recently, Aastveit, Bjørnland, and Thorsrud (2015) studied the role of the increased demand from emerging economies compared to those of developed economies as drivers of the real price of oil using a structural factor-augmented vector auto-regression, with factors from a large panel of emerging and developed activity variables.

Because the factors are latent, they are generally estimated in practice using the principal component method (PCM) assuming the one-level factor model. However, in many economic applications such as international business cycle studies, multi-level structures naturally arise in the specification of the common component. In such a case, specific factors of some groups of countries are allowed in addition to global factors. The illustration in this paper is similar to the one in Kose, Otrok, and Prasad (2012) and contributes to the debate on the existence of specific emerging economy activity factors. Kose, Otrok, and Prasad (2012) have investigated the decoupling between developed and emerging economy activity factors using these models. The motivation is twofold. First, the worldwide economy has become as interconnected as ever through an important increase in trade and free movement of capital. Second, a large share of the global growth has been accounted for by emerging economies. This high economic growth has seemed, at times, to have been unaffected by weak economic activity in developed countries. While the first point of view suggests a strong influence of global economy activity factors, the second suggests a specific emerging economy activity factor different from that in a developed economy.

As is well known, the PCM estimates of factors only converge to a rotation of the true factor space; see Bai (2003). In particular, if one is interested in understanding the role of these specific factors, it is important to know the functional form relating them to the estimated factors, which is impossible in practice. As Breitung and Eickmeier (2014) and Han (2016) also argued, the standard principal component is not able to separately identify specific factor spaces when multi-level common factor structures arise. Han (2016) alternatively suggests a shrinkage estimator and uses it to disentangle global macroeconomic factors that are Europe specific and U.S. specific. He finds that these specific estimated factors have high explanatory power for the leading economic indicators in Europe and the United States. Furthermore, Wang (2010) studies the estimation of multi-level factor models. In his empirical application, he decomposes the comovements within real and financial sectors in the United States economy.

The statistical evidence of the existence of factors that are specific to groups of variables has not been studied. Suppose for example that one is interested in the structural implications of specific real economy activity factors within develop and emerging economies. It is crucial to check whether the data support a coupling of these two economy activity factors or their decoupling into two specific factors. In this paper, we first illustrate the failure of the principal component estimates to separately identify the latent factor when a two-level specification holds. Second, we propose a test, which we justify to be theoretically valid, and document its finite sample performance through simulation experiments.

Although our theoretical results can be applied in other contexts, this work exclusively focuses on the international business cycle developments. It provides a statistical framework that can be used to analyze whether or not developed and emerging economy business cycles decouple. We propose a statistic that formally tests the coupling (corresponding to a one-level factor specification) against the decoupling (corresponding to a two-level factor specification) of developed and emerging activity factors. In the empirical application, we find strong statistical evidence against the null hypothesis that developed and emerging business cycles do not decouple.

Finally, we use a modification of the standard PCM called "sequential" which allows the identification of global and specific latent factors. This procedure considers initial estimated global and specific factors and iterates them until the sum of squared idiosyncratic residuals is minimized. Our analysis of the identified global, emerging and developed economy forces emphasizes their ability to capture major economic events during the sampling period.

The rest of the paper is organized as follows. In Section 2, we present motivational simulation experiments, propose a test of the one-level representation (coupling) against the two-level representation (decoupling) of the comovements within developed and emerging economies and justify its validity theoretically. In Section 3, we investigate the finite sample properties of the proposed test. In Section 4, we apply it to our data and estimate the different economy activity factors from a two-level factor panel model. Section 5 concludes the paper. Assumptions, proofs and the figures are relegated to the Appendix. Throughout the paper, $\lfloor \cdot \rfloor$ and $\|\cdot\|$ respectively denote the integer part of a number and the Euclidean norm.

2 Global, Emerging and Developed Economy Activity Factors

Emerging economies have become major players in the global world economy. On one hand, they have had high economic growth in recent decades compared to many developed economies developed economies. On the other hand, it is generally admitted that globalization has increased world economic interdependence, subsequently increasing the predominance of global economy activity factors. This paper contributes to the debate on the decoupling or not of the developed and emerging economy business cycles. We formally answer that question by investigating whether the distinction between specific factors of developed and emerging economy activities is relevant using a statistical test.

For our study, we consider a large set of N economic activity variables $\left(\boldsymbol{X} = \{X_{it}\}_{t=1,\cdots,T;\, i=1,\cdots,N}\right)$ on developed and emerging countries. We suppose that the $N_1 = \lfloor \alpha N \rfloor$ first rows of $\boldsymbol{X}$ contain information on developed countries. Furthermore, we assume that the comovements within $\boldsymbol{X}$ are reflected by latent global, developed and emerging economy activity factors, which are denoted by f_{0t}, f_{Dt} and f_{Et}, respectively. To separate these common international factors from idiosyncratic national shocks, we model the comovements within $\boldsymbol{X}$ using a two-level factor model specification. This model is motivated by the fact that it allows us to identify some common factors that capture comovements across the entire dataset or across subsets of the series. In our case, the global factor f_{0t} reflects fluctuations that are common across all variables and countries. The developed economy activity factor f_{Dt} and the emerging economy activity factor f_{Et}, respectively, capture fluctuations that are common to developed and emerging economies. The factor panel model can be written as

$$X_{it} = \lambda_{0i} f_{0t} + \lambda_{1i} f_{Dt} + e_{it} \text{ if } 1 \leq i \leq N_1 \tag{1}$$

and

$$X_{it} = \lambda_{0i} f_{0t} + \lambda_{1i} f_{Et} + e_{it} \text{ if } N_1 + 1 \leq i \leq N, \tag{2}$$

where e_{it} are the idiosyncratic errors. The factor loading λ_{0i} measures the exposure of variable i to the global economic activity factor whereas λ_{1i} reflects the variable X_{it} exposure respectively to developed and emerging country economy activity factors. The latent factors f_{Dt} and f_{Et} contain only information specific to each group of countries and not in the global economy activity factor f_{0t}. Therefore, the global factor is allowed to affect all the variables while the specific factors only affect the variables within their corresponding groups. When the distinction between the developed and emerging economy activity factors are irrelevant, there is a factor model representation where $f_{Dt} = f_{Et}$ at any time period.

To identify our factors of interest, we use the log differences of real gross domestic product and industrial production on 22 developed and 29 emerging countries as economy activity variables. The classification of the 51 developed and emerging countries is given in Section 4. The next subsection discusses the PCM estimation of the latent factors.

2.1 Principal Component Method and Identification of Specific Latent Factors

In practice, the activity factors are latent and need to be estimated. A popular approach consists of relying on principal component estimation. This procedure assumes a one-level factor model

$$X_{it} = \boldsymbol{\lambda}_i' \boldsymbol{f}_t + e_{it}, \; i = 1, \ldots, N \text{ and } t = 1, \ldots, T,$$

where $\boldsymbol{f}_t : r \times 1$ contains r latent factors and $\boldsymbol{\lambda}_i$ is the corresponding $r \times 1$ vector of factor loadings. Since the seminal paper of Stock and Watson (2002), where the extracted factors were used in the forecasting context, this approach has received considerable of attention in empirical and theoretical works. Given $\boldsymbol{X}$, the latent factor matrix $\boldsymbol{F} = [\boldsymbol{f}_1 \cdots \boldsymbol{f}_T]'$ is estimated by the matrix

$\tilde{\boldsymbol{F}} = \left[\tilde{\boldsymbol{f}}_1 \cdots \tilde{\boldsymbol{f}}_T\right]' : T \times r$ that is $\sqrt{T}$ times the eigenvectors corresponding to the r largest eigenvalues of $\boldsymbol{X}'\boldsymbol{X}/(TN)$ in decreasing order and using the normalization $\tilde{\boldsymbol{F}}'\tilde{\boldsymbol{F}}/T = \boldsymbol{I}_r$. Moreover, the matrix of factor loadings $\boldsymbol{\Lambda} = [\boldsymbol{\lambda}_1 \cdots \boldsymbol{\lambda}_N]$ is estimated by $\tilde{\boldsymbol{\Lambda}} = \left[\tilde{\boldsymbol{\lambda}}_1 \cdots \tilde{\boldsymbol{\lambda}}_N\right] = \boldsymbol{X}\tilde{\boldsymbol{F}}\left(\tilde{\boldsymbol{F}}'\tilde{\boldsymbol{F}}\right)^{-1} = \boldsymbol{X}\tilde{\boldsymbol{F}}/T$ regressing each column of $\boldsymbol{X}'$ on $\tilde{\boldsymbol{F}}$. As is well known, the PCM only consistently estimates the space spanned by the true factors, i.e., $\tilde{\boldsymbol{f}}_t = \boldsymbol{H}'\boldsymbol{f}_t + o_P(1)$, $t = 1, \ldots, T$, with $\boldsymbol{H}$ an $r \times r$ rotation matrix. Following Bai and Ng (2002) and Bai (2003), $\boldsymbol{H} = \frac{\boldsymbol{\Lambda}'\boldsymbol{\Lambda}}{N}\frac{\boldsymbol{F}'\tilde{\boldsymbol{F}}}{T}\tilde{\boldsymbol{V}}^{-1}$, where $\tilde{\boldsymbol{V}}$ is diagonal and contains the r largest eigenvalues of $\boldsymbol{X}'\boldsymbol{X}/NT$ in decreasing order. Consequently, the vector of the estimated factor loadings $\tilde{\boldsymbol{\lambda}}_i$ converges to a rotation of $\boldsymbol{\lambda}_i$, $\boldsymbol{H}^{-1}\boldsymbol{\lambda}_i$, such that

$$X_{it} = \boldsymbol{\lambda}_i'\boldsymbol{f}_t + e_{it} = \left(\boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\right)'(\boldsymbol{H}'\boldsymbol{f}_t) + e_{it}.$$

In a case where developed and emerging business cycles do not decouple, it becomes irrelevant to consider two different specific factors. Estimating in such a case $r = 2$ latent global factors based on the two most informative eigenvectors of $\boldsymbol{X}'\boldsymbol{X}/(TN)$ is sufficient to capture the comovements within $\boldsymbol{X}$. However, when the developed and emerging economy activity factors decouple, it becomes difficult to identify these two latent factors since $\boldsymbol{H}$ is unobserved and in general not a diagonal matrix asymptotically. Bai and Ng (2013) provide conditions that help identify the factors. For instance, they show that $\boldsymbol{H}$ is diagonal with ± 1 diagonal elements and it is possible to identify the true factors up to the sign when $\boldsymbol{F}'\boldsymbol{F}/T = \boldsymbol{I}_r$ and $\boldsymbol{\Lambda}'\boldsymbol{\Lambda}$ is diagonal. However, these conditions rely on latent factors and cannot be verified in practice. As a consequence, the principal component estimates only jointly identify the true factor space.

To illustrate the identification of developed and emerging economy activity factors by PCM estimates, we perform a Monte Carlo experiment using the previous two-level factor model. More precisely, we analyze how informative are the estimated factors in identifying $\boldsymbol{F}_D = [f_{D1} \cdots f_{DT}]'$ and $\boldsymbol{F}_E = [f_{E1} \cdots f_{ET}]'$ by regressing each of these latent factors on estimated factors individually or on groups of them, reporting the R^2 in each situation. A very high R^2 in this case means the estimated factor or the set of estimated factors have strong predictive ability for the latent factors and strongly contribute to identifying the considered true specific factors. We distinguish the situation where each true specific latent factor is regressed on the first and the second estimated factor individually or jointly, and also the case with three estimated factors as predictors. For the illustration, we assume $N_1 = N/2$ (i.e., $\alpha = 1/2$). For $t = 1, \ldots, T$, the global factor f_{0t} is supposed to be normally and independently distributed (NID) with mean 0 and variance 1. The developed activity factor f_{Dt} and the emerging economy specific factor f_{Et} are assumed to follow a $\mathrm{NID}(0,1)$ distribution. We let the exposure to each factor be $\lambda_{ji} \sim \mathrm{NID}(1,1)$, $j = 0, 1$ and the idiosyncratic errors $e_{it} \sim \mathrm{NID}(0,4)$. The choice of variance for e_{it} helps to insure that the factor model R^2 is 0.5. We set the mean of the latent factor loadings to 1 in order to allow a nonzero expected exposure to the true factors.

Tables 1 and 2 show that the two most informative estimated factors cannot consistently iden-

Table 1: Ability of PCM estimates to predict f_{Dt}

Predictors	R^2	T=50	T=100	T=150	T=200	T=250
$\tilde{f}_{1t}$	N=100	0.1480	0.1273	0.1237	0.1192	0.1189
	N=200	0.1471	0.1269	0.1200	0.1179	0.1147
$\tilde{f}_{2t}$	N=100	0.4477	0.4630	0.4587	0.4627	0.4634
	N=200	0.4667	0.4743	0.4771	0.4785	0.4770
$\tilde{f}_{1t}$, $\tilde{f}_{2t}$	N=100	0.5957	0.5903	0.5824	0.5819	0.5823
	N=200	0.6145	0.6012	0.5972	0.5964	0.5917
$\tilde{f}_{1t}$, $\tilde{f}_{2t}$, $\tilde{f}_{3t}$	N=100	0.9496	0.9519	0.9528	0.9531	0.9535
	N=200	0.9748	0.9759	0.9762	0.9764	0.9764

[a] This table presents the R^2 of the regression of true latent factors on estimated factors. It provides information about how informative the latter are with regard to the latent specific factors considered in this experiment. Because each estimated factor is not identifying one true latent factor, the estimated factors are indexed differently.

Table 2: Ability of PCM estimates to predict f_{Et}

Predictors	R^2	T=50	T=100	T=150	T=200	T=250
$\tilde{f}_{1t}$	N=100	0.1450	0.1285	0.1180	0.1181	0.1156
	N=200	0.1446	0.1245	0.1195	0.1158	0.1136
$\tilde{f}_{2t}$	N=100	0.4597	0.4575	0.4637	0.4658	0.4668
	N=200	0.4669	0.4732	0.4760	0.4774	0.4812
$\tilde{f}_{1t}$, $\tilde{f}_{2t}$	N=100	0.6047	0.5860	0.5853	0.5838	0.5823
	N=200	0.6115	0.5977	0.5955	0.5932	0.5947
$\tilde{f}_{1t}$, $\tilde{f}_{2t}$, $\tilde{f}_{3t}$	N=100	0.9497	0.9519	0.9528	0.9532	0.9533
	N=200	0.9746	0.9758	0.9762	0.9763	0.9765

[b] See Table 1[a].

tify latent emerging and developed activity factors. For any sample sizes, these two estimated factors cannot individually or jointly identify the latent emerging or developed activity factors. In particular, the low R^2 when regressing f_{Dt} and f_{Et} on $\tilde{f}_{1t}$ and $\tilde{f}_{2t}$ reflects their inability to help completely identify the emerging latent factor. Second, the latent factors can only be identified using at least three estimated factors. Therefore, unless we know the linear relationship between them, specific comovement to emerging or developed economies cannot be identified. In the presence of decoupled developed and emerging economy activity factors, the PCM estimates a factor space corresponding to the one-level representation of (1) and (2) given by

$$X_{it} = \phi_{0i} f_{0t} + \phi_{Di} f_{Dt} + \phi_{Ei} f_{Et} + e_{it}, i = 1, \ldots, N, \tag{3}$$

where

$$\phi_{0i} = \lambda_{0i},\ \phi_{Di} = \lambda_{1i},\ \phi_{Ei} = 0,\ \ i = 1, \ldots, N_1,$$
$$\phi_{0i} = \lambda_{0i},\ \phi_{D} = 0,\ \phi_{Ei} = \lambda_{1i},\ \ i = N_1 + 1, \ldots, N.$$

In practice, if one is only interested in simply forecasting a given variable using the PCM estimated factors, then it is sufficient to consider additional common factors as needed. However, if the interest is to understand how potential specific activity factors contribute to the fluctuation of a given variable (e.g., oil prices), it is crucial to be able to identify them individually. Because this is impossible in empirical works using PCM, where the true latent factors and factor loadings are not observed, it could be important to test whether the developed and emerging activity factors decoupled or do not. The next subsection addresses this issue.

2.2 Testing the Decoupling of Specific Factors

To study if we can separately identify the space spanned by the global and specific factor spaces given our data set, we consider the null hypothesis that $H_0 : \boldsymbol{F}_D = \boldsymbol{F}_E$, where $\boldsymbol{F}_D = [f_{D1} \cdots F_{DT}]'$ and $\boldsymbol{F}_E = [f_{E1} \cdots f_{ET}]'$ against $H_1 : \boldsymbol{F}_D \neq \boldsymbol{F}_E$. Therefore, under the null, (1) and (2) are equivalent to the one-level representation

$$\boldsymbol{X} = \boldsymbol{\Lambda} \boldsymbol{F}' + \boldsymbol{e}, \tag{4}$$

where the global latent factor matrix is $\boldsymbol{F} = [\boldsymbol{F}_0 \ \boldsymbol{F}_1]$, with $\boldsymbol{F}_0 = (f_{01} \cdots f_{0T})'$, $\boldsymbol{F}_1 = (f_{11} \cdots f_{1T})'$. The matrix of latent factor loadings is given by $\boldsymbol{\Lambda} = [\boldsymbol{\lambda}_1 \cdots \boldsymbol{\lambda}_N]' : N \times 2$, which reflects the exposure of developed and emerging economy activity variables to the global factors in $\boldsymbol{F}$. Define also the factor loadings second moment matrix under the null hypothesis as

$$\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}} = \mathrm{E}\left(\boldsymbol{\lambda}_i \boldsymbol{\lambda}_i'\right) = \left[\begin{array}{cc} \Sigma_{\boldsymbol{\Lambda}}^{(11)} & \Sigma_{\boldsymbol{\Lambda}}^{(12)} \\ \Sigma_{\boldsymbol{\Lambda}}^{(12)} & \Sigma_{\boldsymbol{\Lambda}}^{(22)} \end{array}\right] : 2 \times 2.$$

This testing problem is non-standard since $\boldsymbol{F}_D$ and $\boldsymbol{F}_E$ are very large $T \times 1$-vectors, and we are not able to directly apply the usual testing approaches. However, it can be viewed as a symmetric problem for testing the structural changes in factor loadings studied by Han and Inoue (2015) ($H_0 : \boldsymbol{\Lambda}_1 = \boldsymbol{\Lambda}_2$ in their set-up), with the factor playing a similar role as the factor loadings. Using arguments that are close to theirs, we compare developed countries and emerging countries' subsample second moments of estimated factor loadings. The intuition behind the proposed test statistic can be understood by analyzing the difference in the infeasible subsample of factor loading second moments.

Consider the matrix notation of the two-level alternative in (3), where the loadings associated with the developed countries are

$$\boldsymbol{\Phi}_D = \left[\begin{array}{ccc} \lambda_{01} & \lambda_{11} & 0 \\ \vdots & \vdots & \vdots \\ \lambda_{0N_1} & \lambda_{1N_1} & 0 \end{array}\right].$$

Those associated with emerging economies are

$$\mathbf{\Phi}_E = \begin{bmatrix} \lambda_{0(N_1+1)} & 0 & \lambda_{1(N_1+1)} \\ \vdots & \vdots & \vdots \\ \lambda_{0N} & 0 & \lambda_{1N} \end{bmatrix}.$$

Under the alternative, (1) and (2) together are equivalent to

$$\boldsymbol{X} = \begin{bmatrix} \mathbf{\Phi}_D \\ \mathbf{\Phi}_E \end{bmatrix} \begin{bmatrix} \boldsymbol{F}_0 & \boldsymbol{F}_D & \boldsymbol{F}_E \end{bmatrix}' + \boldsymbol{e}.$$

It turns out that the limit of the difference of the factor loading second moments in the equivalent one-level representation

$$\frac{1}{N_1}\mathbf{\Phi}_D'\mathbf{\Phi}_D - \frac{1}{N-N_1}\mathbf{\Phi}_E'\mathbf{\Phi}_E = \begin{bmatrix} \frac{1}{N_1}\sum_{i=1}^{N_1}\lambda_{0i}^2 - \frac{1}{N-N_1}\sum_{i=N_1+1}^{N}\lambda_{0i}^2 & \frac{1}{N_1}\sum_{i=1}^{N_1}\lambda_{0i}\lambda_{1i} & -\frac{1}{N-N_1}\sum_{i=N_1+1}^{N}\lambda_{0i}\lambda_{1i} \\ \frac{1}{N_1}\sum_{i=1}^{N_1}\lambda_{1i}\lambda_{0i} & \frac{1}{N_1}\sum_{i=1}^{N_1}\lambda_{1i}^2 & 0 \\ -\frac{1}{N-N_1}\sum_{i=N_1+1}^{N}\lambda_{1i}\lambda_{0i} & 0 & -\frac{1}{N-N_1}\sum_{i=N_1+1}^{N}\lambda_{1i}^2 \end{bmatrix}$$

is nonzero. In particular, from Assumption 3 (a) and (b) in the Appendix, we have

$$\frac{1}{N_1}\mathbf{\Phi}_D'\mathbf{\Phi}_D - \frac{1}{N-N_1}\mathbf{\Phi}_E'\mathbf{\Phi}_E = \begin{bmatrix} 0 & \Sigma_{\mathbf{\Lambda}}^{(12)} & -\Sigma_{\mathbf{\Lambda}}^{(12)} \\ \Sigma_{\mathbf{\Lambda}}^{(12)} & \Sigma_{\mathbf{\Lambda}}^{(22)} & 0 \\ -\Sigma_{\mathbf{\Lambda}}^{(12)} & 0 & -\Sigma_{\mathbf{\Lambda}}^{(22)} \end{bmatrix} + o_P(1).$$

Indeed, $\Sigma_{\mathbf{\Lambda}}^{(22)}$ is necessarily different from zero. Otherwise, the exposure of activity variables to specific factors would be zero. This also implies that when $\boldsymbol{F}_D$ and $\boldsymbol{F}_E$ decouple, the second moments of factor loadings over the subsamples change. Hence, $\sqrt{N}\left(\frac{1}{N_1}\mathbf{\Phi}_D'\mathbf{\Phi}_D - \frac{1}{N-N_1}\mathbf{\Phi}_E'\mathbf{\Phi}_E\right)$ diverges under the alternative.

In contrary, the difference of the second moments of factor loadings over the subsamples is asymptotically a matrix with all elements being zero under the null. Indeed, letting

$$\mathbf{\Psi}_D = \begin{bmatrix} \lambda_{01} & \lambda_{11} \\ \vdots & \vdots \\ \lambda_{0N_1} & \lambda_{1N_1} \end{bmatrix} : N_1 \times 2 \text{ and } \mathbf{\Psi}_E = \begin{bmatrix} \lambda_{0(N_1+1)} & \lambda_{1(N_1+1)} \\ \vdots & \vdots \\ \lambda_{0N} & \lambda_{1N} \end{bmatrix} : (N-N_1) \times 2,$$

we obtain under H_0 that

$$\frac{1}{N_1}\mathbf{\Psi}_D'\mathbf{\Psi}_D - \frac{1}{N-N_1}\mathbf{\Psi}_E'\mathbf{\Psi}_E = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} + o_P(1) = o_P(1).$$

Furthermore, we can derive the asymptotic distribution of $\sqrt{N}\left(\frac{1}{N_1}\mathbf{\Psi}_D'\mathbf{\Psi}_D - \frac{1}{N-N_1}\mathbf{\Psi}_E'\mathbf{\Psi}_E\right)$. However, because $\boldsymbol{F}$ and $\mathbf{\Lambda}$ are latent, they are estimated. Their principal component estimates $\tilde{\boldsymbol{F}}$ and $\tilde{\mathbf{\Lambda}}$ converge respectively to their rotated versions $\boldsymbol{FH}$ and $\mathbf{\Lambda}\boldsymbol{H}^{-1\prime}$ such that $\boldsymbol{X} = \left(\mathbf{\Lambda}\boldsymbol{H}^{-1\prime}\right)(\boldsymbol{FH})' + \boldsymbol{e}$ with $\boldsymbol{H}$ a 2×2 random matrix converging in probability to a nonsingular matrix $\boldsymbol{H}_0$. We therefore construct the test statistic based on the rotated latent factor loadings $\mathbf{\Lambda}\boldsymbol{H}_0^{-1\prime}$.

Define the vectorized rotated latent difference $\boldsymbol{A}\left(\alpha, \mathbf{\Lambda}\boldsymbol{H}_0^{\prime -1}\right)$, a $\frac{2(2+1)}{2} = 3$-dimensional vector, given by

$$\sqrt{N}\text{Vech}\left(\frac{1}{N_1}\sum_{i=1}^{N_1}\boldsymbol{H}_0^{-1}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'\boldsymbol{H}_0'^{-1}-\frac{1}{N-N_1}\sum_{i=N_1+1}^{N}\boldsymbol{H}_0^{-1}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'\boldsymbol{H}_0'^{-1}\right).$$

Under the independence and the existence of slightly more than fourth moment conditions on $\boldsymbol{\lambda}_i$, we show in Lemma A.2 that $\boldsymbol{A}\left(\alpha,\boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)$ is asymptotically normal. Consequently, the infeasible statistic

$$LM_N\left(\alpha,\boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)=\boldsymbol{A}\left(\alpha,\boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)'\left(\boldsymbol{S}\left(\alpha,\boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)\right)^{-1}\boldsymbol{A}\left(\alpha,\boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)\xrightarrow{d}\chi^2(3),$$

where

$$\boldsymbol{S}\left(\alpha,\boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)=\left(\frac{1}{\alpha}+\frac{1}{1-\alpha}\right)\frac{1}{N}\sum_{i=1}^{N}\text{Vech}\left(\boldsymbol{H}_0^{-1}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'\boldsymbol{H}_0'^{-1}-\boldsymbol{V}_0\right)\text{Vech}\left(\boldsymbol{H}_0^{-1}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'\boldsymbol{H}_0'^{-1}-\boldsymbol{V}_0\right)'$$

converges in probability to the variance of $\boldsymbol{A}\left(\alpha,\boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)$ and $\boldsymbol{V}_0=\boldsymbol{H}_0^{-1}\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\boldsymbol{H}_0'^{-1}$ is the rotated second moments of factor loadings. See the Appendix for the details of the proof.

Let $\tilde{\Lambda}=\left[\tilde{\boldsymbol{\lambda}}_1,\ldots,\tilde{\boldsymbol{\lambda}}_N\right]'$ is the PCM estimator of the whole matrix of factor loadings $\boldsymbol{\Lambda}$. Noting that both factors and factor loadings are estimated under the null, we define the statistic by

$$LM_N\left(\alpha,\tilde{\boldsymbol{\Lambda}}\right)=\boldsymbol{A}\left(\alpha,\tilde{\boldsymbol{\Lambda}}\right)'\left(\tilde{\boldsymbol{S}}\left(\alpha,\tilde{\boldsymbol{\Lambda}}\right)\right)^{-1}\boldsymbol{A}\left(\alpha,\tilde{\boldsymbol{\Lambda}}\right),\tag{5}$$

where

$$\boldsymbol{A}\left(\alpha,\tilde{\Lambda}\right)=\text{Vech}\left(\sqrt{N}\left(\frac{1}{N_1}\sum_{i=1}^{N_1}\tilde{\boldsymbol{\lambda}}_i\tilde{\boldsymbol{\lambda}}_i'-\frac{1}{N-N_1}\sum_{i=N_1+1}^{N}\tilde{\boldsymbol{\lambda}}_i\tilde{\boldsymbol{\lambda}}_i'\right)\right)\tag{6}$$

and $\tilde{\boldsymbol{S}}\left(\alpha,\tilde{\Lambda}\right)$ is the restricted long run variance estimator of $A\left(\alpha,\tilde{\Lambda}\right)$, with

$$\tilde{\boldsymbol{S}}\left(\alpha,\tilde{\boldsymbol{\Lambda}}\right)=\left(\frac{1}{\alpha}+\frac{1}{1-\alpha}\right)\frac{1}{N}\sum_{i=1}^{N}\text{Vech}\left(\tilde{\boldsymbol{\lambda}}_i\tilde{\boldsymbol{\lambda}}_i'-\boldsymbol{V}_{NT}\right)\text{Vech}\left(\tilde{\boldsymbol{\lambda}}_i\tilde{\boldsymbol{\lambda}}_i'-\boldsymbol{V}_{NT}\right)',\tag{7}$$

with $\boldsymbol{V}_{NT}=\frac{1}{N}\sum_{i=1}^{N}\tilde{\boldsymbol{\lambda}}_i\tilde{\boldsymbol{\lambda}}_i'$.

To obtain the limit distribution of the test statistic, these observations are combined with Lemma 2.1. This lemma establishes that $\boldsymbol{A}\left(\alpha,\tilde{\Lambda}\right)$ and $\tilde{\boldsymbol{S}}\left(\alpha,\tilde{\Lambda}\right)$ are close enough to their respective infeasible analogue $\boldsymbol{A}\left(\alpha,\boldsymbol{\Lambda}\boldsymbol{H}_0^{-1\prime}\right)$ and $\boldsymbol{S}\left(\alpha,\boldsymbol{\Lambda}\boldsymbol{H}_0^{-1\prime}\right)$ uniformly in $\alpha\in[\alpha_1,\alpha_2]\subset(0,1)$. In order to derive our results, we make Assumptions 1–3 presented in the Appendix. Assumptions 1 and 2 allow for weak dependence and heteroskedasticity in the idiosyncratic errors and are similar to the assumptions A−D of Bai and Ng (2002), 1−3 of Djogbenou, Gonçalves, and Perron (2015) and 1−2 of Djogbenou (2017). However, Assumption 2 (a), (c) and (d) which restricts the dependence between $\boldsymbol{f}_t$, $\boldsymbol{\lambda}_i$ and e_{it} among specific group of variables is slightly stronger. Assumption 3 is useful for deriving the asymptotic distribution of $\boldsymbol{A}\left(\alpha,\boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)$ using the Central Limit Theorem.

Lemma 2.1. *Suppose that Assumptions 1–3 are satisfied. As $N,T\to\infty$, if $\sqrt{N}/T\to 0$, then it holds that under the null,*

$$\left\| \boldsymbol{A}\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right) - \boldsymbol{A}\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0^{-1\prime}\right) \right\| = o_P(1), \tag{8}$$

and

$$\left\| \tilde{\boldsymbol{S}}\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right) - \boldsymbol{S}\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0^{-1\prime}\right) \right\| = o_P(1), \tag{9}$$

uniformly in $\alpha \in [\alpha_1, \alpha_2]$.

The following theorem, formally proved in the appendix, states the asymptotic null distribution of the test statistic.

Theorem 1. *Suppose that Assumptions 1–3 are satisfied. As* $N, T \to \infty$, *if* $\sqrt{N}/T \to 0$, *then it holds that under the null,*

$$LM_N\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right) \xrightarrow{d} \chi^2(3),$$

uniformly in $\alpha \in [\alpha_1, \alpha_2]$.

Theorem 1 suggests that we could test the null hypothesis of the one-level factor specification against the two-level one using the $LM_N\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right)$ statistic based on critical values from a Chi-squared distribution with 3 degrees of freedom. Alternatively to the restricted long run variance estimator, one could have considered a test statistic using an unrestricted estimator of the long run variance based on variance estimates over each subsample. However, the resulting test is asymptotically equivalent to the version suggested here, and we find that both statistics have similar size properties in simulation studies. We therefore focus on the simpler version presented here.

The next theorem shows that the test statistic has power against the alternative. In order to derive our result under the alternative, we use Assumption 4 (a)−(g) (in the Appendix) to complement Assumptions 1–3 as an additional factor arises in the one-level representation of the two-level alternative due to the decoupling. Assumption 4 (h) imposes positive definiteness and is a standard assumption.

Theorem 2. *Suppose that Assumptions 1–4 are satisfied. As* $N, T \to \infty$, *if* $\sqrt{N}/T \to 0$, *then there exists, under the two-level alternative, a non-random matrix* $\boldsymbol{R}_0 \neq \boldsymbol{0}$ *such that*

$$\frac{1}{N_1}\sum_{i=1}^{N_1} \tilde{\boldsymbol{\lambda}}_i \tilde{\boldsymbol{\lambda}}_i' - \frac{1}{N - N_1}\sum_{i=N_1+1}^{N} \tilde{\boldsymbol{\lambda}}_i \tilde{\boldsymbol{\lambda}}_i' \xrightarrow{P} \boldsymbol{R}_0$$

and

$$LM_N\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right) = N \text{Vech}\left(\boldsymbol{R}_0\right)' \boldsymbol{S}_0 \text{Vech}\left(\boldsymbol{R}_0\right) + o_P(N),$$

uniformly in $\alpha \in [\alpha_1, \alpha_2]$, *with* $\boldsymbol{S}_0$ *a constant matrix and* $\text{Vech}\left(\boldsymbol{R}_0\right)' \boldsymbol{S}_0 \text{Vech}\left(\boldsymbol{R}_0\right) > 0$.

Theorem 2 implies that under the alternative $LM_N\left(\alpha, \tilde{\Lambda}\right)$ diverges as sample sizes increase. In fact, $\boldsymbol{A}\left(\alpha, \tilde{\Lambda}\right)$ will tend to infinity given its $\sqrt{N}$ scaling. The proposed test statistic is easy to implement. Given the large $N \times T$ panel $\boldsymbol{X}$ with the first N_1 rows containing the developed countries and the remaining being the emerging countries, the steps for the test can be summarized into the following algorithm.

Algorithm for Implementing the Test Procedure.

1. Compute the estimated factors $(\tilde{\boldsymbol{F}})$: $\sqrt{T}$ times the eigenvectors corresponding to the 2 largest eigenvalues of $\boldsymbol{X}'\boldsymbol{X}/(TN)$ in decreasing order and using the normalization $\tilde{\boldsymbol{F}}'\tilde{\boldsymbol{F}}/T = \boldsymbol{I}_2$.

2. Compute the estimated factor loading $\tilde{\boldsymbol{\Lambda}} = \left[\tilde{\boldsymbol{\lambda}}_1 \cdots \tilde{\boldsymbol{\lambda}}_N\right]'$ is given by $\boldsymbol{X}\tilde{\boldsymbol{F}}/T$.

3. Find the scaled difference between the estimated second moments
$$\boldsymbol{A}\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right) = \sqrt{N}\text{Vech}\left(\frac{1}{N_1}\sum_{i=1}^{N_1}\tilde{\boldsymbol{\lambda}}_i\tilde{\boldsymbol{\lambda}}_i' - \frac{1}{N-N_1}\sum_{i=N_1+1}^{N}\tilde{\boldsymbol{\lambda}}_i\tilde{\boldsymbol{\lambda}}_i'\right).$$

4. Let $\boldsymbol{V}_{NT}$ denote the average over i of $\tilde{\boldsymbol{\lambda}}_i\tilde{\boldsymbol{\lambda}}_i'$. Compute $\tilde{\boldsymbol{S}}\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right)$, the long run variance estimator of $\boldsymbol{A}\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right)$ given by
$$\left(\frac{1}{\alpha} + \frac{1}{1-\alpha}\right)\frac{1}{N}\sum_{i=1}^{N}\text{Vech}\left(\tilde{\boldsymbol{\lambda}}_i\tilde{\boldsymbol{\lambda}}_i' - \boldsymbol{V}_{NT}\right)\text{Vech}\left(\tilde{\boldsymbol{\lambda}}_i\tilde{\boldsymbol{\lambda}}_i' - \boldsymbol{V}_{NT}\right)'.$$

5. Obtain the test statistic
$$LM_N\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right) = \boldsymbol{A}\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right)'\left(\tilde{\boldsymbol{S}}\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right)\right)^{-1}\boldsymbol{A}\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right).$$

6. Reject or do not using critical values or P-values from a $\chi^2(3)$.

Although we focus on the statistical test of the decoupling between developed and emerging business cycles, our results can be applied to other contexts and easily extended to case with more than one global factor or more than one specific factor in each group. Further, one could also investigate the test of factor models with more than two specific group or with a multi-level structure using arguments that are similar. These aspects are beyond the scope of this paper. We leave them for future research.

The asymptotic results presented above suggest that the proposed test should have good control of size and power as the cross-sectional and time dimensions increase. The next section reports simulation studies used to assess its finite sample properties.

3 Simulation Experiments

The Monte Carlo simulations are based on six different data generating processes (DGP), DGP 1-a, DGP 2-a, DGP 3-a, DGP 1-b, DGP 2-b and DGP 3-b. DGP 1-a, DGP 2-a and DGP 3-a are used to investigate the test size control while DGP 1-b, DGP 2-b and DGP 3-b evaluate the power of

Table 3: Rejection frequencies (%) for DGPs 1-a, 2-a and 3-a

DGP 1-a	T=50	T=100	T=150	T=200	T=250
N=100	4.44	4.49	4.48	4.44	4.64
N=200	4.84	4.66	4.71	4.72	4.81
DGP 2-a	T=50	T=100	T=150	T=200	T=250
N=100	6.91	6.09	5.79	5.58	5.62
N=200	6.42	5.87	5.53	5.58	5.41
DGP 3-a	T=50	T=100	T=150	T=200	T=250
N=100	4.43	4.49	4.57	4.58	4.45
N=200	4.83	4.57	4.78	4.69	4.81

[c] This table presents the rejection frequencies over 50000 simulated data when there are no specific factors and the level of the test is 5%.

the test. The six DGPs are all based on modifications of the simulation designs in Han and Inoue (2015) in order to incorporate two-level alternatives.

The first specification called DGP 1-a considers

$$X_{it} = \lambda_{0i} f_{0t} + \lambda_{1i} f_{1t} + \kappa e_{it},\ \ i = 1, \ldots, N \text{ and } t = 1, \ldots, T, \tag{10}$$

with

$$e_{it} \sim \text{NID}\,(0,1)\,,\ \ f_{jt} \sim \text{NID}\,(0,1)\,,\ \ \lambda_{ji} \sim \text{NID}\,(c,1)\,,\ \ j = 0,1.$$

We choose $\kappa = \sqrt{2\left(1+c^2\right)}$ such that $\text{R}^2 = 1 - \frac{\text{trace}(\text{E}(\boldsymbol{ee}'))}{\text{trace}(\text{E}(\boldsymbol{XX}'))} = 0.50$. DGP 2-a allows cross-sectional dependence in idiosyncratic errors

$$e_{it} = \sigma_i \left(u_{it} + \sum_{1 \leq |j| \leq P} \theta u_{(i-j)t} \right),\ \ u_{it} \sim \text{NID}\,(0,1)\,, \tag{11}$$

$$\sigma_i \sim U\,(0.5, 1.5) \ \text{ and } \kappa = \sqrt{\frac{24\left(1+c^2\right)}{13\left(1+2P\theta^2\right)}}, \tag{12}$$

where $\theta = 0.1$ and $P = 4$. DGP 3-a allows time dependence in $\boldsymbol{e}$ and $\boldsymbol{F}$. It differs from DGP 1-a by $f_{jt} = \rho_f f_{j(t-1)} + v_t$, $e_{it} = u_{it}\sigma_i$ and $u_{it} = \rho_e u_{i(t-1)} + w_{it}$, with

$$v_t \sim \text{NID}\left(0, 1-\rho_f^2\right),\ \ w_{it} \sim \text{NID}\left(0, 1-\rho_e^2\right) \text{ and } \sigma_i \sim \text{U}\,(0.5, 1.5)\,, \tag{13}$$

where $\rho_e = 0.5$, $\rho_f = 0.7$, $\kappa = \sqrt{\frac{24(1+c^2)}{13}}$ and U, the uniform distribution. In all settings, we simulate the data $M = 50000$ times, set $c = 1$ and use sample sizes (N, T) that belong to $\{100, 200\} \times \{50, 100, 150, 200, 250\}$. From Table 3, it follows that for different numbers of series and time periods, the sizes of the tests are around the 5% level.

DGP 1-a, DGP 2-a and DGP 3-a are now modified to allow for two-level specification and

Table 4: Rejection frequencies (%) for DGPs 1-b, 2-b and 3-b

DGP 1-b	T=50	T=100	T=150	T=200	T=250
N=100	87.02	86.38	86.16	86.07	86.04
N=200	91.83	90.82	90.18	89.70	89.59
DGP 2-b	T=50	T=100	T=150	T=200	T=250
N=100	89.02	89.20	89.51	89.44	89.31
N=200	92.51	91.85	91.43	91.38	91.30
DGP 3-b	T=50	T=100	T=150	T=200	T=250
N=100	87.29	87.01	87.12	86.56	86.44
N=200	93.06	92.50	91.74	91.17	91.13

[d] This table presents the rejection frequencies over 50000 simulated data when specific factors arise and the level of the test is 5%.

renamed DGP 1-b, DGP 2-b and DGP 3-b. These DGPs differ from DGP 1-a, DGP 2-a and DGP 3-a by the fact that f_{1t} is replaced by f_{Dt} when $i \leq \frac{N}{2}$, and by f_{Et} when $i \geq \frac{N}{2}+1$, which suggests specific factors when $i \leq \frac{N}{2}$ and $i \geq \frac{N}{2}+1$. In DGP 1-b and DGP 2-b,

$$f_{jt} \sim \text{NID}\left(0,1\right),$$

and in DGP 3-b,

$$f_{jt} = \rho_f f_{j(t-1)} + v_{jt} \text{ with } v_{jt} \sim \text{NID}\left(0, 1-\rho_f^2\right),$$

$j = D, E$. However, we introduce a parameter ρ representing the correlation between f_{Dt} and f_{Et} under the alternative hypothesis of decoupling, since they are not required to be independent in our theory. For DGP 1-b and DGP 2-b, this is done by simply drawing f_{Dt} and f_{Et} jointly from a normal distribution with mean $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and variance $\begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$. For DGP 3-b, we set the innovations in the AR(1) representation of f_{Dt} and f_{Et} have a correlation $\rho\left(1-\rho_f^2\right)$, implying a correlation ρ between f_{Dt} and f_{Et}. In particular, we assume $f_{Dt} = \rho_f f_{D(t-1)} + v_{Et}$ and $f_{Et} = \rho_f f_{E(t-1)} + v_{Et}$, with

$$\begin{pmatrix} v_{Dt} \\ v_{Et} \end{pmatrix} \sim \text{NID}\left(\mathbf{0}, \left(1-\rho_f^2\right)\begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right).$$

We set $\rho = 0.3$.

In these three cases, the test rejection frequencies of the null hypothesis over the 50000 replications are above 86% (See Table 4), which is a quiet high power. We observe also that significant increase in the rejection frequencies are associated with an increase in the cross-sectional dimension. In fact, the convergence in distribution of the statistic relies on the $\sqrt{N}$ convergence of $\boldsymbol{A}\left(\alpha, \boldsymbol{\Lambda H}_0'^{-1}\right)$.

We now modify DGP 3-b to investigate how a strong cyclical dependence in the latent factor

Table 5: Rejection frequencies (%) for DGP 3-b when $\rho_f = 0.9$

$\rho_e = 0.5$	T=50	T=100	T=150	T=200	T=250
N=100	88.14	89.33	89.14	88.61	88.51
N=200	94.15	94.55	94.24	93.42	93.40
$\rho_e = 0.7$	T=50	T=100	T=150	T=200	T=250
N=100	84.27	88.18	88.35	88.01	87.92
N=200	92.52	94.01	93.75	93.20	93.19
$\rho_e = 0.9$	T=50	T=100	T=150	T=200	T=250
N=100	66.40	80.47	84.86	85.79	86.22
N=200	81.74	90.78	91.97	92.03	92.14

[e] See Table 4 [d]

affects the power of the test. In this case, the auto-regressive parameter ρ_f is set to 0.9, which allows for a strong persistence of shocks to the latent factors. Table 5 reports rejection frequencies for different values of the time dependence parameter ρ_e associated with the idiosyncratic errors. Although the rejection frequencies decrease when time dependence in e_{it} is near the unit root, we recover the power when sample sizes increase.

Overall, the simulation exercises for different DGPs exhibit good control of size and power as the cross-sectional and the time dimensions change.

4 Application of the Test and Sequential PCM Estimation of the Latent Factors

The dataset covers the period from the third quarter of 1996 to the last quarter of 2014. We employ 89 series of the real gross domestic product and the industrial production of 22 developed and 29 emerging countries combining the classification in Aastveit, Bjørnland, and Thorsrud (2015) and Caldara, Cavallo, and Iacoviello (2016).

The developed countries are: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, Luxembourg, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom and United States. In addition, we have as emerging countries: Argentina, Brazil, Chile, China, Colombia, Czech Republic, Estonia, Hong Kong, Hungary, India, Indonesia, Israel, Jordan, Korea Republic, Latvia, Lithuania, Malaysia, Mexico, Peru, Philippines, Poland, Russia, Singapore, Slovak Republic, Slovenia, South Africa, Taiwan, Thailand and Turkey.

This sample contains the major developed countries which are United States, United Kingdom, Canada, France and Japan. Moreover, the emerging countries include the BRICS countries composed of Brazil, Russia, India, China and South Africa. We downloaded the data from the International Financial Statistics (IFS) database, the Global Economic Monitor (GEM) DataBank and the Organisation for Economic Co-operation and Development (OECD) database. We collected 22 series of real gross domestic product and 22 series of industrial production for developed countries. However, we used 26 series of gross domestic product and 19 series of industrial production for emerging countries. Due to a large number of periods with missing observations, we do not include gross domestic product series for Columbia, Poland and Malaysia. The industrial production series are not included for Argentina, Chile, China, Estonia, Indonesia, Jordan, Lithuania, Peru, Taiwan and Thailand.

Using this dataset on developed and emerging economy activities, we obtain $LM_N\left(\alpha, \tilde{\mathbf{\Lambda}}\right)$ is equal to 27.4499 while the 95% and 99% quantile of the $\chi^2(3)$ are 7.8147 and 11.3449, respectively. Thus, we conclude that there is strong statistical evidence against the null that developed and emerging business cycles do not decouple. This supports the assertion that the comovement in the real activity of emerging economies has decoupled from the one of developed economies in the recent decades. Given this evidence, the purpose in the remaining part of this section is the estimation of the latent factors in the two-level factor model using a procedure that can identify specific comovements in developed and emerging economy activity variables.

As we previously pointed out, the usual PCM is not able to separately identify specific factors under the two-level alternative. To solve that problem, Han (2016) considers a shrinkage estimator to consistently identify the true model specification. He estimates the multi-level factors based on two adaptive group least absolute shrinkage and selection operator (LASSO) estimators using a penalty term. An alternative well known procedure consists of employing a sequential principal component estimation as described by Breitung and Eickmeier (2014) and Wang (2010, Section 4.2). This approach updates an initial factor and factor loading estimates until convergence and is implemented as follows.

Algorithm for Sequential PCM Estimation.

1. Choose an initial global estimated factor $\check{\boldsymbol{F}}_0$ and corresponding factor loading $\check{\mathbf{\Lambda}}_0$ using all countries.

2. Obtain $\check{\boldsymbol{F}}_j$ and $\check{\mathbf{\Lambda}}_j$ by PCM estimation according to

$$X_{it}^{(j)} - \check{\lambda}_{0i}\check{f}_{0t} = \lambda_{ji} f_{jt} + u_{it}^{(j)},$$

   when $j = D$ for the developed country economic activity variables ($i = 1, \ldots, N_1$) and $j = E$ for the emerging country ones ($i = N_1 + 1, \ldots, N$).

3. Obtain new $\check{\boldsymbol{F}}_0$ and $\check{\mathbf{\Lambda}}_0$ by PCM estimation according to

$$y_{it} = \lambda_{0i} f_{0t} + v_{it},$$

Table 6: Ability of sequential PCM estimates to predict f_{0t}, f_{Dt} and f_{Et}

	Predictor	R^2	T=50	T=100	T=150	T=200	T=250
f_{0t}	$\hat{f}_{0t}$	N=100	0.9696	0.9714	0.9721	0.9724	0.9726
		N=200	0.9847	0.9857	0.9861	0.9862	0.9863
	Predictor	R^2	T=50	T=100	T=150	T=200	T=250
f_{Dt}	$\hat{f}_{Dt}$	N=100	0.9298	0.9420	0.9462	0.9481	0.9496
		N=200	0.9554	0.9660	0.9696	0.9714	0.9725
	Predictor	R^2	T=50	T=100	T=150	T=200	T=250
f_{Et}	$\hat{f}_{Et}$	N=100	0.9302	0.9420	0.9460	0.9484	0.9494
		N=200	0.9553	0.9659	0.9696	0.9712	0.9726

[g] See Table 1[a]

where $y_{it}^{(j)} = X_{it}^{(j)} - \check{\lambda}_{ji}\check{f}_{jt}$, $y_{it} = y_{it}^{(D)}$ for any developed economy activity variable and $y_{it} = y_{it}^{(E)}$ for emerging one.

4. Iterate steps 2 and 3 until the change in

$$\sum_{i,t} \left(y_{it} - \check{\lambda}_{0i}\check{f}_{0t} \right)^2$$

is lower than the tolerance level. When the convergence condition is reached, we get from the last iteration $\hat{f}_{0t} = \check{f}_{0t}$ and $\hat{f}_{jt} = \check{f}_{jt}$, $j = D, E$.

We use this method because it is easy to implement, consistently estimates the two-level factor space and minimizes the sum of squared residuals conditionally on the two-level restriction. It yields identical results up to a rotation to the sequential least squares approach, where initial global and specific estimates are also iteratively updated using least squares estimation. See Breitung and Eickmeier (2014) for details and other methods. Although, the sequential PCM algorithm identifies the specific factors, $\hat{\boldsymbol{F}}_{\boldsymbol{0}}'\hat{\boldsymbol{F}}_{\boldsymbol{j}}/T$, $j = D, E$ may be slightly different from zero in practice. As suggested by Wang (2010), one could impose the orthogonality restriction between steps 3 and 4 in the algorithm to obtain the specific factors. This consists of projecting the specific factors on the space orthogonal to the space associated with the global factor. More formally, between step 3 and 4, the specific estimated factors $\hat{\boldsymbol{F}}_j = \boldsymbol{M}_{\hat{\boldsymbol{F}}_0}\check{\boldsymbol{F}}_j$ with $\boldsymbol{M}_{\hat{\boldsymbol{F}}_0} = \boldsymbol{I}_T - \hat{\boldsymbol{F}}_0\left(\hat{\boldsymbol{F}}_0'\hat{\boldsymbol{F}}_0\right)^{-1}\hat{\boldsymbol{F}}_0'$ are used. Table 6 shows that in the two-level illustrative experiment in Section 2.1, the estimated global and specific factors strongly identify their corresponding true latent factors.

In order to understand the economic information behind the estimated global and specific factors from the large panel of economy activity variables, they are plotted over the sample time periods. Figures 1–3 in the Appendix plot the estimated sequential PCM global, developed and emerging economy activity factors. These estimated factors differently match the main developments in the recent global, developed and emerging economy business cycles reported by Kose, Otrok, and

Whiteman (2003), Kose, Otrok, and Prasad (2008), Kose, Otrok, and Prasad (2012) and recently by Charnavoki and Dolado (2014).

The global factor in Figure 1 captures the most important downturn in the recent decades in worldwide growth after the 2008 financial crisis following the collapse of Lehman Brothers. Likewise, Figure 2 captures the 2008 slow-down as well as the 1997 Asian financial crisis, which affected countries such as Thailand, Japan, Indonesia, South Korea, Hong Kong and Malaysia. Beginning in July 1997 after the Thailand foreign debt crisis, the financial woes spread to other Southeast Asian countries, inducing fears of worldwide contagion. The developed activity factor matches the slow economic growth after the August 2011 stock market fall due to fears of contagion of the European sovereign debt crisis (see, Figure 3). Hence, the estimated factor appeared to be able to capture big common shocks to the global economy as well as the one to developed and emerging specific economies. Although the observed large variation in the activity factors can be associated with important economic crisis, the visual analysis also suggests that many of the changes in the variability of economic activity are headed by the developed activity factors during our sampling period. Nevertheless, some variation through the global economy activity factors and some heterogeneous variation that are specific to emerging economies are captured as well.

The investigation of the role of the monetary policies, the changes in aggregate demand due to a change in government spending or the balance of payment surplus, or the fluctuations of commodity prices could certainly help understand these business cycle fluctuations. For instance, Caldara, Cavallo, and Iacoviello (2016) obtain in a dataset containing groups of countries that are close to ours, that emerging economies are net oil producer (+5 %) while the developed countries are net importer (-20 %). They find empirical evidence based on vector-autoregressive model that a drop in oil prices driven by oil supply shocks and global demand induces a boom in the activity of developed economies, while it leads to economic slow-down in emerging economies. Overall, it turns out that the identified factors in this paper could help deepen, in this context and more generally, the study of the interrelation between economic variables and economy activity factors using large scale datasets.

5 Conclusion

This paper contributes to the debate on the existence of a specific business cycle within emerging economies, different from developed ones. It investigates using a test statistic the ability to identify developed and emerging economy activity factors from a large panel of economic activity variables. Furthermore, we show the validity of the proposed testing procedure, provide evidence for its finite sample performance through Monte Carlo experiments and use it to test for decoupling between developed and emerging economy activity factors. In the empirical application, we find strong statistical evidence against the null hypothesis that developed and the emerging economy activity factors do not decouple. Finally, we identify these factors using a sequential PCM and find that they are able to track important economic events since 1996.

The suggested test statistic could be used in other contexts. For instance, one could use it to test whether the risk in the international agricultural market is fully characterized by global risk or whether some low-income and high-income country risks matter as well. Our work could also be extended to cases with more than two specific groups and more than two levels. These aspects are beyond the scope of this paper and are left for future research. In a different paper, we are investigating how this heterogeneity in international business cycles contributes to explain the fluctuation in the price of oil.

Appendix A: Assumptions and Proofs

Throughout this appendix, we let $\delta_{NT} = \min\left[\sqrt{N}, \sqrt{T}\right]$. When $\boldsymbol{M}$ is a matrix, $\boldsymbol{M} > \mathbf{0}$ means that $\boldsymbol{M}$ is positive definite. Denote C a generic finite constant. Note again that under the null, $\boldsymbol{\Lambda} = [\boldsymbol{\lambda}_1 \cdots \boldsymbol{\lambda}_N]' : N \times 2$ and $\boldsymbol{F} = [\boldsymbol{f}_1 \cdots \boldsymbol{f}_T]' : T \times 2$. As is well known, the principal component estimator $\tilde{\boldsymbol{f}}_t$ only consistently estimates a rotation of $\boldsymbol{f}_t$ given by $\boldsymbol{H}'\boldsymbol{f}_t$ (see, e.g., Bai and Ng (2002), Gonçalves and Perron (2014) and Djogbenou (2017)). As Bai and Ng (2002) show, $\boldsymbol{H} = \frac{\boldsymbol{\Lambda}'\boldsymbol{\Lambda}}{N} \frac{\boldsymbol{F}'\tilde{\boldsymbol{F}}}{T} \tilde{\boldsymbol{V}}^{-1}$, where $\tilde{\boldsymbol{V}}$ contains the $r = 2$ largest eigenvalues of $\boldsymbol{X}'\boldsymbol{X}/(NT)$ in decreasing order on the diagonal and is an $r \times r$ diagonal matrix. We also let $\iota_{Di} \equiv \mathrm{I}(i \leq \lfloor \alpha N \rfloor)$ and $\iota_{Ei} \equiv \mathrm{I}(i \geq \lfloor \alpha N \rfloor + 1)$, where $\mathrm{I}(\cdot)$ is an indicator function. Under the alternative, we define $\boldsymbol{G} = [\boldsymbol{F_0}\ \boldsymbol{F_D}\ \boldsymbol{F_E}] = [\boldsymbol{g}_1 \cdots \boldsymbol{g}_T]' : T \times 3$ and $\boldsymbol{\Phi} = [\boldsymbol{\phi}_1 \cdots \boldsymbol{\phi}_N]' : N \times 3$. Suppose that $\tilde{\boldsymbol{G}}$ is the PCM estimator of $\boldsymbol{G}$. Assume that the associated 3×3 rotation matrix is $\boldsymbol{\Xi}$, and $\boldsymbol{\Xi}_0$ is its limit. Let $\tilde{\Phi}$ denote the $N \times 3$ matrix of factor loadings associated with the three estimated factors in $\tilde{\boldsymbol{G}}$. Because we built the statistic using the estimated factor loadings associated with the first two estimated factors ($\tilde{\boldsymbol{F}}$), $\tilde{\boldsymbol{\Lambda}}$ corresponds to the first two columns of $\tilde{\boldsymbol{\Phi}}$ when the underlying DGP has a two-level structure. Suppose also that $\boldsymbol{J}$ is the $3 \times 2-$matrix composed of the first two columns of $\boldsymbol{\Xi}'^{-1}$, and $\boldsymbol{J}_0$ is the $3 \times 2-$matrix composed of the first two columns of $\boldsymbol{\Xi}_0'^{-1}$. To study the limiting null distributions of the suggested test statistics we invoke the following assumptions of the approximate factor model.

Assumption 1. (Factor model and idiosyncratic errors)

(a) $\mathrm{E}\left\|\boldsymbol{f}_t\right\|^4 \leq C$ and $\frac{1}{T}\boldsymbol{F}'\boldsymbol{F} = \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{f}_t\boldsymbol{f}_t' \xrightarrow{P} \boldsymbol{\Sigma_F} > 0$, where $\boldsymbol{\Sigma_F}$ is non-random.

(b) $\mathrm{E}\left\|\boldsymbol{\lambda}_i\right\|^4 \leq C$ and $\frac{1}{N}\boldsymbol{\Lambda}'\boldsymbol{\Lambda} = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{\lambda}_i\boldsymbol{\lambda}_i' \xrightarrow{P} \boldsymbol{\Sigma_\Lambda} > 0$, where $\boldsymbol{\Sigma_\Lambda}$ is non-random.

(c) The eigenvalues of the $r \times r$ matrix $(\boldsymbol{\Sigma_F} \times \boldsymbol{\Sigma_\Lambda})$ are distinct.

(d) $\mathrm{E}(e_{it}) = 0$, $\mathrm{E}\left|e_{it}\right|^8 \leq C$.

(e) $\mathrm{E}\left(e_{it}e_{js}\right)=\sigma_{ij,ts}$, $\left|\sigma_{ij,ts}\right|\leq\overline{\sigma}_{ij}$ for all (t,s) and $\left|\sigma_{ij,ts}\right|\leq\tau_{st}$ for all (i,j), with $\frac{1}{N}\sum_{i,j=1}^{N}\overline{\sigma}_{ij}\leq C$, $\frac{1}{T}\sum_{t,s=1}^{T}\tau_{st}\leq C$ and $\frac{1}{NT}\sum_{i,j,t,s=1}\left|\sigma_{ij,ts}\right|\leq C$.

(f) $\mathrm{E}\left|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(e_{it}e_{is}-E\left(e_{it}e_{is}\right)\right)\right|^{4}\leq C$ for all (t,s).

Assumption 2. (Moment conditions and weak dependence among $\{\boldsymbol{f}_t\}$, $\{\boldsymbol{\lambda}_i\}$ and $\{e_{it}\}$)

(a) $\mathrm{E}\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\boldsymbol{f}_t e_{it}\iota_{mi}\right\|^2\right)\leq C$, $m=D,E$, where $\mathrm{E}\left(\boldsymbol{f}_t e_{it}\right)=0$ for every (i,t).

(b) For each t, $\mathrm{E}\left\|\frac{1}{\sqrt{TN}}\sum_{s=1}^{T}\sum_{i=1}^{N}\boldsymbol{f}_s\left(e_{it}e_{is}-\mathrm{E}\left(e_{it}e_{is}\right)\right)\right\|^2\leq C$.

(c) $\mathrm{E}\left\|\frac{1}{\sqrt{TN}}\sum_{t=1}^{T}\sum_{i=1}^{N}\boldsymbol{f}_t\boldsymbol{\lambda}_i' e_{it}\iota_{mi}\right\|^2\leq C$, $m=D,E$, where $\mathrm{E}\left(\boldsymbol{f}_t\boldsymbol{\lambda}_i' e_{it}\right)=0$ for all (i,t).

(d) $\mathrm{E}\left(\frac{1}{T}\sum_{t=1}^{T}\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\lambda}_i e_{it}\iota_{mi}\right\|^2\right)\leq C$, $m=D,E$, where $\mathrm{E}\left(\boldsymbol{\lambda}_i e_{it}\right)=\mathbf{0}$ for all (i,t).

Assumption 3. (Conditions for the asymptotic normality of $\boldsymbol{A}\left(\alpha,\boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)$)

(a) The factor loadings $\{\boldsymbol{\lambda}_i\}_{i=1,\dots,N}$ are independent across i and $\mathrm{E}\left(\|\boldsymbol{\lambda}_i\|^{4+\xi}\right)\leq C$ for some $\xi>0$.

(b) $\mathrm{E}\left(\boldsymbol{\lambda}_i'\boldsymbol{\lambda}_i\right)=\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}$ and $\mathrm{E}\left(\left(\mathrm{Vech}\left(\boldsymbol{\lambda}_i'\boldsymbol{\lambda}_i\right)\right)'\mathrm{Vech}\left(\boldsymbol{\lambda}_i'\boldsymbol{\lambda}_i\right)\right)=\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}\boldsymbol{\Lambda}}$.

(c) The limit of $\mathrm{Var}\left(\boldsymbol{A}\left(\alpha,\boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)\right)$ is bounded and positive definite.

Assumption 4. (Additional conditions for the two-level factor model)

(a) $\frac{1}{T}\boldsymbol{G}'\boldsymbol{G}=\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{g}_t\boldsymbol{g}_t'\xrightarrow{P}\boldsymbol{\Sigma}_{\boldsymbol{G}}>\mathbf{0}$, where $\boldsymbol{\Sigma}_{\boldsymbol{G}}$ is non-random.

(b) $\frac{1}{N}\boldsymbol{\Phi}'\boldsymbol{\Phi}=\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\phi}_i\boldsymbol{\phi}_i'\xrightarrow{P}\boldsymbol{\Sigma}_{\boldsymbol{\Phi}}>\mathbf{0}$, where $\boldsymbol{\Sigma}_{\boldsymbol{\Phi}}$ is non-random.

(c) The eigenvalues of the matrix $\boldsymbol{\Sigma}_{\boldsymbol{G}}\times\boldsymbol{\Sigma}_{\boldsymbol{\Phi}}$ are distinct.

(d) $\mathrm{E}\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\boldsymbol{g}_t e_{it}\iota_{mi}\right\|^2\right)\leq C$, $m=D,E$, where $\mathrm{E}\left(\boldsymbol{g}_t e_{it}\right)=0$ and $\mathrm{E}\left(\boldsymbol{g}_t\boldsymbol{\phi}_i' e_{it}\right)=\mathbf{0}$ for every (i,t).

(e) For each t, $\mathrm{E}\left\|\frac{1}{\sqrt{TN}}\sum_{s=1}^{T}\sum_{i=1}^{N}\boldsymbol{g}_s\left(e_{it}e_{is}-\mathrm{E}\left(e_{it}e_{is}\right)\right)\right\|^2\leq C$.

(f) $\mathrm{E}\left\|\frac{1}{\sqrt{TN}}\sum_{t=1}^{T}\sum_{i=1}^{N}\boldsymbol{g}_t\boldsymbol{\phi}_i' e_{it}\iota_{mi}\right\|^2\leq C$, $m=D,E$.

(g) $\mathrm{E}\left\|\frac{1}{\sqrt{TN}}\sum_{t=1}^{T}\sum_{i=1}^{N}\boldsymbol{g}_t\boldsymbol{\phi}_i' e_{it}\iota_{mi}\right\|^2\leq C$, $m=D,E$, where $\mathrm{E}\left(\boldsymbol{g}_t\boldsymbol{\phi}_i' e_{it}\right)=0$ for all (i,t).

(h) The limit of the long run variance of $\boldsymbol{A}\left(\alpha,\boldsymbol{\Phi}\boldsymbol{J}\right)$ is positive definite.

We state the following results, which help to prove Lemma 2.1, Theorem 1 and Theorem 2.

Lemma A.1. *Suppose that Assumptions 1–3 are satisfied. If as $N, T \to \infty$, $\sqrt{N}/T \to 0$, then for any $\alpha \in [\alpha_1, \alpha_2]$, it holds that*

$$\frac{1}{\lfloor \alpha N \rfloor} \sum_{i=1}^{\lfloor \alpha N \rfloor} \left(\tilde{\boldsymbol{\lambda}}_i - \boldsymbol{H}^{-1} \boldsymbol{\lambda}_i \right) \boldsymbol{\lambda}_i' = O_P \left(\frac{1}{\delta_{NT}^2} \right), \tag{A.1}$$

$$\frac{1}{N - \lfloor \alpha N \rfloor} \sum_{i=\lfloor \alpha N \rfloor + 1}^{N} \left(\tilde{\boldsymbol{\lambda}}_i - \boldsymbol{H}^{-1} \boldsymbol{\lambda}_i \right) \boldsymbol{\lambda}_i' = O_p \left(\frac{1}{\delta_{NT}^2} \right), \tag{A.2}$$

$$\frac{1}{\lfloor \alpha N \rfloor} \sum_{i=1}^{\lfloor \alpha N \rfloor} \left\| \tilde{\boldsymbol{\lambda}}_i - \boldsymbol{H}^{-1} \boldsymbol{\lambda}_i \right\|^2 = O_p \left(\frac{1}{\delta_{NT}^2} \right), \tag{A.3}$$

$$\frac{1}{N - \lfloor \alpha N \rfloor} \sum_{i=\lfloor \alpha N \rfloor + 1}^{N} \left\| \tilde{\boldsymbol{\lambda}}_i - \boldsymbol{H}^{-1} \boldsymbol{\lambda}_i \right\|^2 = O_p \left(\frac{1}{\delta_{NT}^2} \right) \tag{A.4}$$

uniformly in α. It also holds that

$$\frac{1}{N} \sum_{i=1}^{N} \left\| \tilde{\boldsymbol{\lambda}}_i \tilde{\boldsymbol{\lambda}}_i' - \boldsymbol{H}^{-1} \boldsymbol{\lambda}_i \boldsymbol{\lambda}_i' \boldsymbol{H}'^{-1} \right\|^2 = O_P \left(\frac{N}{T^2} \right) \tag{A.5}$$

and

$$\frac{1}{N} \sum_{i=1}^{N} \left\| \boldsymbol{H}^{-1} \boldsymbol{\lambda}_i \boldsymbol{\lambda}_i' \boldsymbol{H}^{-1\prime} - \boldsymbol{H}_0^{-1} \boldsymbol{\lambda}_i \boldsymbol{\lambda}_i' \boldsymbol{H}_0'^{-1} \right\|^2 = o_P (1) . \tag{A.6}$$

Lemma A.2. *Suppose that Assumption 3 is satisfied. As $N \to \infty$, for any α such that $\alpha \in [\alpha_1, \alpha_2]$, it holds uniformly in α that*

$$\operatorname*{plim}_{N \to \infty} \boldsymbol{S} \left(\alpha, \boldsymbol{\Lambda} \boldsymbol{H}_0'^{-1} \right) = \lim_{N \to \infty} \operatorname{Var} \left(\boldsymbol{A} \left(\alpha, \boldsymbol{\Lambda} \boldsymbol{H}_0'^{-1} \right) \right), \tag{A.7}$$

$$\boldsymbol{A} \left(\alpha, \boldsymbol{\Lambda} \boldsymbol{H}_0'^{-1} \right) \xrightarrow{d} \mathrm{N} \left(\boldsymbol{0}, \lim_{N \to \infty} \operatorname{Var} \left(\boldsymbol{A} \left(\alpha, \boldsymbol{\Lambda} \boldsymbol{H}_0'^{-1} \right) \right) \right), \tag{A.8}$$

$$\frac{\sqrt{N}}{\lfloor \alpha N \rfloor} \sum_{i=1}^{\lfloor \alpha N \rfloor} \left(\boldsymbol{\lambda}_i \boldsymbol{\lambda}_i' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}} \right) = O_P (1) \ \textit{and}\ \frac{\sqrt{N}}{N - \lfloor \alpha N \rfloor} \sum_{i=\lfloor \alpha N \rfloor + 1}^{N} \left(\boldsymbol{\lambda}_i \boldsymbol{\lambda}_i' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}} \right) = O_P (1) . \tag{A.9}$$

The proof of (A.1) and (A.3) uses similar steps with Bai and Ng (2004, Lemma A3). Since (A.2) and (A.4) can be proved following nearly identical steps to (A.1) and (A.3) , they are omitted. The results (A.5) and (A.6) show how close are to the rotated cross-product of factor loading to their estimates. (A.7), (A.8) and (A.9) are useful to derive the limit distribution of the proposed test statistic. To obtain results under the alternative hypothesis, we rely on the following lemma.

Lemma A.3. *Suppose that Assumptions 1–3 and Assumption 4(a)–(g) are satisfied. If as $N, T \to$*

∞, $\sqrt{N}/T \to 0$, *then for any* $\alpha \in [\alpha_1, \alpha_2]$, *it holds uniformly in* α *that*

$$\frac{1}{\lfloor \alpha N \rfloor} \sum_{i=1}^{\lfloor \alpha N \rfloor} \left(\tilde{\boldsymbol{\lambda}}_i - \boldsymbol{J}' \boldsymbol{\phi}_i \right) \boldsymbol{\phi}_i' = O_P \left(\delta_{NT}^{-2} \right), \tag{A.10}$$

$$\frac{1}{N - \lfloor \alpha N \rfloor} \sum_{i=\lfloor \alpha N \rfloor + 1}^{N} \left(\tilde{\boldsymbol{\lambda}}_i - \boldsymbol{J}' \boldsymbol{\phi}_i \right) \boldsymbol{\phi}_i' = O_P \left(\delta_{NT}^{-2} \right), \tag{A.11}$$

$$\frac{1}{\lfloor \alpha N \rfloor} \sum_{i=1}^{\lfloor \alpha N \rfloor} \left\| \tilde{\boldsymbol{\lambda}}_i - \boldsymbol{J}' \boldsymbol{\phi}_i \right\|^2 = O_P \left(\delta_{NT}^{-2} \right), \tag{A.12}$$

$$\frac{1}{N - \lfloor \alpha N \rfloor} \sum_{i=\lfloor \alpha N \rfloor + 1}^{N} \left\| \tilde{\boldsymbol{\lambda}}_i - \boldsymbol{J}' \boldsymbol{\phi}_i \right\|^2 = O_P \left(\delta_{NT}^{-2} \right), \tag{A.13}$$

$$\frac{1}{N} \sum_{i=1}^{N} \left\| \tilde{\boldsymbol{\lambda}}_i \tilde{\boldsymbol{\lambda}}_i' - \boldsymbol{J}' \boldsymbol{\phi}_i \boldsymbol{\phi}_i' \boldsymbol{J} \right\|^2 = O_P \left(\frac{N}{T^2} \right) \tag{A.14}$$

and

$$\frac{1}{N} \sum_{i=1}^{N} \left\| \boldsymbol{J}' \boldsymbol{\phi}_i \boldsymbol{\phi}_i' \boldsymbol{J} - \boldsymbol{J}_0' \boldsymbol{\phi}_i \boldsymbol{\phi}_i' \boldsymbol{J}_0 \right\|^2 = o_P (1). \tag{A.15}$$

.

It also holds that

$$\left\| \tilde{\boldsymbol{S}} \left(\alpha, \tilde{\boldsymbol{\Lambda}} \right) - \boldsymbol{S} \left(\alpha, \boldsymbol{\Phi} \boldsymbol{J}_0 \right) \right\| = o_P (1) \tag{A.16}$$

uniformly in α

We next present the proof of Lemma A.1, Lemma A.2, Lemma 2.1, Theorem 1, Lemma A.3 Theorem 2.

A.1 Proof of Lemma A.1

The proof is subdivided into four parts corresponding to the proofs of (A.1), (A.3), (A.5) and (A.6).

Proof of (A.1) To demonstrate that uniformly in α, $\frac{1}{\lfloor \alpha N \rfloor} \sum_{i=1}^{\lfloor \alpha N \rfloor} \left(\tilde{\boldsymbol{\lambda}}_i - \boldsymbol{H}^{-1} \boldsymbol{\lambda}_i \right) \boldsymbol{\lambda}_i' = O_P \left(\frac{1}{\delta_{NT}^2} \right)$, we use the following identity from the proof of Bai and Ng (2004, Lemma A2).

$$\tilde{\boldsymbol{\lambda}}_i - \boldsymbol{H}^{-1} \boldsymbol{\lambda}_i = \frac{1}{T} \boldsymbol{H}' \boldsymbol{F}' \boldsymbol{e}_i + \frac{1}{T} \tilde{\boldsymbol{F}}' \left(\boldsymbol{F} - \tilde{\boldsymbol{F}} \boldsymbol{H}^{-1} \right) \boldsymbol{\lambda}_i + \frac{1}{T} \left(\tilde{\boldsymbol{F}} - \boldsymbol{F} \boldsymbol{H} \right)' \boldsymbol{e}_i \tag{A.17}$$

where $\boldsymbol{e}_i = (e_{i1}, e_{i2}, \ldots, e_{iT})'$. It follows uniformly in α that

$$\begin{aligned} & \frac{1}{\lfloor \alpha N \rfloor} \sum_{i=1}^{\lfloor \alpha N \rfloor} \left(\tilde{\boldsymbol{\lambda}}_i - \boldsymbol{H}^{-1} \boldsymbol{\lambda}_i \right) \boldsymbol{\lambda}_i' \\ = & \underbrace{\frac{1}{\lfloor \alpha N \rfloor} \sum_{i=1}^{\lfloor \alpha N \rfloor} \frac{1}{T} \boldsymbol{H}' \boldsymbol{F}' \boldsymbol{e}_i \boldsymbol{\lambda}_i'}_{\boldsymbol{A}_1} + \underbrace{\frac{1}{\lfloor \alpha N \rfloor} \sum_{i=1}^{\lfloor \alpha N \rfloor} \frac{1}{T} \tilde{\boldsymbol{F}}' \left(\boldsymbol{F} - \tilde{\boldsymbol{F}} \boldsymbol{H}^{-1} \right) \boldsymbol{\lambda}_i \boldsymbol{\lambda}_i'}_{\boldsymbol{A}_2} + \underbrace{\frac{1}{\lfloor \alpha N \rfloor} \sum_{i=1}^{\lfloor \alpha N \rfloor} \frac{1}{T} \left(\tilde{\boldsymbol{F}} - \boldsymbol{F} \boldsymbol{H} \right)' \boldsymbol{e}_i \boldsymbol{\lambda}_i'}_{\boldsymbol{A}_3}. \end{aligned}$$

Since $\mathrm{E}\left\|\frac{1}{NT}\sum_{i=1}^{\lfloor\alpha N\rfloor}\sum_{t=1}^{T}\boldsymbol{f}_t e_{it}\boldsymbol{\lambda}_i'\right\| = O\left(\frac{1}{\sqrt{NT}}\right)$ by Assumption 2 (c) and $\boldsymbol{H} = O_P(1)$, we have

$$\boldsymbol{A}_1 = \boldsymbol{H}\frac{1}{\lfloor\alpha N\rfloor T}\sum_{i=1}^{\lfloor\alpha N\rfloor}\sum_{t=1}^{T}\boldsymbol{f}_t e_{it}\boldsymbol{\lambda}_i' = O_P\left(\frac{1}{\sqrt{NT}}\right),$$

uniformly in α. Moreover, we also have

$$\boldsymbol{A}_2 = \frac{1}{T}\tilde{\boldsymbol{F}}'\left(\boldsymbol{F} - \tilde{\boldsymbol{F}}\boldsymbol{H}^{-1}\right)\frac{1}{\lfloor\alpha N\rfloor}\sum_{i=1}^{\lfloor\alpha N\rfloor}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i' = O_P\left(\frac{1}{\delta_{NT}^2}\right)$$

uniformly in α as $\frac{1}{T}\tilde{\boldsymbol{F}}'\left(\boldsymbol{F} - \tilde{\boldsymbol{F}}\boldsymbol{H}^{-1}\right) = O_P\left(\frac{1}{\delta_{NT}^2}\right)$ (see Bai (2003, Lemma B3)) and $\mathrm{E}\left\|\frac{1}{\lfloor\alpha N\rfloor}\sum_{i=1}^{\lfloor\alpha N\rfloor}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'\right\|$ is bounded by $\sup_{1\le i\le N}\mathrm{E}\left\|\boldsymbol{\lambda}_i\right\|^2 \le C$ given Assumption 1 (b). To study A_3, we use the following identity

$$\tilde{\boldsymbol{f}}_t - \boldsymbol{H}'\boldsymbol{f}_t = \tilde{\boldsymbol{V}}^{-1}\left(\frac{1}{T}\sum_{s=1}^{T}\tilde{\boldsymbol{f}}_s\gamma_{st} + \frac{1}{T}\sum_{s=1}^{T}\tilde{\boldsymbol{f}}_s\zeta_{st} + \frac{1}{T}\sum_{s=1}^{T}\tilde{\boldsymbol{f}}_s\eta_{st} + \frac{1}{T}\sum_{s=1}^{T}\tilde{\boldsymbol{f}}_s\xi_{st}\right),$$

where

$$\gamma_{st} = \mathrm{E}\left(\frac{1}{N}\sum_{i=1}^{N}e_{is}e_{it}\right),\ \zeta_{st} = \frac{1}{N}\sum_{i=1}^{N}\left(e_{is}e_{it} - \mathrm{E}\left(\frac{1}{N}\sum_{i=1}^{N}e_{is}e_{it}\right)\right),\ \eta_{st} = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\lambda}_i'\boldsymbol{f}_s e_{it},\ \xi_{st} = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\lambda}_i'\boldsymbol{f}_t e_{is}$$

and $\tilde{\boldsymbol{V}}^{-1}$ known to be $O_P(1)$ as $\tilde{\boldsymbol{V}} \xrightarrow{P} \boldsymbol{V} > \boldsymbol{0}$ (e.g., Bai and Ng (2002)). Thus $\boldsymbol{A}_3 = \tilde{\boldsymbol{V}}^{-1}(\boldsymbol{B}_1 + \boldsymbol{B}_2 + \boldsymbol{B}_3 + \boldsymbol{B}_4)$, where

$$\begin{aligned}
\boldsymbol{B}_1 &\equiv \frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\tilde{\boldsymbol{f}}_s\gamma_{st}\left(\frac{1}{\lfloor\alpha N\rfloor}\sum_{i=1}^{\lfloor\alpha N\rfloor}e_{it}\boldsymbol{\lambda}_i'\right),\\
\boldsymbol{B}_2 &\equiv \frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\tilde{\boldsymbol{f}}_s\zeta_{st}\left(\frac{1}{\lfloor\alpha N\rfloor}\sum_{i=1}^{\lfloor\alpha N\rfloor}e_{it}\boldsymbol{\lambda}_i'\right),\\
\boldsymbol{B}_3 &\equiv \frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\tilde{\boldsymbol{f}}_s\eta_{st}\left(\frac{1}{\lfloor\alpha N\rfloor}\sum_{i=1}^{\lfloor\alpha N\rfloor}e_{it}\boldsymbol{\lambda}_i'\right)
\end{aligned}$$

and

$$\boldsymbol{B}_4 \equiv \frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\tilde{\boldsymbol{f}}_s\xi_{st}\left(\frac{1}{\lfloor\alpha N\rfloor}\sum_{i=1}^{\lfloor\alpha N\rfloor}e_{it}\boldsymbol{\lambda}_i'\right).$$

For $\boldsymbol{B}_1$, we write

$$\begin{aligned}
\boldsymbol{B}_1 &= \frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\left(\tilde{\boldsymbol{f}}_s - \boldsymbol{H}'\boldsymbol{f}_s\right)\gamma_{st}\left(\frac{1}{\lfloor\alpha N\rfloor}\sum_{i=1}^{\lfloor\alpha N\rfloor}e_{it}\boldsymbol{\lambda}_i'\right) + \boldsymbol{H}'\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\boldsymbol{f}_s\gamma_{st}\left(\frac{1}{\lfloor\alpha N\rfloor}\sum_{i=1}^{\lfloor\alpha N\rfloor}e_{it}\boldsymbol{\lambda}_i'\right)\\
&\equiv B_{11} + B_{12}.
\end{aligned}$$

By an application of Cauchy-Schwarz inequality, $\boldsymbol{B}_{11}$ is bounded for any $\alpha \in [\alpha_1, \alpha_2]$ by

$$\begin{aligned}
&\frac{1}{T}\sum_{s=1}^{T}\left\|\tilde{\boldsymbol{f}}_s - \boldsymbol{H}'\boldsymbol{f}_s\right\|\left(\frac{1}{T}\sum_{t=1}^{T}\gamma_{st}^2\right)^{1/2}\left(\frac{1}{T}\sum_{t=1}^{T}\left(\left\|\frac{1}{\lfloor\alpha N\rfloor}\sum_{i=1}^{\lfloor\alpha N\rfloor}e_{it}\boldsymbol{\lambda}_i'\right\|\right)^2\right)^{1/2} \\
\leq\ &\left(\frac{1}{T}\sum_{s=1}^{T}\left\|\tilde{\boldsymbol{f}}_s - \boldsymbol{H}'\boldsymbol{f}_s\right\|^2\right)^{1/2}\left(\frac{1}{T^2}\sum_{s=1}^{T}\sum_{t=1}^{T}\gamma_{st}^2\right)^{1/2}\left(\frac{1}{T}\sum_{t=1}^{T}\left(\left\|\frac{1}{\lfloor\alpha N\rfloor}\sum_{i=1}^{\lfloor\alpha N\rfloor}e_{it}\boldsymbol{\lambda}_i'\right\|\right)^2\right)^{1/2} \\
=\ &O_P\left(\frac{1}{\delta_{NT}\sqrt{NT}}\right),
\end{aligned}$$

where from Bai and Ng (2002, Theorem 1), $\frac{1}{T}\sum_{s=1}^{T}\left\|\tilde{\boldsymbol{f}}_s - \boldsymbol{H}'\boldsymbol{f}_s\right\|^2 = O_P\left(1/\delta_{NT}^2\right)$, $\frac{1}{T}\sum_{s=1}^{T}\sum_{t=1}^{T}\gamma_{st}^2 = O\left(1\right)$ (see Bai and Ng (2002, Lemma 1(i))) and $\mathrm{E}\left(\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{\lfloor\alpha N\rfloor}e_{it}\boldsymbol{\lambda}_i'\right\|\right)^2 \leq C$ by Assumption 2 (d). Similarly, the second term $\boldsymbol{B}_{12}$ is bounded by

$$\|\boldsymbol{H}\|\left(\frac{1}{T}\sum_{s=1}^{T}\|\boldsymbol{f}_s\|^2\right)^{1/2}\left(\frac{1}{T^2}\sum_{s=1}^{T}\sum_{t=1}^{T}\gamma_{st}^2\right)^{1/2}\left(\frac{1}{T}\sum_{t=1}^{T}\left(\left\|\frac{1}{\lfloor\alpha N\rfloor}\sum_{i=1}^{\lfloor\alpha N\rfloor}e_{it}\boldsymbol{\lambda}_i'\right\|\right)^2\right)^{1/2} = O_P\left(\frac{1}{\sqrt{TN}}\right),$$

given that $\|\boldsymbol{H}\| = O_P(1)$ and $\mathrm{E}\|\boldsymbol{f}_s\|^2 \leq C$. Because $\boldsymbol{B}_{11} = O_P\left(\frac{1}{\delta_{NT}\sqrt{NT}}\right)$ and $\boldsymbol{B}_{12} = O_P\left(\frac{1}{\sqrt{NT}}\right)$, we deduce that $\boldsymbol{B}_1 = O_P\left(\frac{1}{\sqrt{NT}}\right)$. For $\boldsymbol{B}_2$, we start with the decomposition

$$\begin{aligned}
\boldsymbol{B}_2 &= \frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\left(\tilde{\boldsymbol{f}}_s - \boldsymbol{H}'\boldsymbol{f}_s\right)\zeta_{st}\left(\frac{1}{\lfloor\alpha N\rfloor}\sum_{i=1}^{\lfloor\alpha N\rfloor}e_{it}\boldsymbol{\lambda}_i'\right) + \boldsymbol{H}'\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\boldsymbol{f}_s\zeta_{st}\left(\frac{1}{\lfloor\alpha N\rfloor}\sum_{i=1}^{\lfloor\alpha N\rfloor}e_{it}\boldsymbol{\lambda}_i'\right) \\
&\equiv \boldsymbol{B}_{21} + \boldsymbol{B}_{22}.
\end{aligned}$$

The first term $\boldsymbol{B}_{21}$ is bounded by

$$\left(\frac{1}{T}\sum_{s=1}^{T}\left\|\tilde{\boldsymbol{f}}_s - \boldsymbol{H}'\boldsymbol{f}_s\right\|^2\right)^{1/2}\left(\frac{1}{T^2}\sum_{s=1}^{T}\sum_{t=1}^{T}\zeta_{st}^2\right)^{1/2}\left(\frac{1}{T}\sum_{t=1}^{T}\left(\left\|\frac{1}{\lfloor\alpha N\rfloor}\sum_{i=1}^{\lfloor\alpha N\rfloor}e_{it}\boldsymbol{\lambda}_i'\right\|\right)^2\right)^{1/2} = O_P\left(\frac{1}{\delta_{NT}\sqrt{TN}}\right),$$

where by Jensen inequality and Assumption 1 (f),

$$\mathrm{E}\left(\zeta_{st}^2\right) \leq \left(\mathrm{E}\left(\zeta_{st}^4\right)\right)^{1/2} = \frac{1}{N}\left(\mathrm{E}\left|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(e_{is}e_{it} - \mathrm{E}\left(\frac{1}{N}\sum_{i=1}^{N}e_{is}e_{it}\right)\right)\right|^4\right)^{1/2} \leq \frac{1}{N}C,$$

which implies $\frac{1}{T^2}\sum_{s=1}^{T}\sum_{t=1}^{T}\zeta_{st}^2 = O_P\left(\frac{1}{\sqrt{N}}\right)$. Using similar arguments and uniformly in α,

$$\boldsymbol{B}_{22} \leq \left(\frac{1}{T}\sum_{s=1}^{T}\|\boldsymbol{f}_s\|^2\right)^{1/2}\left(\frac{1}{T^2}\sum_{s=1}^{T}\sum_{t=1}^{T}\zeta_{st}^2\right)^{1/2}\left(\frac{1}{T}\sum_{t=1}^{T}\left(\left\|\frac{1}{\lfloor\alpha N\rfloor}\sum_{i=1}^{\lfloor\alpha N\rfloor}e_{it}\boldsymbol{\lambda}_i'\right\|\right)^2\right)^{1/2} = O_P\left(\frac{1}{\sqrt{TN}}\right).$$

Consequently, $\boldsymbol{B}_2 = O_P\left(\frac{1}{\sqrt{TN}}\right)$. We again write $\boldsymbol{B}_3 = \boldsymbol{B}_{31} + \boldsymbol{B}_{32}$, with

$$\boldsymbol{B}_{31} = \frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\left(\tilde{\boldsymbol{f}}_s - \boldsymbol{H}'\boldsymbol{f}_s\right)\eta_{st}\left(\frac{1}{\lfloor \alpha N \rfloor}\sum_{i=1}^{\lfloor \alpha N \rfloor} e_{it}\boldsymbol{\lambda}_i'\right) \text{ and } \boldsymbol{B}_{32} = \boldsymbol{H}'\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\boldsymbol{f}_s\eta_{st}\left(\frac{1}{\lfloor \alpha N \rfloor}\sum_{i=1}^{\lfloor \alpha N \rfloor} e_{it}\boldsymbol{\lambda}_i'\right).$$

We start with $\boldsymbol{B}_{31}$. By Cauchy-Schwarz inequality, we have

$$\begin{aligned}\frac{1}{T^2}\sum_{s=1}^{T}\sum_{t=1}^{T}\eta_{st}^2 &= \frac{1}{T^2}\sum_{s=1}^{T}\sum_{t=1}^{T}\left(\boldsymbol{f}_s'\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\lambda}_i\boldsymbol{f}_s e_{it}\right)^2 \le \frac{1}{T^2}\sum_{s=1}^{T}\sum_{t=1}^{T}\|\boldsymbol{f}_s\|^2\left\|\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\lambda}_i' e_{it}\right\|^2 \\ &= \left(\frac{1}{T}\sum_{s=1}^{T}\|\boldsymbol{f}_s\|^2\right)\frac{1}{N}\left(\frac{1}{T}\sum_{t=1}^{T}\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\lambda}_i' e_{it}\right\|^2\right) = O_P\left(\frac{1}{N}\right),\end{aligned}$$

as $\mathrm{E}\left(\frac{1}{T}\sum_{t=1}^{T}\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\lambda}_i' e_{it}\right\|^2\right) \le C$ follows from Assumption 2 (d) and the c_r inequality. Thus using also $\frac{1}{T}\sum_{s=1}^{T}\left\|\tilde{\boldsymbol{f}}_s - \boldsymbol{H}'\boldsymbol{f}_s\right\|^2 = O_P\left(\frac{1}{\delta_{NT}^2}\right)$ and Assumption 2 (d),

$$\|\boldsymbol{B}_{31}\| \le \left(\frac{1}{T}\sum_{s=1}^{T}\left\|\tilde{\boldsymbol{f}}_s - \boldsymbol{H}'\boldsymbol{f}_s\right\|^2\right)^{1/2}\left(\frac{1}{T^2}\sum_{s=1}^{T}\sum_{t=1}^{T}\eta_{st}^2\right)^{1/2}\left(\frac{1}{T}\sum_{t=1}^{T}\left(\left\|\frac{1}{\lfloor \alpha N \rfloor}\sum_{i=1}^{\lfloor \alpha N \rfloor} e_{it}\boldsymbol{\lambda}_i'\right\|\right)^2\right)^{1/2} = O_P\left(\frac{1}{\delta_{NT}N}\right).$$

By the same steps, we also obtain

$$\|\boldsymbol{B}_{32}\| \le \|H\|\left(\frac{1}{T}\sum_{s=1}^{T}\|\boldsymbol{f}_s\|^2\right)^{1/2}\left(\frac{1}{T^2}\sum_{s=1}^{T}\sum_{t=1}^{T}\eta_{st}^2\right)^{1/2}\left(\frac{1}{T}\sum_{t=1}^{T}\left(\left\|\frac{1}{\lfloor \alpha N \rfloor}\sum_{i=1}^{\lfloor \alpha N \rfloor} e_{it}\boldsymbol{\lambda}_i'\right\|\right)^2\right)^{1/2} = O_P\left(\frac{1}{N}\right).$$

Hence, $\boldsymbol{B}_3 = O_P\left(\frac{1}{N}\right)$. The proof for B_4 is similar to the proof of B_3 and is therefore omitted. From the order of $\boldsymbol{B}_1$, $\boldsymbol{B}_2$, $\boldsymbol{B}_3$ and $\boldsymbol{B}_4$, we deduce that $\boldsymbol{A}_3 = O_P\left(\delta_{NT}^{-2}\right)$ uniformly in α. Finally, we conclude that $\frac{1}{\lfloor \alpha N \rfloor}\sum_{i=1}^{\lfloor \alpha N \rfloor}\left(\tilde{\boldsymbol{\lambda}}_i - \boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\right)\boldsymbol{\lambda}_i' = O_P\left(\frac{1}{\delta_{NT}^2}\right)$ uniformly in α.

Proof of (A.3) For this proof, we use the decomposition (A.17), and obtain that uniformly in α,

$$\begin{aligned}&\frac{1}{\lfloor \alpha N \rfloor}\sum_{i=1}^{\lfloor \alpha N \rfloor}\left\|\tilde{\boldsymbol{\lambda}}_i - \boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\right\|^2 \\ \le\ & \underbrace{\frac{1}{\lfloor \alpha N \rfloor}\sum_{i=1}^{\lfloor \alpha N \rfloor}\left\|\frac{1}{T}\boldsymbol{H}'\boldsymbol{F}'\boldsymbol{e}_i\right\|^2}_{C_1} + \underbrace{\frac{1}{\lfloor \alpha N \rfloor}\sum_{i=1}^{\lfloor \alpha N \rfloor}\left\|\frac{1}{T}\tilde{\boldsymbol{F}}'\left(\boldsymbol{F} - \tilde{\boldsymbol{F}}\boldsymbol{H}^{-1}\right)\boldsymbol{\lambda}_i\right\|^2}_{C_2} + \underbrace{\frac{1}{\lfloor \alpha N \rfloor}\sum_{i=1}^{\lfloor \alpha N \rfloor}\left\|\frac{1}{T}\left(\tilde{\boldsymbol{F}} - \boldsymbol{F}\boldsymbol{H}\right)'\boldsymbol{e}_i\right\|^2}_{C_3}.\end{aligned}$$

Since $\frac{1}{\lfloor \alpha N \rfloor}\sum_{i=1}^{\lfloor \alpha N \rfloor}\mathrm{E}\left\|\frac{1}{\sqrt{T}}\boldsymbol{F}'\boldsymbol{e}_i\right\|^2 \le C$ by Assumption 2 (a) and $\boldsymbol{H} = O_P(1)$, we note that

$$C_1 \le \frac{1}{T}\|\boldsymbol{H}\|^2\frac{1}{\lfloor \alpha N \rfloor}\sum_{i=1}^{\lfloor \alpha N \rfloor}\mathrm{E}\left\|\frac{1}{\sqrt{T}}\boldsymbol{F}'\boldsymbol{e}_i\right\|^2 = O_P\left(\frac{1}{T}\right),$$

uniformly in α. Furthermore, we also have uniformly in α that

$$C_2 = \left\| \frac{1}{T} \tilde{\boldsymbol{F}}' \left(\boldsymbol{F} - \tilde{\boldsymbol{F}} \boldsymbol{H}^{-1} \right) \right\|^2 \frac{1}{\lfloor \alpha N \rfloor} \sum_{i=1}^{\lfloor \alpha N \rfloor} \|\boldsymbol{\lambda}_i\|^2 = O_P\left(\frac{1}{\delta_{NT}^4} \right)$$

as $\frac{1}{T} \tilde{\boldsymbol{F}}' \left(\boldsymbol{F} - \tilde{\boldsymbol{F}} \boldsymbol{H}^{-1} \right) = O_P\left(\frac{1}{\delta_{NT}^2} \right)$ from Bai (2003, Lemma B3) and $\frac{1}{\lfloor \alpha N \rfloor} \sum_{i=1}^{\lfloor \alpha N \rfloor} \|\boldsymbol{\lambda}_i\|^2 = O_P(1)$ as $\mathrm{E}\left(\|\boldsymbol{\lambda}_i\|^2 \right) \leq C$. We now similarly observe that by Cauchy-Schwarz inequality,

$$C_3 \leq \left(\frac{1}{T} \sum_{s=1}^{T} \left\| \tilde{\boldsymbol{f}}_t - \boldsymbol{H}' \boldsymbol{f}_t \right\|^2 \right) \left(\frac{1}{T \lfloor \alpha N \rfloor} \sum_{i=1}^{\lfloor \alpha N \rfloor} \sum_{t=1}^{T} e_{it}^2 \right) = O_P\left(\frac{1}{\delta_{NT}^2} \right) O_P(1) = O_P\left(\frac{1}{\delta_{NT}^2} \right).$$

From the bound for C_1, C_2 and C_3, $\frac{1}{\lfloor \alpha N \rfloor} \sum_{i=1}^{\lfloor \alpha N \rfloor} \left\| \tilde{\boldsymbol{\lambda}}_i - \boldsymbol{H}^{-1} \boldsymbol{\lambda}_i \right\|^2 = O_P\left(\frac{1}{\delta_{NT}^2} \right)$ uniformly in α.

Proof of (A.5) Using the identity

$$\begin{aligned} & \tilde{\boldsymbol{\lambda}}_i \tilde{\boldsymbol{\lambda}}_i' - \boldsymbol{H}^{-1} \boldsymbol{\lambda}_i \boldsymbol{\lambda}_i' \left(\boldsymbol{H}^{-1} \right)' \\ = \; & \left(\tilde{\boldsymbol{\lambda}}_i - \boldsymbol{H}^{-1} \boldsymbol{\lambda}_i \right) \left(\tilde{\boldsymbol{\lambda}}_i - \boldsymbol{H}^{-1} \boldsymbol{\lambda}_i \right)' + \boldsymbol{H}^{-1} \boldsymbol{\lambda}_i \left(\tilde{\boldsymbol{\lambda}}_i - \boldsymbol{H}^{-1} \boldsymbol{\lambda}_i \right)' + \left(\tilde{\boldsymbol{\lambda}}_i - \boldsymbol{H}^{-1} \boldsymbol{\lambda}_i \right) \boldsymbol{\lambda}_i' \left(\boldsymbol{H}^{-1} \right)', \end{aligned}$$

we can write by the c_r inequality and an application of the Cauchy-Schwarz inequality that $\frac{1}{N} \sum_{i=1}^{N} \left\| \tilde{\boldsymbol{\lambda}}_i \tilde{\boldsymbol{\lambda}}_i' - \boldsymbol{H}^{-1} \boldsymbol{\lambda}_i \boldsymbol{\lambda}_i' (\boldsymbol{H}^{-1})' \right\|^2$ is bounded by

$$\frac{1}{N} \sum_{i=1}^{N} \left\| \tilde{\boldsymbol{\lambda}}_i - \boldsymbol{H}^{-1} \boldsymbol{\lambda}_i \right\|^4 + 2 \left\| \boldsymbol{H}^{-1} \right\|^2 \left(\frac{1}{N} \sum_{i=1}^{N} \|\boldsymbol{\lambda}_i\|^4 \right)^{1/2} \left(\frac{1}{N} \sum_{i=1}^{N} \left\| \tilde{\boldsymbol{\lambda}}_i - \boldsymbol{H}^{-1} \boldsymbol{\lambda}_i \right\|^4 \right)^{1/2}.$$

As $\boldsymbol{H} \xrightarrow{P} \boldsymbol{H}_0$ where $\boldsymbol{H}_0$ is nonsingular, $\boldsymbol{H}^{-1} = O_P(1)$. Moreover, since $\mathrm{E}\left(\|\boldsymbol{\lambda}_i\|^4 \right)$, we also have that $\frac{1}{N} \sum_{i=1}^{N} \|\boldsymbol{\lambda}_i\|^4 = O_P(1)$. Hence, we only need to show that $\frac{1}{N} \sum_{i=1}^{N} \left\| \tilde{\boldsymbol{\lambda}}_i - \boldsymbol{H}^{-1} \boldsymbol{\lambda}_i \right\|^4 \xrightarrow{P} 0$. We use here nearly identical step to the proof of (A.3), with

$$\begin{aligned} & \frac{1}{N} \sum_{i=1}^{N} \left\| \tilde{\boldsymbol{\lambda}}_i - \boldsymbol{H}^{-1} \boldsymbol{\lambda}_i \right\|^4 \\ \leq \; & 3^3 \left(\underbrace{\frac{1}{N} \sum_{i=1}^{N} \left\| \frac{1}{T} \boldsymbol{H}' \boldsymbol{F}' \boldsymbol{e}_i \right\|^4}_{I_1} + \underbrace{\frac{1}{N} \sum_{i=1}^{N} \left\| \frac{1}{T} \tilde{\boldsymbol{F}}' \left(\boldsymbol{F} - \tilde{\boldsymbol{F}} \boldsymbol{H}^{-1} \right) \boldsymbol{\lambda}_i \right\|^4}_{I_2} + \underbrace{\frac{1}{N} \sum_{i=1}^{N} \left\| \frac{1}{T} \left(\tilde{\boldsymbol{F}} - \boldsymbol{F} \boldsymbol{H} \right)' \boldsymbol{e}_i \right\|^4}_{I_3} \right) \end{aligned}$$

based on (A.17) and the c_r inequality. As $\boldsymbol{H} = O_P(1)$ and $\frac{1}{N} \sum_{i=1}^{N} \left\| \frac{1}{\sqrt{T}} \boldsymbol{F}' \boldsymbol{e}_i \right\|^2 = O_P(1)$ from Assumption 2 (b), we note that

$$I_1 \leq \frac{1}{T^2} \|\boldsymbol{H}\|^4 \frac{1}{N} \sum_{i=1}^{N} \left\| \frac{1}{\sqrt{T}} \boldsymbol{F}' \boldsymbol{e}_i \right\|^4 \leq \frac{N}{T^2} \|\boldsymbol{H}\|^4 \left(\frac{1}{N} \sum_{i=1}^{N} \left\| \frac{1}{\sqrt{T}} \boldsymbol{F}' \boldsymbol{e}_i \right\|^2 \right)^2 = O_P\left(\frac{N}{T^2} \right).$$

Moreover, because $\frac{1}{T}\tilde{\boldsymbol{F}}'\left(\boldsymbol{F}-\tilde{\boldsymbol{F}}\boldsymbol{H}^{-1}\right)=O_P\left(\frac{1}{\delta_{NT}^2}\right)$ and $\frac{1}{N}\sum_{i=1}^N\|\boldsymbol{\lambda}_i\|^4=O_P(1)$ given $\mathrm{E}\left(\|\lambda_i\|^4\right)\leq C$, we obtain

$$I_2=\left\|\frac{1}{T}\tilde{\boldsymbol{F}}'\left(\boldsymbol{F}-\tilde{\boldsymbol{F}}\boldsymbol{H}^{-1}\right)\right\|^4\frac{1}{N}\sum_{i=1}^N\|\boldsymbol{\lambda}_i\|^4=O_P\left(\frac{1}{\delta_{NT}^8}\right).$$

We also have by Cauchy-Schwarz inequality that

$$I_3\leq\left(\frac{1}{T}\sum_{s=1}^T\left\|\tilde{\boldsymbol{f}}_t-\boldsymbol{H}'\boldsymbol{f}_t\right\|^2\right)^2\frac{1}{N}\sum_{i=1}^N\left(\frac{1}{T}\sum_{t=1}^Te_{it}^2\right)^2=\left(\frac{1}{T}\sum_{s=1}^T\left\|\tilde{\boldsymbol{f}}_t-\boldsymbol{H}'\boldsymbol{f}_t\right\|^2\right)^2\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^Te_{it}^4=O_P\left(\frac{1}{\delta_{NT}^4}\right),$$

where the first equality employs Jensen inequality. From the bound for I_1, I_2 and I_3, we deduce $\frac{1}{N}\sum_{i=1}^N\left\|\tilde{\boldsymbol{\lambda}}_i-\boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\right\|^4=O_P\left(\frac{N}{T^2}\right)$. Consequently, $\frac{1}{N}\sum_{i=1}^N\left\|\tilde{\boldsymbol{\lambda}}_i\tilde{\boldsymbol{\lambda}}_i'-\boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'\boldsymbol{H}^{-1\prime}\right\|^2=O_P\left(\frac{N}{T^2}\right)$.

Proof of (A.6) The result follows from the decomposition $\boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'\boldsymbol{H}^{-1\prime}-\boldsymbol{H}_0^{-1}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'\boldsymbol{H}_0^{-1\prime}=\boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'\left(\boldsymbol{H}^{-1}-\boldsymbol{H}_0^{-1}\right)'+\left(\boldsymbol{H}^{-1}-\boldsymbol{H}_0^{-1}\right)\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'\boldsymbol{H}_0^{-1\prime}$, and the c_r inequality. Indeed, it holds that

$$\frac{1}{N}\sum_{i=1}^N\left\|\boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'\boldsymbol{H}^{-1\prime}-\boldsymbol{H}_0^{-1}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'\boldsymbol{H}_0^{-1\prime}\right\|^2\leq\left(\left\|\boldsymbol{H}^{-1}\right\|^2+\left\|\boldsymbol{H}_0^{-1}\right\|^2\right)\left\|\boldsymbol{H}^{-1}-\boldsymbol{H}_0^{-1}\right\|^2\frac{1}{N}\sum_{i=1}^N\|\boldsymbol{\lambda}_i\|^4\xrightarrow{P}0,$$

using in particular the fact that $\boldsymbol{H}^{-1}-\boldsymbol{H}_0^{-1}=o_P(1)$ as $\boldsymbol{H}\xrightarrow{P}\boldsymbol{H}_0$, which is nonsingular.

A.2 Proof of Lemma A.2

For simplicity, we will let $N_1=\alpha N$ in this proofs.

Proof of (A.7) Define

$$\boldsymbol{A}_1\left(\alpha,\boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)=\frac{\sqrt{N}}{\alpha N}\sum_{i=1}^{\alpha N}\mathrm{Vech}\left(\boldsymbol{H}_0^{-1}\left(\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'-\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\boldsymbol{H}_0^{-1\prime}\right)\tag{A.18}$$

and

$$\boldsymbol{A}_2\left(\alpha,\boldsymbol{\Lambda}\left(\boldsymbol{H}_0'\right)^{-1}\right)=\frac{\sqrt{N}}{(1-\alpha)N}\sum_{i=1}^{(1-\alpha)N}\mathrm{Vech}\left(\boldsymbol{H}_0\left(\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'-\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\boldsymbol{H}_0\right).\tag{A.19}$$

Since $\boldsymbol{S}\left(\alpha,\boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)=\boldsymbol{S}_1\left(\alpha,\boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)+\boldsymbol{S}_2\left(\alpha,\boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)$, where

$$\begin{aligned}\boldsymbol{S}_1\left(\alpha,\boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)&=\frac{1}{\alpha}\frac{1}{N}\sum_{i=1}^N\left(\mathrm{Vech}\left(\boldsymbol{H}_0^{-1}\left(\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'-\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\boldsymbol{H}_0'^{-1}\right)\mathrm{Vech}\left(\boldsymbol{H}_0^{-1}\left(\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'-\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\boldsymbol{H}_0'^{-1}\right)'\right),\\ \boldsymbol{S}_2\left(\alpha,\boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)&=\frac{1}{1-\alpha}\frac{1}{N}\sum_{i=1}^N\left(\mathrm{Vech}\left(\boldsymbol{H}_0^{-1}\left(\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'-\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\boldsymbol{H}_0'^{-1}\right)\mathrm{Vech}\left(\boldsymbol{H}_0^{-1}\left(\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'-\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\boldsymbol{H}_0'^{-1}\right)'\right),\end{aligned}$$

the proof proceeds by showing that

$$\operatorname*{plim}_{N\to\infty}\boldsymbol{S}_1\left(\alpha,\boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)=\lim_{N\to\infty}\mathrm{Var}\left(\boldsymbol{A}_1\left(\alpha,\boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)\right)\tag{A.20}$$

and

$$\underset{N\to\infty}{\text{plim}}\, \boldsymbol{S}_2\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right) = \lim_{N\to\infty} \text{Var}\left(\boldsymbol{A}_2\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)\right). \tag{A.21}$$

Let start with (A.20). We have from Assumption 3 (a)−(b) that

$$\begin{aligned}\text{Var}\left(\boldsymbol{A}_1\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)\right) &= \frac{1}{\alpha}\frac{1}{\alpha N}\sum_{i=1}^{\alpha N} \text{E}\left(\text{Vech}\left(\boldsymbol{H}_0^{-1}\left(\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\boldsymbol{H}_0'^{-1}\right)\text{Vech}\left(\boldsymbol{H}_0^{-1}\left(\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\boldsymbol{H}_0'^{-1}\right)'\right).\\ &= \text{E}\left(\boldsymbol{S}_1\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)\right).\end{aligned}$$

In consequence, to show (A.20), we only need to prove that $\boldsymbol{a}'\left(\boldsymbol{S}_1\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right) - \text{E}\left(\boldsymbol{S}_1\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)\right)\right)\boldsymbol{a}' = o_P(1)$ where $\boldsymbol{a}$ is a 3-dimensional vector of real such that $\boldsymbol{a}'\boldsymbol{a} = 1$. We have

$$\boldsymbol{a}'\left(\boldsymbol{S}_1\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right) - \text{E}\left(\boldsymbol{S}_1\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)\right)\right)\boldsymbol{a}' = \frac{1}{\alpha N}\sum_{i=1}^{N} s_i,$$

with

$$\begin{aligned}s_i &= \boldsymbol{a}'(\text{Vech}\left(\boldsymbol{H}_0^{-1}\left(\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\boldsymbol{H}_0'^{-1}\right)\text{Vech}\left(\boldsymbol{H}_0^{-1}\left(\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\boldsymbol{H}_0'^{-1}\right)'\\ &\quad -\text{E}\left(\text{Vech}\left(\boldsymbol{H}_0^{-1}\left(\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\boldsymbol{H}_0'^{-1}\right)\text{Vech}\left(\boldsymbol{H}_0^{-1}\left(\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\boldsymbol{H}_0'^{-1}\right)'\right))\boldsymbol{a}.\end{aligned}$$

Since, $\sum_{i=1}^{N} \text{E}(s_i) = 0$, Djogbenou, MacKinnon, and Nielsen (2017, Lemma A1) imply that $\sum_{i=1}^{N} s_i = O_P\left(N^{1/2}\right)$ if $\text{E}(|s_i|) \le C$. Using the triangular inequality and the Jensen inequality,

$$\text{E}(|s_i|) \le 2\text{E}\left|\boldsymbol{a}'\left(\text{Vech}\left(\boldsymbol{H}_0^{-1}\left(\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\boldsymbol{H}_0'^{-1}\right)\text{Vech}\left(\boldsymbol{H}_0^{-1}\left(\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\boldsymbol{H}_0'^{-1}\right)'\right)\boldsymbol{a}\right|.$$

In consequence by an application of Cauchy Schwarz inequality and the c_r inequality,

$$\text{E}(|s_i|) \le 2\text{E}\left\|\text{Vech}\left(\boldsymbol{H}_0^{-1}\left(\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\boldsymbol{H}_0'^{-1}\right)\right\|^2 \le 4\left\|\boldsymbol{H}_0^{-1}\right\|^4 \text{E}\left\|\boldsymbol{\lambda}_i\right\|^4 \le C.$$

Hence $\frac{1}{\alpha N}\sum_{i=1}^{N} s_i = O_P\left(N^{-1/2}\right) = o_P(1)$ and we deduce $\underset{N\to\infty}{\text{plim}}\, \boldsymbol{S}_1\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right) = \lim_{N\to\infty}\text{Var}\left(\boldsymbol{A}_1\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)\right)$. Using a similar argument, $\underset{N\to\infty}{\text{plim}}\, \boldsymbol{S}_2\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right) = \lim_{N\to\infty}\text{Var}\left(\boldsymbol{A}_2\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)\right)$. We finally deduce

$$\underset{N\to\infty}{\text{plim}}\left(\boldsymbol{S}_1\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right) + \boldsymbol{S}_2\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)\right) = \lim_{N\to\infty}\text{Var}\left(\boldsymbol{A}_1\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)\right) + \lim_{N\to\infty}\text{Var}\left(\boldsymbol{A}_2\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)\right)$$

which implies given Assumption 3 (a) that $\underset{N\to\infty}{\text{plim}}\, \boldsymbol{S}\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right) = \lim_{N\to\infty}\text{Var}\left(\boldsymbol{A}\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)\right)$.

Proof of (A.8) Recall that $\boldsymbol{A}_1\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right) = \frac{\sqrt{N}}{\alpha N}\sum_{i=1}^{\alpha N}\text{Vech}\left(\boldsymbol{H}_0^{-1}\left(\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\boldsymbol{H}_0'^{-1}\right)$. In this section, we prove that

$$\boldsymbol{A}_1\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right) \xrightarrow{d} \text{N}\left(0, \lim_{N\to\infty}\text{Var}\left(\boldsymbol{A}_1\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)\right)\right). \tag{A.22}$$

Since the limit of $\text{Var}\left(\boldsymbol{A}_1\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)\right)$ is positive definite under Assumption 3 (c), we simply need to show that $\left(\text{Var}\left(\boldsymbol{A}_1\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)\right)\right)^{-1/2}\boldsymbol{A}_1\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right) \xrightarrow{d} \text{N}(\boldsymbol{0}, \boldsymbol{I}_3)$. Let $\boldsymbol{a}$ be a 3-dimensional vector of real such that $\boldsymbol{a}'\boldsymbol{a} = 1$. Define

$$z_i = \boldsymbol{a}' \left(\text{Var} \left(\boldsymbol{A}_1 \left(\alpha, \boldsymbol{\Lambda} \boldsymbol{H}_0'^{-1} \right) \right) \right)^{-1/2} \frac{\sqrt{N}}{\alpha N} \text{Vech} \left(\boldsymbol{H}_0^{-1} \left(\boldsymbol{\lambda}_i \boldsymbol{\lambda}_i' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}} \right) \boldsymbol{H}_0'^{-1} \right).$$

It follows that, $\text{E} \left(\sum_{i=1}^{\alpha N} z_i \right) = 0$ and $\text{E} \left(\sum_{i=1}^{N} z_i^2 \right) = 1$. Therefore (A.8) follows from the Lyapunov Central Limit Theorem for heterogeneous, independent random variables if for some $\xi > 0$, $\sum_{i=1}^{\alpha N} \text{E}|z_i|^{2+\xi} \to 0$ (Lyapunov's condition). To prove the latter, we use the bound

$$\sum_{i=1}^{\alpha N} \text{E}|z_i|^{2+\xi} \leq \left\| \left(\text{Var} \left(\boldsymbol{A}_1 \left(\alpha, \boldsymbol{\Lambda} \boldsymbol{H}_0'^{-1} \right) \right) \right)^{-1/2} \right\|^{2+\xi} \left(\frac{\sqrt{N}}{\alpha N} \right)^{2+\xi} \sum_{i=1}^{N} \text{E} \left\| \text{Vech} \left(\boldsymbol{H}_0^{-1} \left(\boldsymbol{\lambda}_i \boldsymbol{\lambda}_i' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}} \right) \boldsymbol{H}_0'^{-1} \right) \right\|^{2+\xi}. \tag{A.23}$$

Since by an application of the c_r-inequality, the fact that $\text{E} \left(\boldsymbol{\lambda}_i \boldsymbol{\lambda}_i' \right) = \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}$ and the Jensen inequality, we have $\text{E} \left\| \text{Vech} \left(\boldsymbol{H}_0^{-1} \left(\boldsymbol{\lambda}_i \boldsymbol{\lambda}_i' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}} \right) \boldsymbol{H}_0'^{-1} \right) \right\|^{2+\xi} \leq \left\| \boldsymbol{H}_0^{-1} \right\|^{4+2\xi} 2^{1+\xi} \times 2\text{E} \left\| \boldsymbol{\lambda}_i \boldsymbol{\lambda}_i' \right\|^{2+\xi} \leq C$ if $\text{E} \left\| \boldsymbol{\lambda}_i \right\|^{4+2\xi} \leq C$, we deduce that $\sum_{i=1}^{N} \text{E}|z_i|^{2+\xi} = O \left(\left(\frac{\sqrt{N}}{\alpha N} \right)^{2+\xi} \alpha N \right) = O \left(\frac{1}{N^{\xi/2}} \right) = o_P(1)$ given that $\text{Var} \left(\boldsymbol{A}_1 \left(\alpha, \boldsymbol{\Lambda} \boldsymbol{H}_0'^{-1} \right) \right) = O(1)$. It similarly follows that

$$\boldsymbol{A}_2 \left(\alpha, \boldsymbol{\Lambda} \left(\boldsymbol{H}_0' \right)^{-1} \right) \xrightarrow{d} \text{N} \left(\boldsymbol{0}, \lim_{N \to \infty} \text{Var} \left(\boldsymbol{A}_2 \left(\alpha, \boldsymbol{\Lambda} \boldsymbol{H}_0'^{-1} \right) \right) \right). \tag{A.24}$$

Because $\boldsymbol{A}_1 \left(\alpha, \boldsymbol{\Lambda} \boldsymbol{H}_0'^{-1} \right)$ and $\boldsymbol{A}_2 \left(\alpha, \boldsymbol{\Lambda} \boldsymbol{H}_0'^{-1} \right)$ are independent, with mean zero and

$$\lim_{N \to \infty} \text{Var} \left(\boldsymbol{A} \left(\alpha, \boldsymbol{\Lambda} \boldsymbol{H}_0'^{-1} \right) \right) = \lim_{N \to \infty} \text{Var} \left(\boldsymbol{A}_1 \left(\alpha, \boldsymbol{\Lambda} \boldsymbol{H}_0'^{-1} \right) \right) + \lim_{N \to \infty} \text{Var} \left(\boldsymbol{A}_2 \left(\alpha, \boldsymbol{\Lambda} \boldsymbol{H}_0'^{-1} \right) \right)$$

is positive definite, we conclude that $\boldsymbol{A} \left(\alpha, \boldsymbol{\Lambda} \boldsymbol{H}_0'^{-1} \right) \xrightarrow{d} \text{N} \left(\boldsymbol{0}, \lim_{N \to \infty} \text{Var} \left(\boldsymbol{A} \left(\alpha, \boldsymbol{\Lambda} \boldsymbol{H}_0'^{-1} \right) \right) \right)$. This completes the proof of (A.8).

Proof of (A.9) Let $\boldsymbol{a}$ be a 3-dimensional vector of real such that $\boldsymbol{a}'\boldsymbol{a} = 1$. We have from Assumption 3 (a) that

$$\text{E} \left(\frac{\sqrt{N}}{\alpha N} \sum_{i=1}^{\alpha N} \boldsymbol{a}' \left(\boldsymbol{\lambda}_i \boldsymbol{\lambda}_i' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}} \right) \boldsymbol{a} \right)^2 = \frac{1}{\alpha} \frac{1}{\alpha N} \sum_{i=1}^{\alpha N} \text{E} \left(\boldsymbol{a}' \left(\boldsymbol{\lambda}_i \boldsymbol{\lambda}_i' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}} \right) \boldsymbol{a} \right)^2 \leq \frac{1}{\alpha} \frac{1}{\alpha N} \sum_{i=1}^{\alpha N} \text{E} \left(\boldsymbol{a}' \left(\boldsymbol{\lambda}_i \boldsymbol{\lambda}_i' \right) \boldsymbol{a} \right)^2.$$

Given that $\text{E} \left(\boldsymbol{a}' \left(\boldsymbol{\lambda}_i \boldsymbol{\lambda}_i' \right) \boldsymbol{a} \right)^2 = \left(\text{E} \left(\boldsymbol{a}' \boldsymbol{\lambda}_i \right)^4 \right) \leq \text{E} \left\| \boldsymbol{\lambda}_i \right\|^4 \leq C$, we obtain $\text{E} \left(\frac{\sqrt{N}}{\alpha N} \sum_{i=1}^{\alpha N} \boldsymbol{a}' \left(\boldsymbol{\lambda}_i \boldsymbol{\lambda}_i' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}} \right) \boldsymbol{a} \right)^2 \leq \frac{1}{\alpha} C < \infty$. Thus, $\frac{\sqrt{N}}{\alpha N} \sum_{i=1}^{\alpha N} \boldsymbol{a}' \left(\boldsymbol{\lambda}_i \boldsymbol{\lambda}_i' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}} \right) \boldsymbol{a} = O_P(1)$, for any $\alpha \in [\alpha_1, \alpha_2]$. By similar arguments, it also holds that $\frac{\sqrt{N}}{N - \alpha N} \sum_{i=\alpha N+1}^{N} \left(\boldsymbol{\lambda}_i \boldsymbol{\lambda}_i' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}} \right) = O_P(1)$ for any $\alpha \in [\alpha_1, \alpha_2]$.

A.3 Proof of Lemma 2.1

We first note that the results in this section holds uniformly in α and we start with (8).

Proof of (8) A triangular inequality implies that $\left\| \boldsymbol{A} \left(\alpha, \tilde{\boldsymbol{\Lambda}} \right) - \boldsymbol{A} \left(\alpha, \boldsymbol{\Lambda} \boldsymbol{H}_0'^{-1} \right) \right\| \leq J_1 + J_2$, where $J_1 = \left\| \boldsymbol{A} \left(\alpha, \tilde{\boldsymbol{\Lambda}} \right) - \boldsymbol{A} \left(\alpha, \boldsymbol{\Lambda} \boldsymbol{H}'^{-1} \right) \right\|$ and $J_2 = \left\| \boldsymbol{A} \left(\alpha, \boldsymbol{\Lambda} \boldsymbol{H}'^{-1} \right) - \boldsymbol{A} \left(\alpha, \boldsymbol{\Lambda} \boldsymbol{H}_0'^{-1} \right) \right\|$. Applying again the triangular inequality, we have $J_1 \leq J_{11} + J_{12}$ and $J_2 \leq J_{21} + J_{22}$, with

$$\begin{aligned} J_{11} &= \sqrt{N}\left\|\text{Vech}\left(\frac{1}{\lfloor \alpha N \rfloor}\sum_{i=1}^{\lfloor \alpha N \rfloor}\left(\tilde{\boldsymbol{\lambda}}_i\tilde{\boldsymbol{\lambda}}_i' - \boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'\left(\boldsymbol{H}^{-1}\right)'\right)\right)\right\|, \\ J_{12} &= \sqrt{N}\left\|\text{Vech}\left(\frac{1}{N-\lfloor \alpha N \rfloor}\sum_{i=\lfloor \alpha N \rfloor+1}^{N}\left(\tilde{\boldsymbol{\lambda}}_i\tilde{\boldsymbol{\lambda}}_i' - \boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'\left(\boldsymbol{H}^{-1}\right)'\right)\right)\right\|, \\ J_{21} &= \sqrt{N}\left\|\text{Vech}\left(\left(\boldsymbol{H}^{-1} - \boldsymbol{H}_0^{-1}\right)\left(\frac{1}{\lfloor \alpha N \rfloor}\sum_{i=1}^{\lfloor \alpha N \rfloor}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i' - \frac{1}{N-\lfloor \alpha N \rfloor}\sum_{i=\lfloor \alpha N \rfloor+1}^{N}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'\right)\left(\boldsymbol{H}^{-1}\right)'\right)\right\| \end{aligned}$$

and

$$J_{22} = \sqrt{N}\left\|\text{Vech}\left(\boldsymbol{H}_0^{-1}\left(\frac{1}{\lfloor \alpha N \rfloor}\sum_{i=1}^{\lfloor \alpha N \rfloor}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i' - \frac{1}{N-\lfloor \alpha N \rfloor}\sum_{i=\lfloor \alpha N \rfloor+1}^{N}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'\right)\left(\boldsymbol{H}^{-1} - \boldsymbol{H}_0^{-1}\right)'\right)\right\|.$$

Therefore, to complete the proof of (8), we only need to show that $J_{11} = o_P(1)$, $J_{12} = o_P(1)$. $J_{21} = o_P(1)$ and $J_{22} = o_P(1)$. Note that $J_{12} = o_P(1)$ and $J_{22} = o_P(1)$ follow by identical steps to $J_{11} = o_P(1)$ and $J_{12} = o_P(1)$, and are omitted. From a triangular inequality, J_{11} is bounded by

$$\begin{aligned} & \sqrt{N}\left\|\frac{1}{\lfloor \alpha N \rfloor}\sum_{i=1}^{\lfloor \alpha N \rfloor}\left(\left(\tilde{\boldsymbol{\lambda}}_i - \boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\right)\left(\tilde{\boldsymbol{\lambda}}_i - \boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\right)' + \left(\tilde{\boldsymbol{\lambda}}_i - \boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\right)\boldsymbol{\lambda}_i'\left(\boldsymbol{H}^{-1}\right)' + \boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\left(\tilde{\boldsymbol{\lambda}}_i - \boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\right)'\right)\right\| \\ \leq & \sqrt{N}\left\|\frac{1}{\lfloor \alpha N \rfloor}\sum_{i=1}^{\lfloor \alpha N \rfloor}\left(\tilde{\boldsymbol{\lambda}}_i - \boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\right)\left(\tilde{\boldsymbol{\lambda}}_i - \boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\right)'\right\| + 2\sqrt{N}\left\|\frac{1}{\lfloor \alpha N \rfloor}\sum_{i=1}^{\lfloor \alpha N \rfloor}\boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\left(\tilde{\boldsymbol{\lambda}}_i - \boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\right)'\right\| \\ \leq & \sqrt{N}\frac{1}{\lfloor \alpha N \rfloor}\sum_{i=1}^{\lfloor \alpha N \rfloor}\left\|\tilde{\boldsymbol{\lambda}}_i - \boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\right\|^2 + 2\left\|\boldsymbol{H}^{-1}\right\|\sqrt{N}\left\|\frac{1}{\lfloor \alpha N \rfloor}\sum_{i=1}^{\lfloor \alpha N \rfloor}\boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\left(\tilde{\boldsymbol{\lambda}}_i - \boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\right)'\right\|. \end{aligned}$$

Thus $J_{11} = O_P\left(\sqrt{N}\delta_{NT}^{-2}\right) = o_P(1)$ by an application of (A.1), (A.3), $\boldsymbol{H}_0^{-1} = O_P(1)$ and $\sqrt{N}/T \to 0$. We also have for any $\alpha \in [\alpha_1, \alpha_2]$,

$$\begin{aligned} J_{21} &\leq \left\|\boldsymbol{H}^{-1} - \boldsymbol{H}_0^{-1}\right\|\left\|\frac{\sqrt{N}}{\lfloor \alpha N \rfloor}\sum_{i=1}^{\lfloor \alpha N \rfloor}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i' - \frac{\sqrt{N}}{N-\lfloor \alpha N \rfloor}\sum_{i=\lfloor \alpha N \rfloor+1}^{N}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'\right\|\left\|\boldsymbol{H}^{-1}\right\| \\ &\leq \left\|H^{-1} - H_0^{-1}\right\|\left(\left\|\frac{\sqrt{N}}{\lfloor \alpha N \rfloor}\sum_{i=1}^{\lfloor \alpha N \rfloor}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i' - \boldsymbol{\Sigma}_\Lambda\right\| + \left\|\frac{\sqrt{N}}{N-\lfloor \alpha N \rfloor}\sum_{i=\lfloor \alpha N \rfloor+1}^{N}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right\|\right)\left\|\boldsymbol{H}^{-1}\right\|, \end{aligned}$$

which is $o_P(1)$ using again $H^{-1} - H_0^{-1} = o_P(1)$, $\boldsymbol{H}_0^{-1} = O_P(1)$ and Equation (A.9).

Proof of (9) To show the consistency of the long run variance estimate $\tilde{\boldsymbol{S}}\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right)$ of $\boldsymbol{A}\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right)$, we need to prove that $\tilde{\boldsymbol{S}}\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right) - \boldsymbol{S}\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right) = o_P(1)$. Noting that

$$\tilde{\boldsymbol{S}}\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right) - \boldsymbol{S}\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right) = \left(\tilde{\boldsymbol{S}}\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right) - \boldsymbol{S}\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}'^{-1}\right)\right) + \left(\boldsymbol{S}\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}'^{-1}\right) - \boldsymbol{S}\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)\right), \tag{A.25}$$

the next steps will consist in showing that

$$\left(\tilde{\boldsymbol{S}}\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right) - \boldsymbol{S}\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}'^{-1}\right)\right) = o_P(1) \text{ and } \left(\boldsymbol{S}\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}'^{-1}\right) - \boldsymbol{S}\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)\right) = o_P(1).$$

Expanding $\tilde{\boldsymbol{S}}\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right)$ and $S\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}'^{-1}\right)$, respectively, using

$$\tilde{\boldsymbol{\lambda}}_i\tilde{\boldsymbol{\lambda}}_i' - \boldsymbol{V}_{NT} = \left(\tilde{\lambda}_i\tilde{\boldsymbol{\lambda}}_i' - \boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\boldsymbol{\lambda}\boldsymbol{H}^{-1\prime}\right) + \left(\boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'\boldsymbol{H}^{-1\prime} - \boldsymbol{V}_0\right) + \left(\boldsymbol{V}_0 - \boldsymbol{V}_{NT}\right),$$

$$\boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'\boldsymbol{H}^{-1\prime} - \boldsymbol{V}_0 = \left(\boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'\boldsymbol{H}^{-1\prime} - \tilde{\boldsymbol{\lambda}}_i\tilde{\boldsymbol{\lambda}}_i'\right) + \left(\tilde{\boldsymbol{\lambda}}_i\tilde{\boldsymbol{\lambda}}_i' - \boldsymbol{V}_0\right)$$

and making use of

$$\frac{1}{N}\sum_{i=1}^{N}\text{Vech}\left(\tilde{\boldsymbol{\lambda}}_i\tilde{\boldsymbol{\lambda}}_i' - \boldsymbol{V}_{NT}\right)\text{Vech}\left(\boldsymbol{V} - \boldsymbol{V}_{NT}\right)' = \text{Vech}\left(\frac{1}{N}\sum_{i=1}^{N}\tilde{\boldsymbol{\lambda}}_i\tilde{\boldsymbol{\lambda}}_i' - \boldsymbol{V}_{NT}\right)\text{Vech}\left(\boldsymbol{V} - \boldsymbol{V}_{NT}\right)' = \mathbf{0},$$

we obtain $\tilde{\boldsymbol{S}}\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right) - \boldsymbol{S}\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}'^{-1}\right) = \left(\frac{1}{\alpha} + \frac{1}{1-\alpha}\right)\left(\boldsymbol{C}_1 - \boldsymbol{C}_2 + \boldsymbol{C}_3\right)$, where

$$\begin{aligned}
\boldsymbol{C}_1 &= \frac{1}{N}\sum_{i=1}^{N}\text{Vech}\left(\tilde{\boldsymbol{\lambda}}_i\tilde{\boldsymbol{\lambda}}_i' - \boldsymbol{V}_{NT}\right)\text{Vech}\left(\tilde{\boldsymbol{\lambda}}_i\tilde{\boldsymbol{\lambda}}_i' - \boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'\boldsymbol{H}'^{-1}\right)', \\
\boldsymbol{C}_2 &= \frac{1}{N}\sum_{i=1}^{N}\text{Vech}\left(\boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'\boldsymbol{H}'^{-1} - \tilde{\boldsymbol{\lambda}}_i\tilde{\boldsymbol{\lambda}}_i'\right)\text{Vech}\left(\boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'\boldsymbol{H}'^{-1} - \boldsymbol{V}_0\right)', \\
\boldsymbol{C}_3 &= \frac{1}{N}\sum_{i=1}^{N}\text{Vech}\left(\boldsymbol{V}_0 - \boldsymbol{V}_{NT}\right)\text{Vech}\left(\boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'\boldsymbol{H}'^{-1} - \boldsymbol{V}_0\right)'.
\end{aligned}$$

Consequently from an application of the Cauchy-Schwarz inequality and the c_r inequality,

$$\left\|\tilde{\boldsymbol{S}}\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right) - \boldsymbol{S}\left(\alpha, \boldsymbol{H}^{-1}\boldsymbol{\Lambda}\right)\right\| \leq \left(\frac{1}{\alpha} + \frac{1}{1-\alpha}\right)\left(\|\boldsymbol{C}_1\| + \|\boldsymbol{C}_2\| + \|\boldsymbol{C}_3\|\right),$$

where we have the bounds

$$\begin{aligned}
\|\boldsymbol{C}_1\| &\leq 2^{1/2}\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\tilde{\boldsymbol{\lambda}}_i\tilde{\boldsymbol{\lambda}}_i'\right\|^2 + \|\boldsymbol{V}_{NT}\|^2\right)^{1/2}\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\tilde{\boldsymbol{\lambda}}_i\tilde{\boldsymbol{\lambda}}_i' - \boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'\boldsymbol{H}'^{-1}\right\|^2\right)^{1/2} = O_P\left(\sqrt{N}/T\right), \\
\|\boldsymbol{C}_2\| &\leq 2^{1/2}\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'\boldsymbol{H}'^{-1}\right\|^2 + \|\boldsymbol{V}_0\|^2\right)^{1/2}\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\tilde{\boldsymbol{\lambda}}_i\tilde{\boldsymbol{\lambda}}_i' - \boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'\boldsymbol{H}'^{-1}\right\|^2\right)^{1/2} = O_P\left(\sqrt{N}/T\right),
\end{aligned}$$

and

$$\|\boldsymbol{C}_3\| \leq \left(\frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\right\|^2 + \|\boldsymbol{V}_0\|\right)\|\boldsymbol{V}_{NT} - \boldsymbol{V}_0\| = O_P\left(\sqrt{N}/T\right)$$

from $\boldsymbol{V}_{NT} \xrightarrow{P} \boldsymbol{V}_0 = O_P(1)$, (A.5) which gives $\frac{1}{N}\sum_{i=1}^{N}\left\|\tilde{\boldsymbol{\lambda}}_i\tilde{\boldsymbol{\lambda}}_i' - \boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'\boldsymbol{H}'^{-1}\right\|^2 = O_P\left(\sqrt{N}/T\right)$, the fact that $\frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'\boldsymbol{H}'^{-1}\right\|^2 \leq \left\|\boldsymbol{H}^{-1}\right\|^4\frac{1}{N}\sum_{i=1}^{N}\|\boldsymbol{\lambda}_i\|^4 = O_P(1)$ and $\frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\right\|^2 = \left\|\boldsymbol{H}^{-1}\right\|^2\frac{1}{N}\sum_{i=1}^{N}\|\boldsymbol{\lambda}_i\|^2 = O_P(1)$. Note that by the c_r inequality, we also have

$$\frac{1}{N}\sum_{i=1}^{N}\left\|\tilde{\boldsymbol{\lambda}}_i\tilde{\boldsymbol{\lambda}}_i'\right\|^2 \leq 2\frac{1}{N}\sum_{i=1}^{N}\left\|\tilde{\boldsymbol{\lambda}}_i\tilde{\boldsymbol{\lambda}}_i' - \boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'\boldsymbol{H}'^{-1}\right\|^2 + 2\frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'\boldsymbol{H}'^{-1}\right\|^2 = O_P(1).$$

In consequence, $\tilde{\boldsymbol{S}}\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right) - \boldsymbol{S}\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}'^{-1}\right) = O_P\left(\sqrt{N}/T\right) = o_P(1)$ as $\sqrt{N}/T \to 0$. To show that $\boldsymbol{S}\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}'^{-1}\right) - \boldsymbol{S}\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right) = o_P(1)$, we combine the Cauchy-Schwarz inequality and the c_r inequality to have that $\left\|\boldsymbol{S}\left(\alpha, \boldsymbol{H}^{-1}\boldsymbol{\Lambda}\right) - \boldsymbol{S}\left(\alpha, \boldsymbol{H}_0^{-1}\boldsymbol{\Lambda}\right)\right\|$ is lower than

$$2^{1/2}\left(\frac{1}{\alpha}+\frac{1}{1-\alpha}\right)\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'\boldsymbol{H}'^{-1} - \boldsymbol{H}_0^{-1}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'\boldsymbol{H}_0'^{-1}\right\|^2\right)^{1/2}\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\right\|^4 + \|\boldsymbol{V}_0\|^4\right)^{1/2}$$
$$+2^{1/2}\left(\frac{1}{\alpha}+\frac{1}{1-\alpha}\right)\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'\boldsymbol{H}'^{-1} - \boldsymbol{H}_0^{-1}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'\boldsymbol{H}_0'^{-1}\right\|^2\right)^{1/2}\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{H}_0^{-1}\boldsymbol{\lambda}_i\right\|^4 + \|\boldsymbol{V}_0\|^4\right)^{1/2}.$$

Since we know in particular from (A.6) that $\frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'\boldsymbol{H}'^{-1} - \boldsymbol{H}_0^{-1}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'\boldsymbol{H}_0'^{-1}\right\|^2 = o_P(1)$, and also that $\frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{H}^{-1}\boldsymbol{\lambda}_i\right\|^4 \leq \left\|\boldsymbol{H}^{-1}\right\|^4 \frac{1}{N}\sum_{i=1}^{N}\|\boldsymbol{\lambda}_i\|^4 = O_P(1)$, we finally obtain the second needed result, which is $\left\|\boldsymbol{S}\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}'^{-1}\right) - \boldsymbol{S}\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)\right\| = o_P(1)$.

A.4 Proof of Theorem 1

We begin by proving that $LM_N\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right) - LM_N\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right) = o_P(1)$. Using the triangular inequality, we have $\left|LM_N\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right) - LM_N\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)\right| \leq M_1 + M_2 + M_3$, where uniformly in α,

$$\begin{aligned} M_1 &= \left\|\boldsymbol{A}\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right)'\right\|\left\|\tilde{\boldsymbol{S}}\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right)^{-1} - \boldsymbol{S}\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)^{-1}\right\|\left\|\boldsymbol{A}\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right)\right\| \\ M_2 &= \left\|\boldsymbol{A}\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right) - \boldsymbol{A}\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)\right\|\left\|\boldsymbol{S}\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)^{-1}\right\|\left\|\boldsymbol{A}\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right)\right\| \end{aligned}$$

and

$$M_3 = \left\|\boldsymbol{A}\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)\right\|\left\|\boldsymbol{S}\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)^{-1}\right\|\left\|\boldsymbol{A}\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right) - \boldsymbol{A}\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)\right\|.$$

The proof uses the auxiliary results that (a) $\boldsymbol{A}\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right) - \boldsymbol{A}\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right) = o_P(1)$, (b) $\boldsymbol{A}\left(\alpha, \boldsymbol{\Lambda}\left(\boldsymbol{H}_0'\right)^{-1}\right) = O_P(1)$ and $\boldsymbol{A}\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right) = O_P(1)$ and (c) $\left\|\tilde{\boldsymbol{S}}\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right)^{-1} - \boldsymbol{S}\left(\alpha, \boldsymbol{\Lambda}\left(\boldsymbol{H}_0'\right)^{-1}\right)^{-1}\right\| = o_P(1)$ uniformly in α. We note that (a) follows from (8). Second, note that $\left\|\boldsymbol{A}\left(\alpha, \boldsymbol{\Lambda}\left(\boldsymbol{H}_0'\right)^{-1}\right)\right\|$ is equal to

$$\left\|\sqrt{N}\boldsymbol{H}_0^{-1}\left(\frac{1}{\lfloor\alpha N\rfloor}\sum_{i=1}^{\lfloor\alpha N\rfloor}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i' - \frac{1}{N-\lfloor\alpha N\rfloor}\sum_{i=\lfloor\alpha N\rfloor+1}^{N}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'\right)\boldsymbol{H}_0'^{-1}\right\| = O_P(1)$$

uniformly in α given Lemma A.2 (A.8). We also have uniformly in α, that

$$\left\|\boldsymbol{A}\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right)\right\| \leq \left\|\boldsymbol{A}\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right) - A\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)\right\| + \left\|\boldsymbol{A}\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)\right\| = o_P(1) + O_P(1) = O_P(1).$$

Thus (b) holds. Third, $\left\|\tilde{\boldsymbol{S}}\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right)^{-1} - \boldsymbol{S}\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)^{-1}\right\|$ is dominated by

$$\begin{aligned} &\left\|\boldsymbol{S}\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)^{-1}\left(\boldsymbol{S}\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right) - \tilde{\boldsymbol{S}}\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right)\right)\tilde{\boldsymbol{S}}\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right)^{-1}\right\| \\ \leq\ &\left\|\tilde{\boldsymbol{S}}\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)^{-1}\right\|\left\|\tilde{\boldsymbol{S}}\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right) - \boldsymbol{S}\left(\alpha, \boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)\right\|\left\|\boldsymbol{S}\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right)^{-1}\right\|, \end{aligned}$$

where $\tilde{\boldsymbol{S}}\left(\alpha,\tilde{\boldsymbol{\Lambda}}\right)-\boldsymbol{S}\left(\alpha,\boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)\xrightarrow{P}\mathbf{0}$ by Lemma 2.1 (9) and the limit in probability of $\boldsymbol{S}\left(\alpha,\boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)$ is positive definite from Assumption 3 (d). Hence, $\left\|\tilde{\boldsymbol{S}}\left(\alpha,\tilde{\boldsymbol{\Lambda}}\right)-\boldsymbol{S}\left(\alpha,\boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)\right\|=o_P(1)$, $\left\|\tilde{\boldsymbol{S}}\left(\alpha,\tilde{\boldsymbol{\Lambda}}\right)^{-1}\right\|=O_P(1)$, $\left\|\tilde{\boldsymbol{S}}\left(\alpha,\boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)^{-1}\right\|=O_P(1)$. Thus, we can see that $\left\|\tilde{\boldsymbol{S}}\left(\alpha,\tilde{\boldsymbol{\Lambda}}\right)^{-1}-\boldsymbol{S}\left(\alpha,\boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)^{-1}\right\|=o_P(1)$ which is result (c). From (a), (b) and (c), $\left|LM_N\left(\alpha,\tilde{\boldsymbol{\Lambda}}\right)-LM_N\left(\alpha,\boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)\right|=o_P(1)$. Because $LM_N\left(\alpha,\boldsymbol{\Lambda}\boldsymbol{H}_0'^{-1}\right)\xrightarrow{d}\chi^2(3)$ from Lemma A.2 (A.8), the result follows.

A.5 Proof of Lemma A.3

The proof of this lemma relies on showing that the required conditions in Lemma A.1 and Lemma 2.1 (9) are satisfied in the context of the one-level representation of the two-level alternative. If in addition to Assumption 4 (a)−(g), the conditions (i) $\mathrm{E}\left\|\boldsymbol{\phi}_i\right\|^4\leq C$, (ii) $\mathrm{E}\left(\frac{1}{T}\sum_{t=1}^T\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^N\boldsymbol{\phi}_ie_{it}\iota_{mi}\right\|^2\right)\leq C$, $m=D,E$ where $\mathrm{E}\left(\boldsymbol{\phi}_ie_{it}\right)=\mathbf{0}$ for all (i,t) hold, then the needed assumptions are satisfied. First, given the definition of $\boldsymbol{\Phi}$, $\mathrm{E}\left\|\boldsymbol{\Phi}_i\right\|^4=\mathrm{E}\left\|\boldsymbol{\lambda}_i\right\|^4\leq C$ under Assumption 1 (b). Second, for $m=D$,

$$\mathrm{E}\left(\frac{1}{T}\sum_{t=1}^T\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^N\boldsymbol{\phi}_ie_{it}\iota_{mi}\right\|^2\right)=\mathrm{E}\left(\frac{1}{T}\sum_{t=1}^T\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N_1}\boldsymbol{\lambda}_ie_{it}\right\|^2\right)\leq C$$

and for $m=E$,

$$\mathrm{E}\left(\frac{1}{T}\sum_{t=1}^T\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^N\boldsymbol{\phi}_ie_{it}\iota_{mi}\right\|^2\right)=\mathrm{E}\left(\frac{1}{T}\sum_{t=1}^T\left\|\frac{1}{\sqrt{N}}\sum_{i=N_1+1}^{N}\boldsymbol{\lambda}_ie_{it}\right\|^2\right)\leq C,$$

using the definition of $\boldsymbol{\Phi}$ and Assumption 2 (d). Similarly,

$$\mathrm{E}\left(\boldsymbol{\phi}_ie_{it}\right)=\mathrm{E}\left(\left[\lambda_{0i}\ \lambda_{1i}\ 0\right]'e_{it}\right)=\mathbf{0}\text{ for }i=1,\ldots,N_1$$

and

$$\mathrm{E}\left(\boldsymbol{\phi}_ie_{it}\right)=\mathrm{E}\left(\left[\lambda_{0i}\ 0\ \lambda_{1i}\right]'e_{it}\right)=\mathbf{0}\text{ for }i=N_1+1,\ldots,N.$$

Hence, the analogue of Lemma A.1 follows with $\boldsymbol{\Xi}^{-1}$ and $\boldsymbol{\Xi}_0^{-1}$ in the place $\boldsymbol{H}^{-1}$ and $\boldsymbol{H}_0^{-1}$, respectively, using the same the step as in the proof of Lemma A.1. Furthermore, the analogue of Lemma 2.1 (9) also follows using similar steps and given the definition of $\boldsymbol{J}$ and $\boldsymbol{J}_0$ as submatrices of $\boldsymbol{\Xi}'^{-1}$ and $\boldsymbol{\Xi}_0'^{-1}$, respectively.

A.6 Proof of Theorem 2

Recall the matrix of factor loadings $\boldsymbol{\Phi}=\left[\boldsymbol{\phi}_1\cdots\boldsymbol{\phi}_N\right]':N\times 3$ of the one-level representation under the two-level alternative. From Bai (2003), the first three PCM estimates $\tilde{\boldsymbol{\Phi}}$ converge to a rotation $\boldsymbol{\Phi}\boldsymbol{\Xi}'^{-1}$ with $\boldsymbol{\Xi}$ the 3×3 rotation matrix. In particular, we can find a 3×2 submatrix J of $\boldsymbol{\Xi}'^{-1}$

such that $\tilde{\boldsymbol{\Lambda}} = \boldsymbol{\Phi}\boldsymbol{J} + O_P\left(\delta_{NT}^{-2}\right)$, with $\tilde{\boldsymbol{\Lambda}}$, the first two PCM estimates. Thus, we can write

$$\frac{1}{N_1}\sum_{i=1}^{N_1}\tilde{\boldsymbol{\lambda}}_i\tilde{\boldsymbol{\lambda}}_i' - \frac{1}{N-N_1}\sum_{i=N_1+1}^{N}\tilde{\boldsymbol{\lambda}}_i\tilde{\boldsymbol{\lambda}}_i' = \boldsymbol{R}_1 + \boldsymbol{R}_2 - \boldsymbol{R}_3, \tag{A.26}$$

where

$$\begin{aligned}
\boldsymbol{R}_1 &= \frac{1}{N_1}\sum_{i=1}^{N_1}\boldsymbol{J}'\boldsymbol{\phi}_i\boldsymbol{\phi}_i'\boldsymbol{J} - \frac{1}{N-N_1}\sum_{i=N_1+1}^{N}\boldsymbol{J}'\boldsymbol{\phi}_i\boldsymbol{\phi}_i'\boldsymbol{J} \\
\boldsymbol{R}_2 &= \frac{1}{N_1}\sum_{i=1}^{N_1}\tilde{\boldsymbol{\lambda}}_i\tilde{\boldsymbol{\lambda}}_i' - \frac{1}{N_1}\sum_{i=1}^{N_1}\boldsymbol{J}'\boldsymbol{\phi}_i\boldsymbol{\phi}_i'\boldsymbol{J} \\
\boldsymbol{R}_3 &= \frac{1}{N-N_1}\sum_{i=N_1+1}^{N}\tilde{\boldsymbol{\lambda}}_i\tilde{\boldsymbol{\lambda}}_i' - \frac{1}{N-N_1}\sum_{i=N_1+1}^{N}\boldsymbol{J}'\boldsymbol{\phi}_i\boldsymbol{\phi}_i'\boldsymbol{J}.
\end{aligned}$$

Given that

$$\boldsymbol{R}_2 = \frac{1}{N_1}\sum_{i=1}^{N_1}\left(\tilde{\boldsymbol{\lambda}}_i - \boldsymbol{J}'\boldsymbol{\phi}_i\right)\left(\tilde{\boldsymbol{\lambda}}_i - \boldsymbol{J},\boldsymbol{\phi}_i\right)' + \frac{1}{N_1}\sum_{i=1}^{N_1}\boldsymbol{J}'\boldsymbol{\phi}_i\left(\tilde{\boldsymbol{\lambda}}_i - \boldsymbol{J}'\boldsymbol{\phi}_i\right)' + \frac{1}{N_1}\sum_{i=1}^{N_1}\left(\tilde{\boldsymbol{\lambda}}_i - \boldsymbol{J}'\boldsymbol{\phi}_i\right)\boldsymbol{\phi}_i'\boldsymbol{J},$$

we deduce from Lemma A.3 that $\boldsymbol{R}_2 = O_P\left(\delta_{NT}^{-2}\right)$. Similarly, $\boldsymbol{R}_3 = O_P\left(\delta_{NT}^{-2}\right)$. Hence

$$\frac{1}{N_1}\sum_{i=1}^{N_1}\tilde{\boldsymbol{\lambda}}_i\tilde{\boldsymbol{\lambda}}_i' - \frac{1}{N-N_1}\sum_{i=N_1+1}^{N}\tilde{\boldsymbol{\lambda}}_i\tilde{\boldsymbol{\lambda}}_i' = \boldsymbol{R}_1 + O_P\left(\delta_{NT}^{-2}\right).$$

Noting that $\boldsymbol{\Xi}^{-1} = \boldsymbol{\Xi}_0^{-1} + o_P(1)$, and letting $\boldsymbol{J}_0$ the submatrix of $\boldsymbol{\Xi}_0^{-1}$ corresponding to the limit of $\boldsymbol{J}$. Using Assumption 3 (a) and the steps as in Lemma A.2 (A.9), $\frac{1}{N_1}\sum_{i=1}^{N_1}\boldsymbol{\phi}_i\boldsymbol{\phi}_i' = O_P(1)$ and $\frac{1}{N-N_1}\sum_{i=N_1+1}^{N}\boldsymbol{\phi}_i\boldsymbol{\phi}_i' = O_P(1)$. Consequently, $\boldsymbol{R}_1 = \boldsymbol{R}_0 + o_P(1)$, where

$$\boldsymbol{R}_0 = \boldsymbol{J}_0' \operatorname*{plim}_{N\to\infty}\left(\frac{1}{N_1}\sum_{i=1}^{N_1}\boldsymbol{\phi}_i\boldsymbol{\phi}_i' - \frac{1}{N-N_1}\sum_{i=N_1+1}^{N}\boldsymbol{\phi}_i\boldsymbol{\phi}_i'\right)\boldsymbol{J}_0.$$

Moreover, $\boldsymbol{R}_0 = \boldsymbol{J}_0'\begin{bmatrix} 0 & \Sigma_\Lambda^{12} & -\Sigma_\Lambda^{12} \\ \Sigma_\Lambda^{21} & \Sigma_\Lambda^{22} & 0 \\ -\Sigma_\Lambda^{21} & 0 & -\Sigma_\Lambda^{22}\end{bmatrix}\boldsymbol{J}_0$ is different from $\mathbf{0}$ since the second term is nonzero as $\Sigma_\Lambda^{22} \neq 0$ and the rows of $\boldsymbol{J}_0'$ are linearly independent. Hence

$$\frac{1}{N_1}\sum_{i=1}^{N_1}\tilde{\boldsymbol{\lambda}}_i\tilde{\boldsymbol{\lambda}}_i' - \frac{1}{N-N_1}\sum_{i=N_1+1}^{N}\tilde{\boldsymbol{\lambda}}_i\tilde{\boldsymbol{\lambda}}_i' = \boldsymbol{R}_0 + o_P(1), \tag{A.27}$$

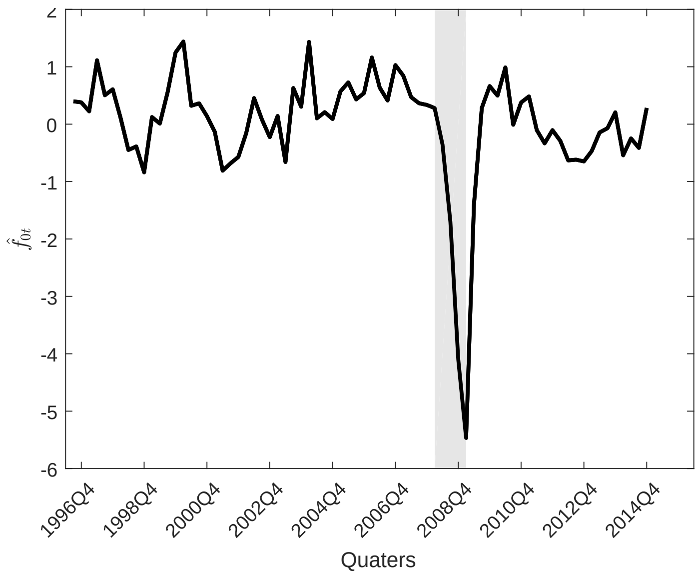
with $\boldsymbol{R}_0 \neq \mathbf{0}$. The second part of Theorem 2 follows from the fact that

$$\left\|\tilde{\boldsymbol{S}}\left(\alpha, \tilde{\boldsymbol{\Lambda}}\right) - \boldsymbol{S}\left(\alpha, \boldsymbol{\Phi}\boldsymbol{J}_0\right)\right\| = o_P(1). \tag{A.28}$$

given Lemma A.3, the positive definiteness of the limit in probability $\boldsymbol{S}_0$ of $\boldsymbol{S}(\alpha, \boldsymbol{\Phi}\boldsymbol{J}_0)$ given Assumption 4 (h) and an application of Equation (A.28).

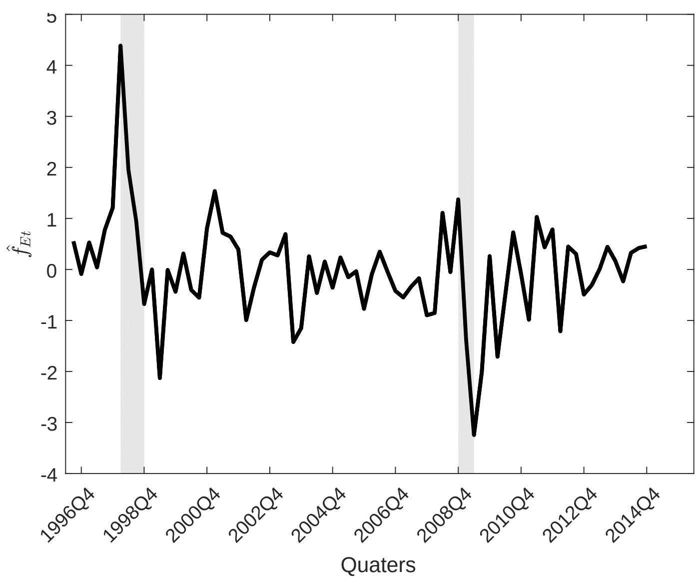
A.7 Figures: Economy Activity Factors

Figure 1: Global economy activity factor



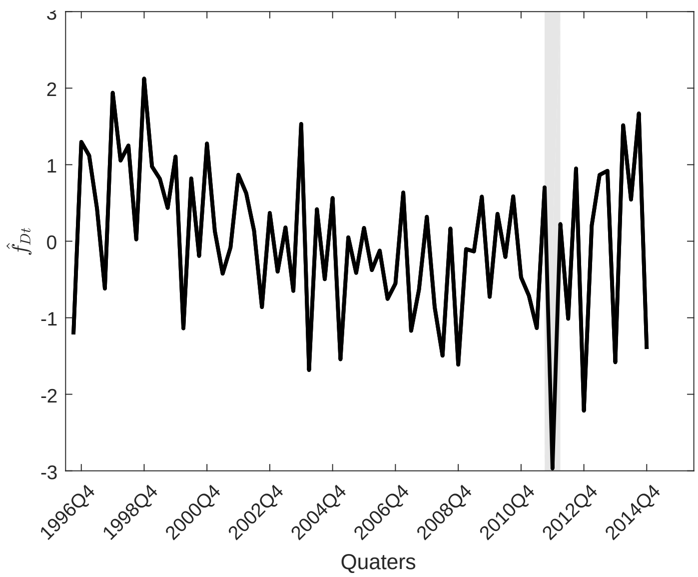
[g] This figure plots the estimated global economy activity factor over quarters.

Figure 2: Emerging economy activity factor



[h] This figure plots the estimated emerging economy activity factor over quarters.

Figure 3: Developed economy activity factor



[i] This figure plots the estimated developed economy activity factor over quarters.

References

Aastveit, K. A., H. C. Bjørnland, and L. A. Thorsrud (2015). What drives oil prices? emerging versus developed economies. *Journal of Applied Econometrics 30*, 1013–1028.

Bai, J. (2003). Inferential theory for factor models of large dimensions. *Econometrica 71*(1), 135–171.

Bai, J. and S. Ng (2002). Determining the number of factors in approximate factor models. *Econometrica 70*(1), 191–221.

Bai, J. and S. Ng (2004). Confidence intervals for diffusion index forecasts with a large number of predictor. Econometrics 0408006, EconWPA.

Bai, J. and S. Ng (2013). Principal component estimation and identificaton of static factor model. *Journal of Econometrics 176*(1), 18–29.

Breitung, J. and S. Eickmeier (2014). Analyzing business and financial cycles using multi-level factor models. Discussion Papers 11/2014, Deutsche Bundesbank, Research Centre.

Caldara, D., M. Cavallo, and M. Iacoviello (2016). Oil Price Elasticities and Oil Price Fluctuations. International Finance Discussion Papers 1173, Board of Governors of the Federal Reserve System (U.S.).

Charnavoki, V. and J. Dolado (2014). The effects of global shocks on small commodity-exporting economies: Lessons from canada. *American Economic Journal: Macroeconomics 6*(2), 207–37.

Djogbenou, A., S. Gonçalves, and B. Perron (2015). Bootstrap inference in regressions with estimated factors and serial correlation. *Journal of Time Series Analysis 36*(3), 481–502.

Djogbenou, A., J. G. MacKinnon, and M. Ø. Nielsen (2017). Validity of Wild Bootstrap Inference with Clustered Errors. Working Papers 1383, Queen's University, Department of Economics.

Djogbenou, A. A. (2017). Model Selection in Factor-Augmented Regressions with Estimated Factors. Working Papers 1391, Queen's University, Department of Economics.

Gonçalves, S. and B. Perron (2014). Bootstrapping factor-augmented regression models. *Journal of Econometrics 182*(1), 156–173.

Han, X. (2016). Shrinkage estimation of factor models with global and group-specific factors. Working papers, Department of Economics and Finance, City University of Hong Kong.

Han, X. and A. Inoue (2015). Tests For Parameter Instability In Dynamic Factor Models. *Econometric Theory 31*(05), 1117–1152.

Kose, M. A., C. Otrok, and C. H. Whiteman (2003). International Business Cycles: World, Region, and Country-Specific Factors. *American Economic Review 93*(4), 1216–1239.

Kose, M. A., C. M. Otrok, and E. S. Prasad (2008). Understanding the evolution of world business cycles. *Journal of International Economics 75*(1), 110–130.

Kose, M. A., C. M. Otrok, and E. S. Prasad (2012). Global business cycles: convergence or decoupling? *International Economic Review 53*(2), 511–538.

Ludvigson, S. and S. Ng (2007). The empirical risk-return relation: A factor analysis approach. *Journal of Financial Economics 83*(1), 171–222.

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

Wang, P. (2010). Large Dimensional Factor Models with a Multi-Level Factor Structure: Identification, Estimation and Inference. Working papers, Department of Economics, Hong Kong University of Science and Technology.