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

Antoine A. Djogbenou†
York University
daa@yorku.ca

June 30, 2019

Summary

Although globalization has shaped the world economy in recent decades, emerging economies have experienced impressive growth compared to developed economies, suggesting specific comovements within developed and emerging business cycles. Using observed developed and emerging real 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 global components, emerging or developed common components, 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 Monte Carlo simulations that document the testing procedure. An application of the test to various datasets of developed and emerging countries leads to strong statistical evidence of specific comovements within these two groups.

Keywords: statistical test, latent factors, specific comovements, 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 for bridging various types of information related to economic agents. Factor models generally assume a one-level structure, where the comovements within a large panel

*I am grateful to three anonymous referees, Sílvia Gonçalves, Christian Gouriéroux, Joann Jasiak, James G. MacKinnon, Morten Ø. Nielsen, Benoit Perron, Dalibor Stevanovic, Razvan Sufana, and participants at the Queen's University Economics Department Quantitative Workshop, the New York Camp Econometrics XIII, the 5th International Association for Applied Econometrics Annual Conference, the 52th Annual Conference of the Canadian Economics Association, the 2018 Africa Meeting of the Econometric Society, and the 35th Meeting of the Canadian Econometric Study Group, for useful discussions and comments. I thank Queen's University Economics Department for financial support.

†Address: Department of Economics, 4700 Keele Street, York University, Toronto, Ontario M3J 1P3, Canada. Email: daa@yorku.ca.

of variables can be summarized into a few latent factors affecting all variables. In particular, each variable in the large panel can decomposed into an idiosyncratic error component and a common component. See, for example, Stock and Watson (2002), Bai and Ng (2002), and Bai (2003) for details. In many prediction and policy exercises, empirical researchers have found it useful to extract factors from a large set of series. Among others, Ludvigson and Ng (2007) investigated the risk-return relation in the U.S. 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 that 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 economies 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 developed and emerging economy activity factors. Kose, Otrok, and Prasad (2012) investigated the decoupling between developed and emerging economy activity factors, using these models. Their motivation was twofold. First, the global economy has become interconnected through a large 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 real activity factors, the second suggests specific emerging economy activity factors different from those in developed economies.

As is well-known, the PCM estimates of factors 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) suggested a shrinkage estimator as an alternative and used it to disentangle global macroeconomic factors that are Europe specific and U.S. specific. He found that these specific estimated factors have high explanatory power for the leading economic indicators in Europe and the United States. Furthermore, Wang (2010) studied the estimation of multi-level factor models. In his empirical application, he decomposed the comovements within real and financial sectors in the U.S. economy.

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 developed and emerging economies. It is crucial to check whether the data is mainly driven by global factors, or whether factors specific to developed and emerging economies arise. In this paper, we propose a test, which we justify to be theoretically valid, and document its finite sample performance through simulation experiments.

Although the theoretical results can be applied in other contexts, this work focuses exclusively on international business cycle developments. It provides a statistical framework that can be used to analyze whether developed and emerging economy business cycles decouple. We propose a statistic that formally tests the one-level factor model specification against the two-level factor model specification with developed and emerging activity factors. In the empirical application, we find strong statistical evidence against the null hypothesis that developed and emerging business cycles emerge.

The rest of the paper is organized as follows. In Section 2, we present the two-level factor model, discuss the limit of the usual PCM for estimating specific factors, and propose a test of the one-level representation without specific comovements within developed and emerging economies against the two-level representation with specific comovements within developed and emerging economies. In Section 3, we investigate the finite sample properties of the proposed test. In Section 4, we apply the test to various datasets of developed and emerging economies real activity variables. In Section 5, we conclude. Proofs are relegated to the appendix. Throughout the paper, C, $\boldsymbol{I}$, $\lfloor \cdot \rfloor$, and $\|\cdot\|$ denote a generic finite constant, a generic identity matrix, the integer part of a number, and the Euclidean norm, respectively. When $\boldsymbol{M}$ is a matrix, $\boldsymbol{M} > \mathbf{0}$ means that $\boldsymbol{M}$ is positive definite.

2 Global, Emerging, and Developed Economy Activity Factors

Emerging economies have become major players in the global economy. On one hand, they have had high economic growth in recent decades compared to many 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 (Kose, Otrok, and Prasad, 2012). This paper contributes to the debate on the existence of specific comovements within the real activity of developed and emerging economies. 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 this study, we consider a large set of N economic activity variables $\left(\boldsymbol{X} = \{X_{it}\}_{t=1,\dots,T;\ i=1,\dots,N}\right)$ for 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 captured by latent global, developed, and emerging economy activity factors, which are denoted by $\boldsymbol{f}_{0t} : r_0 \times 1$, $\boldsymbol{f}_{Dt} : r_D \times 1$, and $\boldsymbol{f}_{Et} : r_E \times 1$, respectively. Therefore, more than one factor for developed or emerging countries is allowed. Further, given the heterogeneity within developed and emerging countries, it may happen that they are driven by a different number of specific factors. 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 allows us to identify some common factors that capture comovements across the entire dataset or across subsets of the series. In this case, the global factors $\boldsymbol{f}_{0t}$ reflect fluctuations that are common across all variables and countries. The developed economy activity factors $\boldsymbol{f}_{Dt}$ and the emerging economy activity factors $\boldsymbol{f}_{Et}$ capture fluctuations that are common to developed and emerging economies, respectively. The factor panel model can be written as

$$X_{it} = \boldsymbol{\lambda}_{0i}'\boldsymbol{f}_{0t} + \boldsymbol{\lambda}_{Di}'\boldsymbol{f}_{Dt} + e_{it} \text{ if } i = 1, \ldots, N_1, t = 1, \ldots, T, \tag{1}$$

and

$$X_{it} = \boldsymbol{\lambda}_{0i}'\boldsymbol{f}_{0t} + \boldsymbol{\lambda}_{Ei}'\boldsymbol{f}_{Et} + e_{it} \text{ if } i = N_1 + 1, \ldots, N, t = 1, \ldots, T, \tag{2}$$

where e_{it} are the idiosyncratic errors. The vector of factor loadings $\boldsymbol{\lambda}_{0i}$ measures the exposure of variable i to the global economic activity factors, whereas $\boldsymbol{\lambda}_{Di}$ and $\boldsymbol{\lambda}_{Ei}$ reflect the variable X_{it} exposure to developed and emerging country economy activity factors, respectively. The vectors of latent factors $\boldsymbol{f}_{Dt}$ and $\boldsymbol{f}_{Et}$ contain only information specific to each group of countries and not in the global economy activity factors $\boldsymbol{f}_{0t}$. Therefore, the global factors are allowed to affect all the variables, while the specific factors affect only the variables within their corresponding groups. In consequence, specificities like a change in the volatility within developed and emerging economies will be assumed to be generated by the specific factors $\boldsymbol{f}_{Dt}$ and $\boldsymbol{f}_{Et}$, respectively. Although, these specific factors are distinct, they are allowed to be correlated. When the distinction between the developed and emerging economy real activity factors is irrelevant, there is a factor model representation where only global factors matter, i.e.,

$$X_{it} = \boldsymbol{\lambda}_{0i}'\boldsymbol{f}_{0t} + e_{it}, i = 1, \ldots, N, t = 1, \ldots, T. \tag{3}$$

In the following subsection, we discuss the PCM estimation of the latent factors.

2.1 Principal Component Method and Specific Latent Factors

In practice, the real economic activity factors are latent, and must be estimated. When all latent factors are global, a popular approach consists of relying on principal component estimation. Since the seminal paper of Stock and Watson (2002), where the extracted factors were used in the forecasting context, this approach has received considerable attention in empirical and theoretical works.

When no specific factors emerge, the PCM estimates $\hat{\boldsymbol{F}} = \left[\hat{\boldsymbol{f}}_1 \cdots \hat{\boldsymbol{f}}_T\right]' : T \times r_0$ and $\hat{\boldsymbol{\Lambda}}_0 = \left[\hat{\boldsymbol{\lambda}}_{01} \cdots \hat{\boldsymbol{\lambda}}_{0N}\right]' : N \times r_0$ of the latent factor matrix $\boldsymbol{F}_0 = [\boldsymbol{f}_{01} \cdots \boldsymbol{f}_{0T}]'$ and the latent factor loading matrix $\boldsymbol{\Lambda} = [\boldsymbol{\lambda}_{01} \cdots \boldsymbol{\lambda}_{0N}]'$, respectively, are obtained by minimizing the sum of the squared idiosyncratic residuals under the restriction that $\hat{\boldsymbol{\Lambda}}'\hat{\boldsymbol{\Lambda}}/N = \boldsymbol{I}$. The resulting matrix of the estimated factor loadings corresponds to $\sqrt{N}$ times the eigenvectors corresponding to the r_0 largest eigenvalues of $\boldsymbol{X}\boldsymbol{X}'/(TN)$ in decreasing order. The matrix of the latent factors is estimated with $\hat{\boldsymbol{F}} = \boldsymbol{X}'\hat{\boldsymbol{\Lambda}}\left(\hat{\boldsymbol{\Lambda}}'\hat{\boldsymbol{\Lambda}}\right)^{-1} = \boldsymbol{X}'\hat{\boldsymbol{\Lambda}}/N$ (Bai and Ng, 2008). Alternatively, the PCM estimates $\tilde{\boldsymbol{F}} = \left[\tilde{\boldsymbol{f}}_1 \cdots \tilde{\boldsymbol{f}}_T\right]' : T \times r_0$ and $\tilde{\boldsymbol{\Lambda}} = \left[\tilde{\boldsymbol{\lambda}}_1 \cdots \tilde{\boldsymbol{\lambda}}_N\right]' : N \times r_0$ of the latent factors matrix $\boldsymbol{F}_0$ and the latent factor loading matrix $\boldsymbol{\Lambda}$, respectively, can be obtained by minimizing the sum of the squared idiosyncratic residuals under the restriction that $\tilde{\boldsymbol{F}}'\tilde{\boldsymbol{F}}/T = \boldsymbol{I}$. In this case, $\tilde{\boldsymbol{F}}$ is $\sqrt{T}$ times the eigenvectors corresponding to the r_0 largest eigenvalues of $\boldsymbol{X}'\boldsymbol{X}/(TN)$ in decreasing order, and $\tilde{\boldsymbol{\Lambda}} = \boldsymbol{X}\tilde{\boldsymbol{F}}/T$.

As is well-known, the PCM consistently estimates the space spanned by the true factors, i.e., $\hat{\boldsymbol{f}}_t = \hat{\boldsymbol{H}}'\boldsymbol{f}_{0t} + o_P(1)$, $t = 1, \ldots, T$, where we denote the rotation matrix by $\hat{\boldsymbol{H}} : r_0 \times r_0$. See also Bai (2003). Consequently, the vector of the estimated factor loadings $\hat{\boldsymbol{\lambda}}_i$ converges to a rotation of $\boldsymbol{\lambda}_i$, $\hat{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_i$, such that

$$X_{it} = \boldsymbol{\lambda}_{0i}'\boldsymbol{f}_{0t} + e_{it} = \left(\hat{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\right)'\left(\hat{\boldsymbol{H}}'\boldsymbol{f}_{0t}\right) + e_{it}. \tag{4}$$

Bai and Ng (2013) provided conditions that help identify the factors. However, these conditions rely on latent factors, and cannot be verified in practice. As a consequence, the principal component estimates jointly identify the true factor space.

In the presence of specific developed or emerging economy activity factors, we can write the one-level representation of Equation (1) and Equation (2) by

$$X_{it} = \boldsymbol{\phi}_{0i}'\boldsymbol{f}_{0t} + \boldsymbol{\phi}_{Di}'\boldsymbol{f}_{Dt} + \boldsymbol{\phi}_{Ei}'\boldsymbol{f}_{Et} + e_{it}, i = 1, \ldots, N, t = 1, \ldots, T, \tag{5}$$

where

$$\boldsymbol{\phi}_{0i} = \boldsymbol{\lambda}_{0i},\ \boldsymbol{\phi}_{Di} = \boldsymbol{\lambda}_{Di},\ \boldsymbol{\phi}_{Ei} = 0,\ i = 1, \ldots, N_1,$$
$$\boldsymbol{\phi}_{0i} = \boldsymbol{\lambda}_{0i},\ \boldsymbol{\phi}_{D} = 0,\ \boldsymbol{\phi}_{Ei} = \boldsymbol{\lambda}_{Ei},\ i = N_1 + 1, \ldots, N.$$

This implies that the real activity variables of developed economies have zero exposure to the real activity variables of emerging economies ($\boldsymbol{\phi}_{Ei} = 0, i = 1, \ldots, N_1$). Similarly, the real activity variables of emerging economies have zero exposure to the real activity variables of developed economies ($\boldsymbol{\phi}_{Di} = 0, i = N_1, \ldots, N$). Equivalently, Equation (5) can be written as

$$X_{it} = \boldsymbol{\phi}_i'\boldsymbol{g}_t + e_{it}, i = 1, \ldots, N, t = 1, \ldots, T, \tag{6}$$

where the vector of factors is $\boldsymbol{g}_t = [\boldsymbol{f}_{0t}'\ \boldsymbol{f}_{Dt}'\ \boldsymbol{f}_{Et}']' : r \times 1$, the vector of the factor loadings is $\boldsymbol{\phi}_i = [\boldsymbol{\phi}_{0t}'\ \boldsymbol{\phi}_{Dt}'\ \boldsymbol{\phi}_{Et}']' : r \times 1$, and $r = r_0 + r_D + r_E$.

Therefore, unless we know the linear relationship between the true latent factors and the estimated factors, comovements specific to emerging or developed economies cannot be identified directly applying the PCM. In practice, if one is interested only in simply forecasting a given variable using the PCM estimated factors, then it is sufficient to consider $(r_D + r_E)$ 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. Alternatively, Han (2016) proposed a shrinkage estimator to consistently identify the true model specification. This method estimates multi-level factors based on two adaptive group least absolute shrinkage and selection operator (LASSO) estimators using a penalty term. An alternative 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 the initial factor and factor loading estimates until convergence, and is easy to implement.

An important question that arises is whether there is statistical evidence for specific comovements within the real economic activity of developed and emerging economies. In the following subsection, we propose a statistical test that can be used to answer this question.

2.2 Testing for Specific Comovements

To investigate whether we can separately identify the specific real economic activity comovements, we develop a statistical test for the one-level factor model representation against the two-level one. Consider the matrix notation of the two-level alternative in Equation (5), where the factor loadings associated with developed countries are defined as

$$\boldsymbol{\Phi}_D = [\boldsymbol{\phi}_1 \cdots \boldsymbol{\phi}_{N_1}]' = \begin{bmatrix} \boldsymbol{\lambda}'_{01} & \boldsymbol{\lambda}'_{D1} & \mathbf{0} \\ \vdots & \vdots & \vdots \\ \boldsymbol{\lambda}'_{0N_1} & \boldsymbol{\lambda}'_{DN_1} & \mathbf{0} \end{bmatrix} : N_1 \times r.$$

The associated factor loadings with emerging economies are

$$\boldsymbol{\Phi}_E = [\boldsymbol{\phi}_{N_1+1} \cdots \boldsymbol{\phi}_N]' = \begin{bmatrix} \boldsymbol{\lambda}'_{0(N_1+1)} & \mathbf{0} & \boldsymbol{\lambda}'_{E(N_1+1)} \\ \vdots & \vdots & \vdots \\ \boldsymbol{\lambda}'_{0N} & \mathbf{0} & \boldsymbol{\lambda}'_{EN} \end{bmatrix} : (N - N_1) \times r.$$

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

$$\boldsymbol{X} = \boldsymbol{\Phi} G' + \boldsymbol{e}, \tag{7}$$

where $G = [\boldsymbol{g}_1 \cdots \boldsymbol{g}_T]' : T \times r$ and $\boldsymbol{\Phi} = \begin{bmatrix} \boldsymbol{\Phi}_D \\ \boldsymbol{\Phi}_E \end{bmatrix}$. This testing problem can be viewed as similar to the problem for testing the structural changes in factor loadings studied by Han and Inoue (2015), with the factor playing a similar role as the factor loadings. Using arguments that are close to theirs, we compare developed and emerging countries' subsample second moments of estimated factor loadings. The intuition behind the proposed statistical test can be understood by analyzing the difference in the infeasible subsample of factor loading second moments. We find that the limit of the difference of the subsample factor loading scaled products $\frac{1}{N_1}\boldsymbol{\Phi}_D'\boldsymbol{\Phi}_D - \frac{1}{N-N_1}\boldsymbol{\Phi}_E'\boldsymbol{\Phi}_E$ in the equivalent one-level representation Equation (7) of Equation (1) and Equation (2),

$$\begin{bmatrix} \frac{1}{N_1}\sum_{i=1}^{N_1}\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}' - \frac{1}{N-N_1}\sum_{i=N_1+1}^{N}\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}' & \frac{1}{N_1}\sum_{i=1}^{N_1}\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{Di}' & -\frac{1}{N-N_1}\sum_{i=N_1+1}^{N}\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{Ei}' \\ \frac{1}{N_1}\sum_{i=1}^{N_1}\boldsymbol{\lambda}_{Di}\boldsymbol{\lambda}_{0i}' & \frac{1}{N_1}\sum_{i=1}^{N_1}\boldsymbol{\lambda}_{Di}\boldsymbol{\lambda}_{Di}' & \mathbf{0} \\ -\frac{1}{N-N_1}\sum_{i=N_1+1}^{N}\boldsymbol{\lambda}_{Ei}\boldsymbol{\lambda}_{0i}' & \mathbf{0} & -\frac{1}{N-N_1}\sum_{i=N_1+1}^{N}\boldsymbol{\lambda}_{Ei}\boldsymbol{\lambda}_{Ei}' \end{bmatrix},$$

is asymptotically nonzero. Define $\boldsymbol{\Sigma}_{DD} = \mathrm{E}(\boldsymbol{\lambda}_{Di}\boldsymbol{\lambda}_{Di}')$, $\boldsymbol{\Sigma}_{EE} = \mathrm{E}(\boldsymbol{\lambda}_{Ei}\boldsymbol{\lambda}_{Ei}')$, $\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}} = \mathrm{E}(\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}')$, $\boldsymbol{\Sigma}_{0D} = \mathrm{E}(\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0D}')$, and $\boldsymbol{\Sigma}_{0E} = \mathrm{E}(\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0E}')$. Under the assumptions that $\boldsymbol{\lambda}_{Di}, \boldsymbol{\lambda}_{Ei}$, and $\boldsymbol{\lambda}_{0i}$ are independent across i, and have a finite fourth moment, we have

$$\frac{1}{N_1}\boldsymbol{\Phi}_D'\boldsymbol{\Phi}_D - \frac{1}{N-N_1}\boldsymbol{\Phi}_E'\boldsymbol{\Phi}_E = \begin{bmatrix} \mathbf{0} & \boldsymbol{\Sigma}_{0D} & -\boldsymbol{\Sigma}_{0E} \\ \boldsymbol{\Sigma}_{0D}' & \boldsymbol{\Sigma}_{DD} & \mathbf{0} \\ -\boldsymbol{\Sigma}_{0E}' & \mathbf{0} & -\boldsymbol{\Sigma}_{EE} \end{bmatrix} + o_P(1). \tag{8}$$

This limit is necessarily nonzero, because $\boldsymbol{\Sigma}_{DD}$ or $\boldsymbol{\Sigma}_{EE}$ is necessarily different from zero if specific developed or emerging real economic activity factors appear. Otherwise, the exposure of activity variables to specific factors would be zero. Thus, the scaled products of the factor loadings over the subsamples change. Therefore, $\sqrt{N}\left(\frac{1}{N_1}\boldsymbol{\Phi}_D'\boldsymbol{\Phi}_D - \frac{1}{N-N_1}\boldsymbol{\Phi}_E'\boldsymbol{\Phi}_E\right)$ diverges under the alternative.

In contrast, the difference in the second moments of the factor loadings over the subsamples is asymptotically a matrix with all elements being zero under the null hypothesis. Equation (3) is equivalent to

$$\boldsymbol{X} = \boldsymbol{\Lambda}\boldsymbol{F}_0' + \boldsymbol{e}, \tag{9}$$

where the matrix of the global factors and the associated matrices of the latent factor loadings are $\boldsymbol{F}_0 = [\boldsymbol{f}_{01}\cdots\boldsymbol{f}_{0T}]' : T \times r_0$ and $\boldsymbol{\Lambda} = \begin{bmatrix} \boldsymbol{\Lambda}_D \\ \boldsymbol{\Lambda}_E \end{bmatrix}$, with $\boldsymbol{\Lambda}_D = [\boldsymbol{\lambda}_{01}\cdots\boldsymbol{\lambda}_{0N_1}]' : N_1 \times r_0$, and $\boldsymbol{\Lambda}_E = \left[\boldsymbol{\lambda}_{0(N-N_1)}\cdots\boldsymbol{\lambda}_{0N}\right]' : (N-N_1) \times r_0$. In this case, the difference in the second moments of the factor loadings over the subsamples, $\frac{1}{N_1}\boldsymbol{\Lambda}_D'\boldsymbol{\Lambda}_D - \frac{1}{N-N_1}\boldsymbol{\Lambda}_E'\boldsymbol{\Lambda}_E = \mathbf{0}_{r_0\times r_0} + o_P(1) = o_P(1)$. Therefore, we propose a Lagrange multiplier-type statistic, which tests if the second moment of the factor loadings changes over developed and emerging economies.

2.2.1 The Statistical Test

Because the factors and the factor loadings are latent, they need to be estimated. Suppose r number of factors are selected using existing model selection criteria. See, for example, Bai and Ng (2002) or Onatski (2010). Under the null hypothesis, the principal component estimates $\hat{\boldsymbol{F}} : T \times r_0$ and $\hat{\boldsymbol{\Lambda}} : N \times r_0$ converge to their rotated versions $\boldsymbol{F}_0\hat{\boldsymbol{H}}$ and $\boldsymbol{\Lambda}\hat{\boldsymbol{H}}'^{-1}$, respectively, such that

$$\boldsymbol{X} = \boldsymbol{\Lambda}\boldsymbol{F}_0' + \boldsymbol{e} = \left(\boldsymbol{\Lambda}\hat{\boldsymbol{H}}'^{-1}\right)\left(\boldsymbol{F}_0\hat{\boldsymbol{H}}\right)' + \boldsymbol{e},$$

with $\hat{\boldsymbol{H}}$ an $r_0 \times r_0$ random matrix converging in probability to a nonsingular matrix $\check{\boldsymbol{H}}$. Therefore, we construct the test statistic based on the rotated latent factor loadings $\boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}$. We derive the asymptotic distribution of the $\frac{r_0(r_0+1)}{2}$-dimensional vectorized rotated latent difference, given by $\boldsymbol{A}\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right) = \text{Vech}\left(\sqrt{N}\check{\boldsymbol{H}}^{-1}\left(\frac{1}{N_1}\boldsymbol{\Lambda}_D'\boldsymbol{\Lambda}_D - \frac{1}{N-N_1}\boldsymbol{\Lambda}_E'\boldsymbol{\Lambda}_E\right)\check{\boldsymbol{H}}'^{-1}\right)$.

Under the independence and the existence of slightly more than fourth-moment conditions on $\boldsymbol{\lambda}_i$, we show in Lemma A.2 (Equation (A.11)) that

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

is asymptotically normal. Consequently, the infeasible statistic

$$LM_N\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right) = \boldsymbol{A}\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right)'\left(\boldsymbol{S}\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right)\right)^{-1}\boldsymbol{A}\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right) \xrightarrow{d} \chi^2\left(\frac{r_0(r_0+1)}{2}\right),$$

where

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

converges in probability to the variance of $\boldsymbol{A}\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right)$. See Lemma A.2 (Equation (A.10)) . Thus, we define the proposed statistic by

$$LM_N\left(\alpha, \hat{\boldsymbol{\Lambda}}\right) = \boldsymbol{A}\left(\alpha, \hat{\boldsymbol{\Lambda}}\right)'\left(\hat{\boldsymbol{S}}\left(\alpha, \hat{\boldsymbol{\Lambda}}\right)\right)^{-1}\boldsymbol{A}\left(\alpha, \hat{\boldsymbol{\Lambda}}\right), \tag{10}$$

where

$$\boldsymbol{A}\left(\alpha, \hat{\Lambda}\right) = \text{Vech}\left(\sqrt{N}\left(\frac{1}{N_1}\sum_{i=1}^{N_1}\hat{\boldsymbol{\lambda}}_i\hat{\boldsymbol{\lambda}}_i' - \frac{1}{N-N_1}\sum_{i=N_1+1}^{N}\hat{\boldsymbol{\lambda}}_i\hat{\boldsymbol{\lambda}}_i'\right)\right), \tag{11}$$

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

$$\boldsymbol{S}\left(\alpha, \hat{\boldsymbol{\Lambda}}\right) = \left(\frac{1}{\alpha} + \frac{1}{1-\alpha}\right)\frac{1}{N}\sum_{i=1}^{N}\text{Vech}\left(\hat{\boldsymbol{\lambda}}_i\hat{\boldsymbol{\lambda}}_i' - \boldsymbol{I}\right)\text{Vech}\left(\hat{\boldsymbol{\lambda}}_i\hat{\boldsymbol{\lambda}}_i' - \boldsymbol{I}\right)'. \tag{12}$$

However, under the two-level alternative, r, greater than r_0, factors are selected, and estimated. Using the fact that $\sqrt{N}\left(\frac{1}{N_1}\boldsymbol{\Phi}_D'\boldsymbol{\Phi}_D - \frac{1}{N-N_1}\boldsymbol{\Phi}_E'\boldsymbol{\Phi}_E\right)$ diverges under the alternative, we show that $A\left(\alpha, \hat{\Lambda}\right)$ diverges. Therefore, the test statistic, $LM_N\left(\alpha, \hat{\Lambda}\right)$, becomes unbounded, and leads to the rejection of the null hypothesis. In the following section, we present the asymptotic results.

2.2.2 Asymptotics under the Null Hypothesis

To obtain the limit distribution of the test statistic under the null hypothesis, the previous discussions are combined with Lemma 2.1. This lemma establishes that $\boldsymbol{A}\left(\alpha, \hat{\Lambda}\right)$ and $\boldsymbol{S}\left(\alpha, \hat{\Lambda}\right)$ are close enough to their respective infeasible analogue $\boldsymbol{A}\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right)$ and $\boldsymbol{S}\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right)$ uniformly in $\alpha \in [\alpha_1, \alpha_2] \subset (0, 1)$. To study the limiting null distribution of the proposed statistical test, we invoke the following assumptions of the approximate factor model, $X_{it} = \boldsymbol{\lambda}_{\mathbf{0}i}' \boldsymbol{f}_{0t} + e_{it}, i = 1, \ldots, N, t = 1, \ldots, T$.

Assumption 1. (Factor model and idiosyncratic errors)

(a) $\mathrm{E}\|\boldsymbol{f}_{0t}\|^4 \leq C$ and $\frac{1}{T}\boldsymbol{F}_0'\boldsymbol{F}_0 = \frac{1}{T}\sum_{t=1}^T \boldsymbol{f}_{0t}\boldsymbol{f}_{0t}' \xrightarrow{P} \boldsymbol{\Sigma_F} > 0$, where $\boldsymbol{\Sigma_F}$ is non-random.

(b) $\mathrm{E}\|\boldsymbol{\lambda}_{0i}\|^{4+\xi} \leq C$ for some $\xi > 0$ and the factor loadings $\{\boldsymbol{\lambda}_{0i}\}_{i=1,\ldots,N}$ are independent across i, $\mathrm{E}\left(\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'\right) = \boldsymbol{\Sigma_\Lambda} > 0$, where $\boldsymbol{\Sigma_\Lambda}$ is non-random.

(c) The eigenvalues of the $r_0 \times r_0$ matrix $(\boldsymbol{\Sigma_F} \times \boldsymbol{\Sigma_\Lambda})$ are distinct.

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

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

(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 $\{\boldsymbol{f}_{0t}\}$, $\{\boldsymbol{\lambda}_{0i}\}$, and $\{e_{it}\}$)

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

(b) For each t, $\mathrm{E}\left\|\frac{1}{\sqrt{TN}}\sum_{s=1}^T\sum_{i=1}^N \boldsymbol{f}_{0s}\left(e_{it}e_{is} - \mathrm{E}\left(e_{it}e_{is}\right)\right)\right\|^2 \leq C$.

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

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

(e) $\mathrm{E}\left(\left(\mathrm{Vech}\left(\check{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'\check{\boldsymbol{H}}'^{-1}\right)\right)'\mathrm{Vech}\left(\check{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'\check{\boldsymbol{H}}'^{-1}\right)\right) = \boldsymbol{\Sigma_{\Lambda\Lambda}}$ and the limit of $\mathrm{Var}\left(\boldsymbol{A}\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right)\right)$ is bounded and positive definite.

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 (2019). However, Assumption 2 (a), (c) and (d) which restrict the dependence between $\boldsymbol{f}_{0t}$, $\boldsymbol{\lambda}_{0i}$ and e_{it} among specific groups of variables are slightly stronger, and Assumption 1 (b) and Assumption 2 (e) are stronger and useful for deriving the asymptotic distribution of $\boldsymbol{A}\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right)$ using the central limit theorem.

Lemma 2.1. *Suppose that Assumptions 1 and 2 are satisfied. As $N, T \to \infty$, if $\sqrt{N}/T \to 0$, then it holds that under the null hypothesis,*

$$\left\| \boldsymbol{A}\left(\alpha, \hat{\boldsymbol{\Lambda}}\right) - \boldsymbol{A}\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right) \right\| = o_P\left(1\right), \tag{13}$$

and

$$\left\| \boldsymbol{S}\left(\alpha, \hat{\boldsymbol{\Lambda}}\right) - \boldsymbol{S}\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right) \right\| = o_P\left(1\right), \tag{14}$$

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

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

Theorem 1. *Suppose that Assumptions 1 and 2 are satisfied. As* $N, T \to \infty$*, if* $\sqrt{N}/T \to 0$*, then it holds that under the null hypothesis,*

$$LM_N\left(\alpha, \hat{\boldsymbol{\Lambda}}\right) \xrightarrow{d} \chi^2\left(\frac{r_0(r_0+1)}{2}\right),$$

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 $LM_N\left(\alpha, \hat{\boldsymbol{\Lambda}}\right)$ statistic, based on critical values from a chi-squared distribution with $\frac{r_0(r_0+1)}{2}$ degrees of freedom. As an alternate to the restricted 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 statistical test is asymptotically equivalent to the version suggested here, and in the simulation studies, we find that both statistics have similar size properties. Therefore, we focus on the simpler version presented here.

2.2.3 Asymptotics under the Alternative

The next theorem shows that the test statistic has power against the alternative hypothesis. To derive the result under the alternative, we make additional assumptions. We let $\iota_{Di} \equiv \mathrm{I}\left(i \leq \lfloor \alpha N \rfloor\right)$ and $\iota_{Ei} \equiv \mathrm{I}\left(i \geq \lfloor \alpha N \rfloor + 1\right)$, where $\mathrm{I}\left(\cdot\right)$ is an indicator function. Under the alternative, we recall that $\boldsymbol{G} = [\boldsymbol{g}_1 \cdots \boldsymbol{g}_T]' : T \times r$ and $\boldsymbol{\Phi} = [\boldsymbol{\phi}_1 \cdots \boldsymbol{\phi}_N]' : N \times r$. In this case, the PCM estimate of the factors, $\hat{\boldsymbol{F}}$, is a rotation the true latent factors $\boldsymbol{G}$. We denote $\check{\boldsymbol{\Xi}} : r \times r$ as the limit of the associated $r \times r$ rotation matrix $\hat{\boldsymbol{\Xi}}$ when the underlying data-generating process (DGP) has a two-level structure. Moreover, the PCM estimate of the factor loadings, $\hat{\boldsymbol{\Lambda}}$, is a rotation of the true latent factor loadings $\boldsymbol{\Phi}$.

Assumption 3. (Additional conditions for the two-level factor model)

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

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

(c) The eigenvalues of the $r \times r$ matrix $(\boldsymbol{\Sigma_G} \times \boldsymbol{\Sigma_\Phi})$ 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 $\mathrm{E}(\boldsymbol{g}_t e_{it}) = 0$ for every (i, t).

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

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

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

(h) The factor loadings $\{\boldsymbol{\phi}_i\}_{i=1,\ldots,N}$ are independent across i and $\mathrm{E}(\boldsymbol{\lambda}_{mi}\boldsymbol{\lambda}_{mi}') = \boldsymbol{\Sigma}_{mm}$, $m = D, E$.

(i) The limit $\boldsymbol{S}_0$ of $\boldsymbol{S}\left(\boldsymbol{A}\left(\alpha, \boldsymbol{\Phi}\check{\boldsymbol{\Xi}}'^{-1}\right)\right)$ is bounded and positive definite.

Assumption 3 (a)$-$(g) to complement Assumptions 1 and 2 as additional factors arise in the one-level representation of the two-level alternative hypothesis. Assumption 3 (h) allows distinct second moments for specific factor loadings. Assumption 3 (i) imposes positive definiteness on the limit of the variance estimator under the alternative.

Theorem 2. *Suppose that Assumptions 1–3 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 $\boldsymbol{R}_0 \neq \mathbf{0}$ such that*

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

and

$$LM_N\left(\alpha, \hat{\boldsymbol{\Lambda}}\right) = N\mathrm{Vech}(\boldsymbol{R}_0)'\,\boldsymbol{S}_0^{-1}\mathrm{Vech}(\boldsymbol{R}_0) + o_P(N),$$

uniformly in $\alpha \in [\alpha_1, \alpha_2]$, with $\boldsymbol{S}_0$ a constant matrix and $\mathrm{Vech}(\boldsymbol{R}_0)'\,\boldsymbol{S}_0^{-1}\mathrm{Vech}(\boldsymbol{R}_0) > 0$.

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

Algorithm for Implementing the Test Procedure.

1. Compute the estimated factors $(\hat{\boldsymbol{\Lambda}})$: $\sqrt{N}$ times the eigenvectors corresponding to the r largest eigenvalues of $\boldsymbol{X}\boldsymbol{X}'/(TN)$ in decreasing order and using the normalization $\hat{\boldsymbol{\Lambda}}'\hat{\boldsymbol{\Lambda}}/N = \boldsymbol{I}$.

2. Find the scaled difference between the estimated second moments

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

3. Compute $\boldsymbol{S}\left(\alpha, \hat{\boldsymbol{\Lambda}}\right)$, the variance estimator of $\boldsymbol{A}\left(\alpha, \hat{\boldsymbol{\Lambda}}\right)$ given by

$$\left(\frac{1}{\alpha}+\frac{1}{1-\alpha}\right)\frac{1}{N}\sum_{i=1}^{N}\text{Vech}\left(\hat{\boldsymbol{\lambda}}_i\hat{\boldsymbol{\lambda}}_i'-\boldsymbol{I}\right)\text{Vech}\left(\hat{\boldsymbol{\lambda}}_i\hat{\boldsymbol{\lambda}}_i'-\boldsymbol{I}\right)'.$$

4. Obtain the test statistic

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

5. Reject or do not using critical values or P-values from a $\chi^2\left(\frac{r(r+1)}{2}\right)$.

The number of factors r is greater than r_0 under the alternative hypothesis, but equal to r_0 under the null hypothesis. In practice, the appropriate number of factors can be selected using the criteria suggested by Bai and Ng (2002) or Onatski (2010).

2.3 Discussion of Some Special Situations

We now discuss a number of alternatives with developed and emerging specific heterogeneity ignored for the ease of exposition. The first alternative concerns cases where there is no global factor, $i.e., r_0 = 0$. The two-level alternative remains the same as in the one-level representation in Equation (7), with the number of $r = r_D + r_E$ factors and loading factors, such that the developed economies and the emerging economies real activity factor loadings are

$$\boldsymbol{\Phi}_D=[\boldsymbol{\phi}_1\cdots\boldsymbol{\phi}_{N_1}]'=\begin{bmatrix}\boldsymbol{\lambda}_{D1}' & \mathbf{0}\\ \vdots & \vdots\\ \boldsymbol{\lambda}_{DN_1}' & \mathbf{0}\end{bmatrix}:N_1\times r,$$

and

$$\boldsymbol{\Phi}_E=[\boldsymbol{\phi}_{N_1+1}\cdots\boldsymbol{\phi}_{N}]'=\begin{bmatrix}\mathbf{0} & \boldsymbol{\lambda}_{E(N_1+1)}'\\ \vdots & \vdots\\ \mathbf{0} & \boldsymbol{\lambda}_{EN}'\end{bmatrix}:(N-N_1)\times r$$

respectively. Further, $\frac{1}{N_1}\boldsymbol{\Phi}_D'\boldsymbol{\Phi}_D-\frac{1}{N-N_1}\boldsymbol{\Phi}_E'\boldsymbol{\Phi}_E=\begin{bmatrix}\boldsymbol{\Sigma}_{DD} & \mathbf{0}\\ \mathbf{0} & -\boldsymbol{\Sigma}_{EE}\end{bmatrix}+o_P(1)$, where $\boldsymbol{\Sigma}_{DD}\neq\mathbf{0}$ or $\boldsymbol{\Sigma}_{EE}\neq\mathbf{0}$, which ensures as described above, the divergence of the test statistic. This is a special case of the two-level model discussed previously. The results that our test statistic has power against such alternatives immediately follow from Theorem 2.

It may occur that there is no specific factor within one of the groups of countries ($r_D = 0$ or $r_E = 0$). As an illustration, suppose that is no specific factor within the developed economies, and $r = r_0 + r_E$. The factor loadings could be written as

$$\boldsymbol{\Phi}_D=[\boldsymbol{\phi}_1\cdots\boldsymbol{\phi}_{N_1}]'=\begin{bmatrix}\boldsymbol{\lambda}_{01}' & \mathbf{0}\\ \vdots & \vdots\\ \boldsymbol{\lambda}_{0N_1}' & \mathbf{0}\end{bmatrix}:N_1\times r$$

and

$$\boldsymbol{\Phi}_E = [\boldsymbol{\phi}_{N_1+1} \cdots \boldsymbol{\phi}_N]' = \begin{bmatrix} \boldsymbol{\lambda}'_{0(N_1+1)} & \boldsymbol{\lambda}'_{E(N_1+1)} \\ \vdots & \vdots \\ \boldsymbol{\lambda}'_{0(N_1+1)} & \boldsymbol{\lambda}'_{EN} \end{bmatrix} : (N - N_1) \times r.$$

Consequently, $\frac{1}{N_1}\boldsymbol{\Phi}'_D\boldsymbol{\Phi}_D - \frac{1}{N-N_1}\boldsymbol{\Phi}'_E\boldsymbol{\Phi}_E = \begin{bmatrix} \mathbf{0} & \boldsymbol{\Sigma}_{0E} \\ -\boldsymbol{\Sigma}'_{0E} & -\boldsymbol{\Sigma}_{EE} \end{bmatrix} + o_P(1)$, with $\boldsymbol{\Sigma}_{EE} \neq \mathbf{0}$, which helps to obtain the divergence of the proposed statistic. It is also a particular case of the setting in Section 2.2.3. Thus, the proposed test statistic has power against this type of alternative.

In all the situations discussed in this subsection, the suggested algorithm remains valid. First, the appropriate number of factors in the one-level representation of the two-level alternative is selected. Second, the emergence of specific developed and emerging real economic activity factors are tested by comparing the change in exposure to this factors from developed to emerging economies as argued above.

Although we focus on the statistical test of the existence of specific comovements within developed and emerging business cycles, these results can be applied to other contexts, and easily extended to a case with more than two specific groups, or with a multi-level structure, using arguments that are similar. These aspects are beyond the scope of this paper. The asymptotic results presented in Section 2.2.2 and Section 2.2.3 suggest that the proposed statistical test should have good control of the size and power, as the cross-sectional and time dimensions increase. In Section 3, we report simulation studies we used to assess its finite sample properties.

3 Simulation Experiments

We conduct Monte Carlo simulations based on six different DGPs, 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, to investigate the test size, while DGP 1-b, DGP 2-b, and DGP 3-b evaluate the power of the test statistic. The six DGPs are all based on modifications of the simulation designs in Han and Inoue (2015) work to incorporate two-level alternatives.

The first specification called DGP 1-a considers

$$X_{it} = \lambda_{0i} f_{0t} + \lambda_{1i} f_{1t} + \kappa e_{it}, \; i = 1, \ldots, N \text{ and } t = 1, \ldots, T, \tag{15}$$

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 $\kappa = \sqrt{2\,(1+c^2)}$ 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

$$e_{it} = \sigma_i \left(u_{it} + \sum_{1 \le |j| \le P} \theta u_{(i-j)t} \right), \; u_{it} \sim \text{NID}(0,1), \tag{16}$$

$$\sigma_i \sim U(0.5, 1.5) \text{ and } \kappa = \sqrt{\frac{24(1+c^2)}{13(1+2P\theta^2)}}, \tag{17}$$

where $\theta = 0.1$ and $P = 2$. 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

$$v_t \sim \text{NID}\left(0, 1-\rho_f^2\right), \; w_{it} \sim \text{NID}\left(0, 1-\rho_e^2\right) \text{ and } \sigma_i \sim \text{U}(0.5, 1.5), \tag{18}$$

where $\rho_e = 0.1$, $\rho_f = 0.1$, $\kappa = \sqrt{\frac{24(1+c^2)}{13}}$, and U is the uniform distribution. In all settings, we simulate the data $M = 10,000$ times, set $c = 1$, and use sample sizes (N, T) that belong to $\{50, 100\} \times \{25, 50, 75, 100, 125\}$. From Table 1, it follows that for different numbers of series and time periods, the sizes of the tests improves as N and T increase.

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

DGP 1-a	$T = 25$	$T = 50$	$T = 75$	$T = 100$	$T = 125$
$N = 50$	3.99	3.95	4.09	4.08	3.86
$N = 100$	4.70	4.30	4.61	4.99	4.63
DGP 2-a	$T = 25$	$T = 50$	$T = 75$	$T = 100$	$T = 125$
$N = 50$	5.66	5.42	4.91	5.02	5.06
$N = 100$	6.05	5.58	4.85	5.38	5.02
DGP 3-a	$T = 25$	$T = 50$	$T = 75$	$T = 100$	$T = 125$
$N = 50$	3.75	4.11	3.74	3.80	3.79
$N = 100$	4.80	4.08	4.69	4.47	4.88

[a] This table presents the rejection frequencies over 10,000 simulated data when there are no specific factors, and the significance level of the test is 5%.

DGP 1-a, DGP 2-a, and DGP 3-a are now modified to allow for two-level specification, and 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 $i \le \frac{N}{2}$, and by f_{Et} when $i \ge \frac{N}{2} + 1$, which suggests specific factors when $i \le \frac{N}{2}$ and $i \ge \frac{N}{2} + 1$. In DGP 1-b and DGP 2-b,

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

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 considered alternative hypothesis, because 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} to 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 $\rho = 0.5$.

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

DGP 1-b	$T = 25$	$T = 50$	$T = 75$	$T = 100$	$T = 125$
$N = 50$	76.44	89.24	95.27	9766	98.71
$N = 100$	87.45	98.19	99.79	99.96	100.00
DGP 2-b	$T = 25$	$T = 50$	$T = 75$	$T = 100$	$T = 125$
$N = 50$	84.61	95.01	98.13	99.16	99.66
$N = 100$	90.31	98.79	99.88	100.00	100.30
DGP 3-b	$T = 25$	$T = 50$	$T = 75$	$T = 100$	$T = 125$
$N = 50$	79.29	91.03	96.38	97.93	98.79
$N = 100$	88.37	98.05	99.76	99.98	100.00

[b] This table presents the rejection frequencies over $10,000$ simulated data when specific factors arise, and the significance level of the test is 5%.

In these three cases, the test rejection frequencies of the null hypothesis over the $10,000$ simulations are above 76% (see Table 2), which is a quiet high power. We also observe that the increase in the rejection frequencies is associated with an increase in the cross-sectional dimension.

We now modify DGP 3-b to investigate how an increase in the correlation between the two specific factors affects the power of the test. In this case, the correlation parameter ρ is set to $0.4, 0.6, 0.7$ and 0.8. Although the rejection frequencies decrease when the correlation between f_{Dt} and f_{Et} increases, we recover the power as the sample sizes increase. See Table 3. Overall, the simulation exercises for different DGPs exhibit good control of size and power, as the cross-sectional and time dimensions change.

4 Application of the Test

The high economic growth in emerging economies has led researchers to investigate the implication of specific comovements within developed and emerging economies for key economic variables,

Table 3: Rejection frequencies (%) for DGP 3-b when ρ changes

$\rho = 0.4$	$T = 25$	$T = 50$	$T = 75$	$T = 100$	$T = 125$
$N = 100$	87.94	96.62	99.05	99.59	99.87
$N = 200$	94.58	99.67	99.96	100.00	100.00
$\rho = 0.6$	$T = 25$	$T = 50$	$T = 75$	$T = 100$	$T = 125$
$N = 50$	64.60	76.84	85.39	90.00	92.96
$N = 100$	74.82	90.99	97.75	99.41	99.91
$\rho = 0.7$	$T = 25$	$T = 50$	$T = 75$	$T = 100$	$T = 125$
$N = 50$	43.94	49.97	57.16	63.37	67.94
$N = 100$	52.44	65.96	81.61	90.74	95.72
$\rho = 0.8$	$T = 25$	$T = 50$	$T = 75$	$T = 100$	$T = 125$
$N = 50$	22.27	19.86	19.92	20.49	21.13
$N = 100$	27.55	26.15	31.46	38.90	48.76

[c] See Table 2 [b]

for example, oil prices. For instance, Aastveit, Bjørnland, and Thorsrud (2015) examined how the heterogeneity within developed and emerging real economic activities contributes to make the price of oil fluctuate. In other papers, the interest has been to quantify the contribution of these specific comovements to the international real economic activity. See, for example, Kose, Otrok, and Prasad (2012). More recently, Hirata, Kose, and Otrok (2013) studied the implication of regionalization, facilitated by regional policies or regional integration agreements, for the emergence of specific comovements within international business cycles. Although most papers in the literature agree on the existence of specific factors within developed and emerging economies, few others have argued that the worldwide economy has been mainly driven by global factors (e.g., Flood and Rose (2010), Wälti (2012)). In this application section, we investigate whether we can find statistical evidence supporting developed and emerging specific heterogeneity in the international real economic activity. In other words, we answered the following question using various datasets: Are there specific comovements in the real activity within emerging economies that are different from those in the real activity within developed economies?

We construct a dataset similar to the one used by Aastveit, Bjørnland, and Thorsrud (2015). This dataset has the advantage of being quarterly, so that intra-annual international movements could be captured, and was downloaded from the Global Economic Monitor DataBank. The dataset covers the period from the third quarter of 1996 to the last quarter of 2018. We consider 22 developed and 33 emerging countries combining the classification in Aastveit, Bjørnland, and Thorsrud (2015) and Caldara, Cavallo, and Iacoviello (2016) and a recent classification by the International

Monetary Fund. The developed countries are Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, Luxembourg, the Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, the United Kingdom, and the United States. In addition, we have as emerging countries: Argentina, Brazil, Bulgaria, Chile, China, Costa Rica, Cyprus, Czech Republic, Estonia, Hong Kong, Hungary, India, Indonesia, Israel, Jordan, South Korea, Latvia, Lithuania, Malaysia, Mexico, Paraguay, Peru, Philippines, Poland, Romania, Russia, Singapore, Slovakia, Slovenia, South Africa, Taiwan, Thailand, and Turkey. This sample contains the major developed countries which are United States, United Kingdom, Canada, France, and Japan. The emerging countries include the BRICS countries composed of Brazil, Russia, India, China, and South Africa. We employ 88 series of the real gross domestic product (GDP) and industrial production including 22 series of real gross domestic product and 15 series of industrial production for developed countries. However, we use 30 series of gross domestic product and 21 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 and industrial production series for some of the countries.

Using this dataset, we obtain a P-value of 1.1927×10^{-5}. Thus, we conclude that there is strong statistical evidence against the null hypothesis that specific factors do not emerge within developed and emerging economies. This finding supports the assertion that the comovements in the real activity of emerging economies have become different from those of developed economies in recent decades. The finding is in line with that of Aastveit, Bjørnland, and Thorsrud (2015), who used a similar quarterly dataset. They found that there is heterogeneity in the fluctuation of developed and emerging economies' real activities, and the associated heterogeneous demand source are important in explaining the fluctuation of the price of oil. The present result suggests the existence of at least one specific factor within developed and emerging economies which contributes to making their real economic activity heterogeneous. We now connect the previous result to the existing literature on specific comovements within developed and emerging economies. To that end, we consider alternative datasets used in this literature on heterogeneous real economic activity, and apply the proposed testing procedure.

In a seminal paper on global business cycles, Kose, Otrok, and Prasad (2012) analyzed the evolution of the degree of global cyclical interdependence among developed economies and among emerging economies. The authors decomposed the real economic activity measured by output, consumption, and investment into a global factor, factors specific to country groups, and country-specific factors using a factor model. Comparing the magnitude of the variance decomposition of the macroeconomic variables explained by the developed and emerging specific factors, the authors found an increase in the relative importance in explaining the international real economic activity during the period $1985 - 2008$, and concluded convergence of business cycle fluctuations occurred among developed economies and among emerging economies.

We reconstruct their dataset[1] which includes the growth rates of output, consumption, and investment over the period $1961 - 2008$, and apply the proposed test statistic, to find statistical evidence of specific comovements. We follow Kose, Otrok, and Prasad (2012) and use real GDP, the real private consumption, and real fixed asset investment to measure the national output, consumption, and investment, respectively. To focus on the specific analysis of developed and emerging economies, we restrict our attention to the 23 developed economies and 24 emerging economies in Kose, Otrok, and Prasad (2012). Their data contains two more developed countries (Finland and Iceland), and excludes 14 (mostly recent) emerging economies (Bulgaria, Costa Rica, Cyprus, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Paraguay, Poland, Romania, Russia, Slovakia, and Slovenia) but includes others (Colombia, Egypt, Morocco, Pakistan, and Venezuela). To model the corresponding large dataset of real economic activity variables, we impose the two-level factor structure in Equation (1) and Equation (2). Although that specification does not include country-specific components, the proposed statistical test is derived under flexible enough assumptions on the idiosyncratic error to accommodate cross-correlated idiosyncratic components.

The computed P-value is 1.1102×10^{-16}, which suggests statistical evidence against the null hypothesis that no specific real economic activity factor emerges in developed and emerging economies. This evidence remains strong when the sampling period is extended to 2015, to incorporate the recent development in the international economy. The findings are in line with other works where the heterogeneity within developed and emerging business cycles is captured through economic regions.

Hirata, Kose, and Otrok (2013) employed a dynamic factor model to study the implications of the emergence of specific comovements for the evolution of global and regional business cycles. The authors showed that, there is an increase in economic region business cycles. Moreover, this regional cyclical dependence is stronger in regions where financial and trade flows have increased. Mumtaz, Simonelli, and Surico (2011) investigated the existence of comovements between national inflation rates and national real activities. Among other results, they found that economic regions have contributed considerably to the international business cycle heterogeneity. Breitung and Eickmeier (2014) revisited their analysis using the sequential least-squares and canonical correlation analyses and compared them to the two-step PCM and the quasi maximum-likelihood estimation, as well as the Bayesian approach for estimating the factors. Their results suggested that regionalization has increased, and supported the emergence of specificities in the international real economic activity. The obtained P-values based on the real economic activity variables in Hirata, Kose, and Otrok (2013) and Mumtaz, Simonelli, and Surico (2011) relying on the previous classification of countries

[1] This dataset was reconstructed based on the data used by Hirata, Kose, and Otrok (2013). The author is grateful to Hideaki Hirata and Ayhan M. Kose, who kindly provided their dataset.

confirmed the existence of specific comovements within developed and emerging economies.

Nevertheless, Flood and Rose (2010) established that there is limited empirical evidence suggesting the importance of specific factors within developed and emerging economies, known as decoupling. Wälti (2012) argued that there is a myth of decoupling. The paper finds that decoupling has not occurred over the periods from 1980 to 2007. Wälti (2012) contended that the increase in the trend of the growth rate within emerging economies during recent decades did not necessarily indicate “deviations from [the] trend have diverged between emerging markets and advanced economies.” To circumvent this issue, the author suggested the use of the output gap of each country in the sample as a measure of business cycles. First, Wälti (2012) inspected a measure of synchronicity based on the distance between developed and emerging economies output gap. Second, the author regressed the output gap of emerging economies on the aggregate output gap of developed economies. In both cases, that paper concluded that there was no evidence of decoupling, as the tests indicated statistically significant slope parameters, meaning that emerging and developed economies strongly comove. The business cycles were measured at an annual frequency, and the dataset covered the period from 1983 to 2008.

We also reconstruct Wälti’s dataset by computing the output gap are the detrended GDP using the Hodrick-Prescott filter, divided by the trend GDP. We collected the data from the World Economic Outlook database, and considered a sample of 26 developed countries and 30 emerging countries. The main difference with the sets of countries in the first application is that Cyprus, Israel, Hong Kong, and South Korea are treated as developed economies. In addition, Iceland was included in the group of developed countries, while Ecuador was in the group of emerging countries. Due to a large number of missing values, we dropped Czech Republic, Estonia, Latvia, Lithuania, Russia, Slovak Republic, Ukraine, and Venezuela while reconstructing this dataset. We find using the proposed test statistic a P-value of 2.5998×10^{-5} and conclude strong evidence of developed and emerging specific factors. Increasing the sample size up to 2017 makes this evidence stronger as we notice a drop in the P-value to 1.2529×10^{-5}.

The difference with the results in Wälti (2012) may be explained by the fact that the developed (emerging) economies output gap also includes the global real economic activity. Using these variables as regressors will not only capture the sensitivity to each group of economies. In this paper, we view business cycles forces as latent variables that are global or specific factors affecting countries’ real economic activity differently. The specific factors capture commonalities that are not already captured by the global factors. Whenever the impact of the specific factors on developed and emerging economies, captured through their loadings, are large enough, the test will hypothesis reject the null that no specific factor emerges. Overall, the statistical evidence for specific comovements within developed and emerging economies remains important, independently of whether economic activity is measured at a quarterly or yearly frequency, or detrended real activity variables are used or not.

5 Conclusion

In this paper, we investigated the existence of specific real economic activity factors within emerging economies, different from the ones within developed economies. The paper proposed a statistic to test the ability to identify specific economy activity factors from a large panel of economic activity variables. We showed the validity of the proposed testing procedure, provided evidence for its finite sample performance through Monte Carlo experiments, and applied the test to various datasets. In the empirical application, we found strong statistical evidence against the null hypothesis that no developed or emerging economy activity factors emerged.

The proposed statistical test 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- and high-income country risks matter as well. This 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 explaining the fluctuation in the price of oil.

Appendix A: Proofs

Throughout this appendix, we let $\delta_{NT} = \min\left[\sqrt{N}, \sqrt{T}\right]$. Note again that under the null, $\boldsymbol{\Lambda} = [\boldsymbol{\lambda}_1 \cdots \boldsymbol{\lambda}_N]' : N \times r_0$ and $\boldsymbol{F}_0 = [\boldsymbol{f}_{01} \cdots \boldsymbol{f}_{0T}]' : T \times r_0$. The PCM estimates are obtained by minimizing the sum of the squared idiosyncratic residuals under the identification condition that $\hat{\boldsymbol{\Lambda}}'\hat{\boldsymbol{\Lambda}}/N = \boldsymbol{I}$. These estimates are denoted by $\hat{\boldsymbol{F}} = \left[\hat{\boldsymbol{f}}_1 \cdots \hat{\boldsymbol{f}}_T\right]'$ and $\hat{\boldsymbol{\Lambda}} = \left[\hat{\boldsymbol{\lambda}}_1 \cdots \hat{\boldsymbol{\lambda}}_N\right]'$. From Bai and Ng (2008, Equation 6),

$$\hat{\boldsymbol{F}} = \tilde{\boldsymbol{F}}{\boldsymbol{V}_{NT}}^{1/2}, \hat{\boldsymbol{\Lambda}} = \tilde{\boldsymbol{\Lambda}}\boldsymbol{V}_{NT}^{-1/2}, \tag{A.1}$$

where $\tilde{\boldsymbol{F}} = \left[\tilde{\boldsymbol{f}}_1 \cdots \tilde{\boldsymbol{f}}_T\right]'$ and $\tilde{\boldsymbol{\Lambda}} = \left[\tilde{\boldsymbol{\lambda}}_1 \cdots \tilde{\boldsymbol{\lambda}}_N\right]'$ are the PCM estimates by minimizing the sum of squared idiosyncratic residuals under the condition that $\tilde{\boldsymbol{F}}'\tilde{\boldsymbol{F}}/T = \boldsymbol{I}$. The matrix $\boldsymbol{V}_{NT}$ is diagonal, and contains the r_0 largest eigenvalues of $\boldsymbol{X}'\boldsymbol{X}/(NT)$ in decreasing order on the diagonal. As is well-known, the principal component estimator $\tilde{\boldsymbol{f}}_t$ consistently estimates a rotation of $\boldsymbol{f}_{0t}$ given by $\tilde{\boldsymbol{H}}'\boldsymbol{f}_{0t}$. See, for example, Bai and Ng (2002), and Gonçalves and Perron (2014). As shown by Bai and Ng (2002), $\tilde{\boldsymbol{H}} = \frac{\boldsymbol{\Lambda}'\boldsymbol{\Lambda}}{N}\frac{\boldsymbol{F}_0'\tilde{\boldsymbol{F}}}{T}\boldsymbol{V}_{NT}^{-1}$. Therefore, $\hat{\boldsymbol{f}}_t = \boldsymbol{V}_{NT}^{1/2}\tilde{\boldsymbol{f}}_t$ estimates a rotation $\hat{\boldsymbol{H}}'\boldsymbol{f}_{0t}$ of $\boldsymbol{f}_{0t}$, where the rotation matrix $\hat{\boldsymbol{H}} = \tilde{\boldsymbol{H}}\boldsymbol{V}_{NT}^{1/2} = \frac{\boldsymbol{\Lambda}'\boldsymbol{\Lambda}}{N}\frac{\boldsymbol{F}_0'\hat{\boldsymbol{F}}}{T}\boldsymbol{V}_{NT}^{-1}$. Consequently, $\hat{\boldsymbol{\lambda}}_i = {\boldsymbol{V}_{NT}^{-1/2}}\tilde{\boldsymbol{\lambda}}_i$ also estimates a rotation $\hat{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}$ of $\boldsymbol{\lambda}_{0i}$. In the proofs of the results, we rely on the relationship between these two types of identification of the PCM. The discussed relationships also hold under the alternative with $\boldsymbol{F}_0$, $\boldsymbol{\Lambda}$ and $\boldsymbol{H}$ replaced by $\boldsymbol{G} : T \times r$, $\boldsymbol{\Phi} : N \times r$ and $\boldsymbol{\Xi} : T \times r$.

To establish the results under the null hypothesis, we impose the identification condition $\hat{\boldsymbol{\Lambda}}'\hat{\boldsymbol{\Lambda}}/N = \boldsymbol{I}$, which implies that $\check{\boldsymbol{H}}^{-1}\mathrm{E}(\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}')\check{\boldsymbol{H}}'^{-1} = \boldsymbol{I}$. To obtain that result, we first observe that $\mathrm{plim}\frac{\boldsymbol{\Lambda}'\boldsymbol{\Lambda}}{N} = \Sigma_{\boldsymbol{\Lambda}}$ from Assumption 1 (b). Second, $\mathrm{plim}\frac{\boldsymbol{F}_0'\tilde{\boldsymbol{F}}}{T} = \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}^{-1/2}\boldsymbol{\Upsilon}\boldsymbol{V}^{1/2}$ (Bai, 2003, Proposition 1), with $\boldsymbol{V} = \mathrm{plim}\boldsymbol{V}_{NT}$, the diagonal matrix with the eigenvalues of $\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}^{1/2}\boldsymbol{\Sigma}_{\boldsymbol{F}}\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}^{1/2}$ in decreasing order, and $\boldsymbol{\Upsilon}$ the associated matrix of eigenvectors such that $\boldsymbol{\Upsilon}'\boldsymbol{\Upsilon} = \boldsymbol{I}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}^{1/2}\boldsymbol{\Sigma}_{\boldsymbol{F}}\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}^{1/2}\boldsymbol{\Upsilon} = \boldsymbol{V}\boldsymbol{\Upsilon}$. See also Han and Inoue (2015, Equation (2.8)) for a similar argument. Therefore,

$$\check{\boldsymbol{H}} = \mathrm{plim}\hat{\boldsymbol{H}} = \mathrm{plim}\tilde{\boldsymbol{H}}\boldsymbol{V}_{NT}^{1/2} = \mathrm{plim}\frac{\boldsymbol{\Lambda}'\boldsymbol{\Lambda}}{N}\frac{\boldsymbol{F}_0'\tilde{\boldsymbol{F}}}{T}\boldsymbol{V}_{NT}^{-1/2} = \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}^{-1/2}\boldsymbol{\Upsilon}\boldsymbol{V}^{1/2}\boldsymbol{V}^{-1/2} = \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}^{1/2}\boldsymbol{\Upsilon}. \tag{A.2}$$

Thus,

$$\check{\boldsymbol{H}}^{-1}\mathrm{E}\left(\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'\right)\check{\boldsymbol{H}}'^{-1} = \left(\check{\boldsymbol{H}}'\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}^{-1}\check{\boldsymbol{H}}\right)^{-1} = \left(\boldsymbol{\Upsilon}\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}^{1/2}\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}^{-1}\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}^{1/2}\boldsymbol{\Upsilon}\right)^{-1} = \boldsymbol{I}, \tag{A.3}$$

and the estimated factor loadings converge to a rotation $\check{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}$, in which the second moment is $\boldsymbol{I}$. We state the following results, which help to prove Lemma 2.1, Theorem 1, and Theorem 2.

Lemma A.1. *Suppose that Assumptions 1 and 2 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*

$$\frac{1}{\lfloor \alpha N \rfloor}\sum_{i=1}^{\lfloor \alpha N \rfloor}\left(\hat{\boldsymbol{\lambda}}_i - \hat{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\right)\boldsymbol{\lambda}_{0i}' = O_P\left(\frac{1}{\delta_{NT}^2}\right), \tag{A.4}$$

$$\frac{1}{N - \lfloor \alpha N \rfloor}\sum_{i=\lfloor \alpha N \rfloor + 1}^{N}\left(\hat{\boldsymbol{\lambda}}_i - \hat{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\right)\boldsymbol{\lambda}_{0i}' = O_P\left(\frac{1}{\delta_{NT}^2}\right), \tag{A.5}$$

$$\frac{1}{\lfloor \alpha N \rfloor}\sum_{i=1}^{\lfloor \alpha N \rfloor}\left\|\hat{\boldsymbol{\lambda}}_i - \hat{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\right\|^2 = O_P\left(\frac{1}{\delta_{NT}^2}\right), \tag{A.6}$$

$$\frac{1}{N - \lfloor \alpha N \rfloor}\sum_{i=\lfloor \alpha N \rfloor + 1}^{N}\left\|\hat{\boldsymbol{\lambda}}_i - \hat{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\right\|^2 = O_P\left(\frac{1}{\delta_{NT}^2}\right) \tag{A.7}$$

uniformly in α. It also holds that

$$\frac{1}{N}\sum_{i=1}^{N}\left\|\hat{\boldsymbol{\lambda}}_i\hat{\boldsymbol{\lambda}}_i' - \hat{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'\hat{\boldsymbol{H}}'^{-1}\right\|^2 = O_P\left(\frac{N}{T^2}\right) \tag{A.8}$$

and

$$\frac{1}{N}\sum_{i=1}^{N}\left\|\hat{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'\hat{\boldsymbol{H}}^{-1\prime} - \check{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'\check{\boldsymbol{H}}'^{-1}\right\|^2 = o_P(1). \tag{A.9}$$

,

Lemma A.2. *Suppose that Assumption 1 (b) and Assumption 2 (e) are satisfied. As $N \to \infty$, for any α such that $\alpha \in [\alpha_1, \alpha_2]$, it holds uniformly in α that*

$$\underset{N\to\infty}{\mathrm{plim}}\, \boldsymbol{S}\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right) = \lim_{N\to\infty}\mathrm{Var}\left(\boldsymbol{A}\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right)\right), \tag{A.10}$$

$$
\boldsymbol{A}\left(\alpha,\boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right)\xrightarrow{d}\mathrm{N}\left(\mathbf{0},\lim_{N\to\infty}\mathrm{Var}\left(\boldsymbol{A}\left(\alpha,\boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right)\right)\right), \tag{A.11}
$$

$$
\frac{\sqrt{N}}{\lfloor\alpha N\rfloor}\sum_{i=1}^{\lfloor\alpha N\rfloor}\left(\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'-\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)=O_P\left(1\right)\ \ \textit{and}\ \frac{\sqrt{N}}{N-\lfloor\alpha N\rfloor}\sum_{i=\lfloor\alpha N\rfloor+1}^{N}\left(\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'-\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)=O_P\left(1\right). \tag{A.12}
$$

The proof of (A.4) and (A.6) uses similar steps with the ones in Bai and Ng (2004, Lemma A3). Since (A.5) and (A.7) can be proved following nearly identical steps to (A.4) and (A.6) , they are omitted. The results (A.8) and (A.9) show how close are to the rotated cross-product of factor loading to their estimates. (A.10), (A.11) and (A.12) 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 are satisfied. If as $N,T\to\infty$, $\sqrt{N}/T\to 0$, then for any $\alpha\in[\alpha_1,\alpha_2]$, it holds uniformly in α that*

$$
\begin{aligned}
\frac{1}{\lfloor\alpha N\rfloor}\sum_{i=1}^{\lfloor\alpha N\rfloor}\left(\hat{\boldsymbol{\lambda}}_i-\hat{\boldsymbol{\Xi}}^{-1}\boldsymbol{\phi}_i\right)\boldsymbol{\phi}_i' &= O_P\left(\delta_{NT}^{-2}\right), && \text{(A.13)}\\
\frac{1}{N-\lfloor\alpha N\rfloor}\sum_{i=\lfloor\alpha N\rfloor+1}^{N}\left(\hat{\boldsymbol{\lambda}}_i-\hat{\boldsymbol{\Xi}}^{-1}\boldsymbol{\phi}_i\right)\boldsymbol{\phi}_i' &= O_P\left(\delta_{NT}^{-2}\right), && \text{(A.14)}\\
\frac{1}{\lfloor\alpha N\rfloor}\sum_{i=1}^{\lfloor\alpha N\rfloor}\left\|\hat{\boldsymbol{\lambda}}_i-\hat{\boldsymbol{\Xi}}^{-1}\boldsymbol{\phi}_i\right\|^2 &= O_P\left(\delta_{NT}^{-2}\right), && \text{(A.15)}\\
\frac{1}{N-\lfloor\alpha N\rfloor}\sum_{i=\lfloor\alpha N\rfloor+1}^{N}\left\|\hat{\boldsymbol{\lambda}}_i-\hat{\boldsymbol{\Xi}}^{-1}\boldsymbol{\phi}_i\right\|^2 &= O_P\left(\delta_{NT}^{-2}\right), && \text{(A.16)}
\end{aligned}
$$

$$
\frac{1}{N}\sum_{i=1}^{N}\left\|\hat{\boldsymbol{\lambda}}_i\hat{\boldsymbol{\lambda}}_i'-\hat{\boldsymbol{\Xi}}^{-1}\boldsymbol{\phi}_i\boldsymbol{\phi}_i'\hat{\boldsymbol{\Xi}}'^{-1}\right\|^2=O_P\left(\frac{N}{T^2}\right) \tag{A.17}
$$

and

$$
\frac{1}{N}\sum_{i=1}^{N}\left\|\hat{\boldsymbol{\Xi}}^{-1}\boldsymbol{\phi}_i\boldsymbol{\phi}_i'\hat{\boldsymbol{\Xi}}'^{-1}-\check{\boldsymbol{\Xi}}^{-1}\boldsymbol{\phi}_i\boldsymbol{\phi}_i'\check{\boldsymbol{\Xi}}'^{-1}\right\|^2=o_P\left(1\right). \tag{A.18}
$$

.

It also holds that

$$
\left\|\boldsymbol{S}\left(\alpha,\hat{\boldsymbol{\Lambda}}\right)-\boldsymbol{S}\left(\alpha,\boldsymbol{\Phi}\check{\boldsymbol{\Xi}}'^{-1}\right)\right\|=o_P\left(1\right) \tag{A.19}
$$

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.4), (A.6), (A.8) and (A.9). Because $\hat{\mathbf{\Lambda}} = \tilde{\mathbf{\Lambda}} \boldsymbol{V}_{NT}^{-1/2}$, $\hat{\boldsymbol{H}} = \tilde{\boldsymbol{H}} \boldsymbol{V}_{NT}^{1/2}$ and $\boldsymbol{V}_{NT} \xrightarrow{P} \boldsymbol{V}$, where $\boldsymbol{V}$ is positive definite (Bai, 2003, Lemma A3), proving (A.4), (A.6), (A.8) and (A.9) are equivalent to proving the results with the hat replaced by a tilde.

Proof of (A.4) To demonstrate that uniformly in α, $\frac{1}{\lfloor \alpha N \rfloor} \sum_{i=1}^{\lfloor \alpha N \rfloor} \left(\tilde{\boldsymbol{\lambda}}_i - \tilde{\boldsymbol{H}}^{-1} \boldsymbol{\lambda}_{0i} \right) \boldsymbol{\lambda}_{0i}' = O_P\left(\frac{1}{\delta_{NT}^2} \right)$, we use the following identity from the proof of Bai and Ng (2004, Lemma A2).

$$\tilde{\boldsymbol{\lambda}}_i - \tilde{\boldsymbol{H}}^{-1} \boldsymbol{\lambda}_{0i} = \frac{1}{T} \tilde{\boldsymbol{H}}' \boldsymbol{F}_0' \boldsymbol{e}_i + \frac{1}{T} \tilde{\boldsymbol{F}}' \left(\boldsymbol{F}_0 - \tilde{\boldsymbol{F}} \tilde{\boldsymbol{H}}^{-1} \right) \boldsymbol{\lambda}_{0i} + \frac{1}{T} \left(\tilde{\boldsymbol{F}} - \boldsymbol{F}_0 \tilde{\boldsymbol{H}} \right)' \boldsymbol{e}_i, \tag{A.20}$$

where $\boldsymbol{e}_i = (e_{i1}, e_{i2}, \ldots, e_{iT})'$. It follows uniformly in α that $\frac{1}{\lfloor \alpha N \rfloor} \sum_{i=1}^{\lfloor \alpha N \rfloor} \left(\tilde{\boldsymbol{\lambda}}_i - \tilde{\boldsymbol{H}}^{-1} \boldsymbol{\lambda}_{0i} \right) \boldsymbol{\lambda}_{0i}'$ equals

$$\underbrace{\frac{1}{\lfloor \alpha N \rfloor} \sum_{i=1}^{\lfloor \alpha N \rfloor} \frac{1}{T} \tilde{\boldsymbol{H}}' \boldsymbol{F}_0' \boldsymbol{e}_i \boldsymbol{\lambda}_{0i}'}_{\boldsymbol{A}_1} + \underbrace{\frac{1}{\lfloor \alpha N \rfloor} \sum_{i=1}^{\lfloor \alpha N \rfloor} \frac{1}{T} \tilde{\boldsymbol{F}}' \left(\boldsymbol{F}_0 - \tilde{\boldsymbol{F}} \tilde{\boldsymbol{H}}^{-1} \right) \boldsymbol{\lambda}_{0i} \boldsymbol{\lambda}_{0i}'}_{\boldsymbol{A}_2} + \underbrace{\frac{1}{\lfloor \alpha N \rfloor} \sum_{i=1}^{\lfloor \alpha N \rfloor} \frac{1}{T} \left(\tilde{\boldsymbol{F}} - \boldsymbol{F}_0 \tilde{\boldsymbol{H}} \right)' \boldsymbol{e}_i \boldsymbol{\lambda}_i'}_{\boldsymbol{A}_3}.$$

Since $\mathrm{E} \left\| \frac{1}{NT} \sum_{i=1}^{\lfloor \alpha N \rfloor} \sum_{t=1}^{T} \boldsymbol{f}_{0t} e_{it} \boldsymbol{\lambda}_{0i}' \right\| = O\left(\frac{1}{\sqrt{NT}} \right)$ by Assumption 2 (c) and $\tilde{\boldsymbol{H}} = O_P(1)$, we have $\boldsymbol{A}_1 = \tilde{\boldsymbol{H}} \frac{1}{\lfloor \alpha N \rfloor T} \sum_{i=1}^{\lfloor \alpha N \rfloor} \sum_{t=1}^{T} \boldsymbol{f}_{0t} e_{it} \boldsymbol{\lambda}_{0i}' = O_P\left(\frac{1}{\sqrt{NT}} \right)$, uniformly in α. Moreover, we also have $\boldsymbol{A}_2 = \frac{1}{T} \tilde{\boldsymbol{F}}' \left(\boldsymbol{F}_0 - \tilde{\boldsymbol{F}} \tilde{\boldsymbol{H}}^{-1} \right) \frac{1}{\lfloor \alpha N \rfloor} \sum_{i=1}^{\lfloor \alpha N \rfloor} \boldsymbol{\lambda}_{0i} \boldsymbol{\lambda}_{0i}' = O_P\left(\frac{1}{\delta_{NT}^2} \right)$, uniformly in α because $\frac{1}{T} \tilde{\boldsymbol{F}}' \left(\boldsymbol{F}_0 - \tilde{\boldsymbol{F}} \tilde{\boldsymbol{H}}^{-1} \right) = O_P\left(\frac{1}{\delta_{NT}^2} \right)$ given Bai (2003, Lemma B3), and $\mathrm{E} \left\| \frac{1}{\lfloor \alpha N \rfloor} \sum_{i=1}^{\lfloor \alpha N \rfloor} \boldsymbol{\lambda}_{0i} \boldsymbol{\lambda}_{0i}' \right\|$ 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

$$\tilde{\boldsymbol{f}}_{0t} - \tilde{\boldsymbol{H}}' \boldsymbol{f}_{0t} = \boldsymbol{V}_{NT}^{-1} \left(\frac{1}{T} \sum_{s=1}^{T} \tilde{\boldsymbol{f}}_s \gamma_{st} + \frac{1}{T} \sum_{s=1}^{T} \tilde{\boldsymbol{f}}_s \zeta_{st} + \frac{1}{T} \sum_{s=1}^{T} \tilde{\boldsymbol{f}}_s \eta_{st} + \frac{1}{T} \sum_{s=1}^{T} \tilde{\boldsymbol{f}}_s \xi_{st} \right),$$

where $\gamma_{st} = \mathrm{E}\left(\frac{1}{N} \sum_{i=1}^{N} e_{is} e_{it} \right)$, $\zeta_{st} = \frac{1}{N} \sum_{i=1}^{N} \left(e_{is} e_{it} - \mathrm{E}\left(\frac{1}{N} \sum_{i=1}^{N} e_{is} e_{it} \right) \right)$, $\eta_{st} = \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{\lambda}_i' \boldsymbol{f}_{0s} e_{it}$, $\xi_{st} = \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{\lambda}_{0i}' \boldsymbol{f}_{0t} e_{is}$ and $\boldsymbol{V}_{NT}^{-1} = O_P(1)$ as $\boldsymbol{V}_{NT} \xrightarrow{P} \boldsymbol{V} > \mathbf{0}$. Thus $\boldsymbol{A}_3 = \boldsymbol{V}_{NT}^{-1} (\boldsymbol{B}_1 + \boldsymbol{B}_2 + \boldsymbol{B}_3 + \boldsymbol{B}_4)$, with

$$\begin{aligned}
\boldsymbol{B}_1 &\equiv \frac{1}{T^2} \sum_{t=1}^{T} \sum_{s=1}^{T} \tilde{\boldsymbol{f}}_s \gamma_{st} \left(\frac{1}{\lfloor \alpha N \rfloor} \sum_{i=1}^{\lfloor \alpha N \rfloor} e_{it} \boldsymbol{\lambda}_{0i}' \right), \\
\boldsymbol{B}_2 &\equiv \frac{1}{T^2} \sum_{t=1}^{T} \sum_{s=1}^{T} \tilde{\boldsymbol{f}}_s \zeta_{st} \left(\frac{1}{\lfloor \alpha N \rfloor} \sum_{i=1}^{\lfloor \alpha N \rfloor} e_{it} \boldsymbol{\lambda}_{0i}' \right), \\
\boldsymbol{B}_3 &\equiv \frac{1}{T^2} \sum_{t=1}^{T} \sum_{s=1}^{T} \tilde{\boldsymbol{f}}_s \eta_{st} \left(\frac{1}{\lfloor \alpha N \rfloor} \sum_{i=1}^{\lfloor \alpha N \rfloor} e_{it} \boldsymbol{\lambda}_{0i}' \right)
\end{aligned}$$

and

$$\boldsymbol{B}_4 \equiv \frac{1}{T^2} \sum_{t=1}^{T} \sum_{s=1}^{T} \tilde{\boldsymbol{f}}_s \xi_{st} \left(\frac{1}{\lfloor \alpha N \rfloor} \sum_{i=1}^{\lfloor \alpha N \rfloor} e_{it} \boldsymbol{\lambda}_{0i}' \right).$$

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

$$\begin{aligned}\boldsymbol{B}_1 &= \frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\left(\tilde{\boldsymbol{f}}_s - \tilde{\boldsymbol{H}}'\boldsymbol{f}_{0s}\right)\gamma_{st}\left(\frac{1}{\lfloor \alpha N\rfloor}\sum_{i=1}^{\lfloor \alpha N\rfloor} e_{it}\boldsymbol{\lambda}_{0i}'\right) + \tilde{\boldsymbol{H}}'\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\boldsymbol{f}_{0s}\gamma_{st}\left(\frac{1}{\lfloor \alpha N\rfloor}\sum_{i=1}^{\lfloor \alpha N\rfloor} e_{it}\boldsymbol{\lambda}_{0i}'\right)\\ &\equiv \boldsymbol{B}_{11} + \boldsymbol{B}_{12}.\end{aligned}$$

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

$$\begin{aligned}&\frac{1}{T}\sum_{s=1}^{T}\left\|\tilde{\boldsymbol{f}}_s - \tilde{\boldsymbol{H}}'\boldsymbol{f}_{0s}\right\|\left(\frac{1}{T}\sum_{t=1}^{T}\gamma_{st}^2\right)^{1/2}\left(\frac{1}{T}\sum_{t=1}^{T}\left(\left\|\frac{1}{\lfloor \alpha N\rfloor}\sum_{i=1}^{\lfloor \alpha N\rfloor} e_{it}\boldsymbol{\lambda}_{0i}'\right\|\right)^2\right)^{1/2}\\ \leq\ &\left(\frac{1}{T}\sum_{s=1}^{T}\left\|\tilde{\boldsymbol{f}}_s - \tilde{\boldsymbol{H}}'\boldsymbol{f}_{0s}\right\|^2\right)^{1/2}\left(\frac{1}{T^2}\sum_{s=1}^{T}\sum_{t=1}^{T}\gamma_{st}^2\right)^{1/2}\left(\frac{1}{T}\sum_{t=1}^{T}\left(\left\|\frac{1}{\lfloor \alpha N\rfloor}\sum_{i=1}^{\lfloor \alpha N\rfloor} e_{it}\boldsymbol{\lambda}_{0i}'\right\|\right)^2\right)^{1/2} = O_P\left(\frac{1}{\delta_{NT}\sqrt{NT}}\right),\end{aligned}$$

where from Bai and Ng (2002, Theorem 1), $\frac{1}{T}\sum_{s=1}^{T}\left\|\tilde{\boldsymbol{f}}_s - \tilde{\boldsymbol{H}}'\boldsymbol{f}_{0s}\right\|^2 = O_P\left(1/\delta_{NT}^2\right)$, $\frac{1}{T}\sum_{s=1}^{T}\sum_{t=1}^{T}\gamma_{st}^2 = O\left(1\right)$ (see Bai and Ng (2002, Lemma 1(i))) and $\mathrm{E}\left(\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{\lfloor \alpha N\rfloor} e_{it}\boldsymbol{\lambda}_{0i}'\right\|\right)^2 \leq C$ by Assumption 2 (d). Similarly, the second term $\boldsymbol{B}_{12}$ is bounded by

$$\left\|\tilde{\boldsymbol{H}}\right\|\left(\frac{1}{T}\sum_{s=1}^{T}\left\|\boldsymbol{f}_{0s}\right\|^2\right)^{1/2}\left(\frac{1}{T^2}\sum_{s=1}^{T}\sum_{t=1}^{T}\gamma_{st}^2\right)^{1/2}\left(\frac{1}{T}\sum_{t=1}^{T}\left(\left\|\frac{1}{\lfloor \alpha N\rfloor}\sum_{i=1}^{\lfloor \alpha N\rfloor} e_{it}\boldsymbol{\lambda}_{0i}'\right\|\right)^2\right)^{1/2} = O_P\left(\frac{1}{\sqrt{TN}}\right),$$

given that $\left\|\tilde{\boldsymbol{H}}\right\| = O_P\left(1\right)$ and $\mathrm{E}\left\|\boldsymbol{f}_{0s}\right\|^2 \leq C$. 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

$$\begin{aligned}\boldsymbol{B}_2 &= \frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\left(\tilde{\boldsymbol{f}}_s - \tilde{\boldsymbol{H}}'\boldsymbol{f}_{0s}\right)\zeta_{st}\left(\frac{1}{\lfloor \alpha N\rfloor}\sum_{i=1}^{\lfloor \alpha N\rfloor} e_{it}\boldsymbol{\lambda}_{0i}'\right) + \tilde{\boldsymbol{H}}'\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\boldsymbol{f}_{0s}\zeta_{st}\left(\frac{1}{\lfloor \alpha N\rfloor}\sum_{i=1}^{\lfloor \alpha N\rfloor} e_{it}\boldsymbol{\lambda}_{0i}'\right)\\ &\equiv \boldsymbol{B}_{21} + \boldsymbol{B}_{22}.\end{aligned}$$

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

$$\left(\frac{1}{T}\sum_{s=1}^{T}\left\|\tilde{\boldsymbol{f}}_{0s} - \boldsymbol{H}'\boldsymbol{f}_{0s}\right\|^2\right)^{1/2}\left(\frac{1}{T^2}\sum_{s=1}^{T}\sum_{t=1}^{T}\zeta_{st}^2\right)^{1/2}\left(\frac{1}{T}\sum_{t=1}^{T}\left(\left\|\frac{1}{\lfloor \alpha N\rfloor}\sum_{i=1}^{\lfloor \alpha N\rfloor} e_{it}\boldsymbol{\lambda}_{0i}'\right\|\right)^2\right)^{1/2} = O_P\left(\frac{1}{\delta_{NT}\sqrt{TN}}\right),$$

where by Jensen inequality and Assumption 1 (f),

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

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

$$\boldsymbol{B}_{22} \leq \left(\frac{1}{T}\sum_{s=1}^{T}\left\|\boldsymbol{f}_{0s}\right\|^2\right)^{1/2}\left(\frac{1}{T^2}\sum_{s=1}^{T}\sum_{t=1}^{T}\zeta_{st}^2\right)^{1/2}\left(\frac{1}{T}\sum_{t=1}^{T}\left(\left\|\frac{1}{\lfloor \alpha N\rfloor}\sum_{i=1}^{\lfloor \alpha N\rfloor} e_{it}\boldsymbol{\lambda}_{0i}'\right\|\right)^2\right)^{1/2} = O_P\left(\frac{1}{\sqrt{TN}}\right).$$

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

$$\boldsymbol{B}_{31} = \frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\left(\tilde{\boldsymbol{f}}_{0s} - \tilde{\boldsymbol{H}}'\boldsymbol{f}_{0s}\right)\eta_{st}\left(\frac{1}{\lfloor \alpha N\rfloor}\sum_{i=1}^{\lfloor \alpha N\rfloor} e_{it}\boldsymbol{\lambda}_{0i}'\right) \text{ and } \boldsymbol{B}_{32} = \tilde{\boldsymbol{H}}'\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\boldsymbol{f}_{0s}\eta_{st}\left(\frac{1}{\lfloor \alpha N\rfloor}\sum_{i=1}^{\lfloor \alpha N\rfloor} e_{it}\boldsymbol{\lambda}_{0i}'\right).$$

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

$$\begin{aligned}\frac{1}{T^2}\sum_{s=1}^{T}\sum_{t=1}^{T}\eta_{st}^2 &= \frac{1}{T^2}\sum_{s=1}^{T}\sum_{t=1}^{T}\left(\boldsymbol{f}_{0s}'\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\lambda}_i\boldsymbol{f}_{0s}e_{it}\right)^2 \leq \frac{1}{T^2}\sum_{s=1}^{T}\sum_{t=1}^{T}\|\boldsymbol{f}_{0s}\|^2\left\|\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\lambda}_{0i}'e_{it}\right\|^2\\ &= \left(\frac{1}{T}\sum_{s=1}^{T}\|\boldsymbol{f}_{0s}\|^2\right)\frac{1}{N}\left(\frac{1}{T}\sum_{t=1}^{T}\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\lambda}_{0i}'e_{it}\right\|^2\right) = O_P\left(\frac{1}{N}\right),\end{aligned}$$

as $\mathrm{E}\left(\frac{1}{T}\sum_{t=1}^{T}\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\lambda}_{0i}'e_{it}\right\|^2\right) \leq C$ follows from Assumption 2 (d) and the c_r inequality. Thus, using also $\frac{1}{T}\sum_{s=1}^{T}\left\|\tilde{\boldsymbol{f}}_s - \tilde{\boldsymbol{H}}'\boldsymbol{f}_{0s}\right\|^2 = O_P\left(\frac{1}{\delta_{NT}^2}\right)$ and Assumption 2 (d),

$$\|\boldsymbol{B}_{31}\| \leq \left(\frac{1}{T}\sum_{s=1}^{T}\left\|\tilde{\boldsymbol{f}}_s - \tilde{\boldsymbol{H}}'\boldsymbol{f}_{0s}\right\|^2\right)^{1/2}\left(\frac{1}{T^2}\sum_{s=1}^{T}\sum_{t=1}^{T}\eta_{st}^2\right)^{1/2}\left(\frac{1}{T}\sum_{t=1}^{T}\left(\left\|\frac{1}{\lfloor \alpha N\rfloor}\sum_{i=1}^{\lfloor \alpha N\rfloor}e_{it}\boldsymbol{\lambda}_{0i}'\right\|\right)^2\right)^{1/2} = O_P\left(\frac{1}{\delta_{NT}N}\right).$$

By the same steps, we also obtain

$$\|\boldsymbol{B}_{32}\| \leq \left\|\tilde{\boldsymbol{H}}\right\|\left(\frac{1}{T}\sum_{s=1}^{T}\|\boldsymbol{f}_{0s}\|^2\right)^{1/2}\left(\frac{1}{T^2}\sum_{s=1}^{T}\sum_{t=1}^{T}\eta_{st}^2\right)^{1/2}\left(\frac{1}{T}\sum_{t=1}^{T}\left(\left\|\frac{1}{\lfloor \alpha N\rfloor}\sum_{i=1}^{\lfloor \alpha N\rfloor}e_{it}\boldsymbol{\lambda}_{0i}'\right\|\right)^2\right)^{1/2} = O_P\left(\frac{1}{N}\right).$$

Hence, $\boldsymbol{B}_3 = O_P\left(\frac{1}{N}\right)$. The proof for $\boldsymbol{B}_4$ is similar to the proof of $\boldsymbol{B}_3$ 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 $\frac{1}{\lfloor \alpha N\rfloor}\sum_{i=1}^{\lfloor \alpha N\rfloor}\left(\tilde{\boldsymbol{\lambda}}_i - \tilde{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\right)\boldsymbol{\lambda}_{0i}' = O_P\left(\frac{1}{\delta_{NT}^2}\right)$ uniformly in α.

Proof of (A.6) For this proof, we use the decomposition (A.20), and obtain that uniformly in α,

$$\begin{aligned}&\frac{1}{\lfloor \alpha N\rfloor}\sum_{i=1}^{\lfloor \alpha N\rfloor}\left\|\tilde{\boldsymbol{\lambda}}_i - \tilde{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\right\|^2\\ \leq\; &\underbrace{\frac{1}{\lfloor \alpha N\rfloor}\sum_{i=1}^{\lfloor \alpha N\rfloor}\left\|\frac{1}{T}\tilde{\boldsymbol{H}}'\boldsymbol{F}_0'\boldsymbol{e}_i\right\|^2}_{C_1} + \underbrace{\frac{1}{\lfloor \alpha N\rfloor}\sum_{i=1}^{\lfloor \alpha N\rfloor}\left\|\frac{1}{T}\tilde{\boldsymbol{F}}'\left(\boldsymbol{F}_0 - \tilde{\boldsymbol{F}}\tilde{\boldsymbol{H}}^{-1}\right)\boldsymbol{\lambda}_{0i}\right\|^2}_{C_2} + \underbrace{\frac{1}{\lfloor \alpha N\rfloor}\sum_{i=1}^{\lfloor \alpha N\rfloor}\left\|\frac{1}{T}\left(\tilde{\boldsymbol{F}} - \tilde{\boldsymbol{H}}\boldsymbol{F}_0\right)'\boldsymbol{e}_i\right\|^2}_{C_3}.\end{aligned}$$

Since $\frac{1}{\lfloor \alpha N\rfloor}\sum_{i=1}^{\lfloor \alpha N\rfloor}\mathrm{E}\left\|\frac{1}{\sqrt{T}}\boldsymbol{F}_0'\boldsymbol{e}_i\right\|^2 \leq C$ by Assumption 2 (a) and $\tilde{\boldsymbol{H}} = O_P(1)$, we note that

$$C_1 \leq \frac{1}{T}\left\|\tilde{\boldsymbol{H}}\right\|^2\frac{1}{\lfloor \alpha N\rfloor}\sum_{i=1}^{\lfloor \alpha N\rfloor}\mathrm{E}\left\|\frac{1}{\sqrt{T}}\boldsymbol{F}_0'\boldsymbol{e}_i\right\|^2 = O_P\left(\frac{1}{T}\right),$$

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

$$C_2 = \left\| \frac{1}{T} \tilde{\boldsymbol{F}}' \left(\boldsymbol{F}_0 - \tilde{\boldsymbol{F}} \tilde{\boldsymbol{H}}^{-1} \right) \right\|^2 \frac{1}{\lfloor \alpha N \rfloor} \sum_{i=1}^{\lfloor \alpha N \rfloor} \|\boldsymbol{\lambda}_{0i}\|^2 = O_P\left(\frac{1}{\delta_{NT}^4} \right)$$

because $\frac{1}{T} \tilde{\boldsymbol{F}}' \left(\boldsymbol{F}_0 - \tilde{\boldsymbol{F}} \tilde{\boldsymbol{H}}^{-1} \right) = O_P\left(\frac{1}{\delta_{NT}^2} \right)$ from Bai (2003, Lemma B3) and $\frac{1}{\lfloor \alpha N \rfloor} \sum_{i=1}^{\lfloor \alpha N \rfloor} \|\boldsymbol{\lambda}_{0i}\|^2 = O_P(1)$ as $\mathrm{E}\left(\|\boldsymbol{\lambda}_{0i}\|^2 \right) \leq C$. We now similarly observe that by Cauchy-Schwarz inequality,

$$C_3 \leq \left(\frac{1}{T} \sum_{s=1}^{T} \left\| \tilde{\boldsymbol{f}}_t - \tilde{\boldsymbol{H}}' \boldsymbol{f}_{0t} \right\|^2 \right) \left(\frac{1}{T \lfloor \alpha N \rfloor} \sum_{i=1}^{\lfloor \alpha N \rfloor} \sum_{t=1}^{T} e_{it}^2 \right) = O_P\left(\frac{1}{\delta_{NT}^2} \right) O_P(1) = O_P\left(\frac{1}{\delta_{NT}^2} \right).$$

From the bound for C_1, C_2 and C_3, $\frac{1}{\lfloor \alpha N \rfloor} \sum_{i=1}^{\lfloor \alpha N \rfloor} \left\| \tilde{\boldsymbol{\lambda}}_i - \tilde{\boldsymbol{H}}^{-1} \boldsymbol{\lambda}_{0i} \right\|^2 = O_P\left(\frac{1}{\delta_{NT}^2} \right)$ uniformly in α.

Proof of (A.8) Using the identity that $\tilde{\boldsymbol{\lambda}}_i \tilde{\boldsymbol{\lambda}}_i' - \tilde{\boldsymbol{H}}^{-1} \boldsymbol{\lambda}_{0i} \boldsymbol{\lambda}_{0i}' \left(\tilde{\boldsymbol{H}}^{-1} \right)'$ equals

$$\left(\tilde{\boldsymbol{\lambda}}_i - \tilde{\boldsymbol{H}}^{-1} \boldsymbol{\lambda}_{0i} \right) \left(\tilde{\boldsymbol{\lambda}}_i - \tilde{\boldsymbol{H}}^{-1} \boldsymbol{\lambda}_{0i} \right)' + \tilde{\boldsymbol{H}}^{-1} \boldsymbol{\lambda}_i \left(\tilde{\boldsymbol{\lambda}}_i - \tilde{\boldsymbol{H}}^{-1} \boldsymbol{\lambda}_{0i} \right)' + \left(\tilde{\boldsymbol{\lambda}}_i - \tilde{\boldsymbol{H}}^{-1} \boldsymbol{\lambda}_{0i} \right) \boldsymbol{\lambda}_{0i}' \left(\tilde{\boldsymbol{H}}^{-1} \right)',$$

we can write by the c_r inequality and an application of the Cauchy-Schwarz inequality that $\frac{1}{N} \sum_{i=1}^{N} \left\| \tilde{\boldsymbol{\lambda}}_i \tilde{\boldsymbol{\lambda}}_i' - \tilde{\boldsymbol{H}}^{-1} \boldsymbol{\lambda}_{0i} \boldsymbol{\lambda}_{0i}' \left(\tilde{\boldsymbol{H}}^{-1} \right)' \right\|^2$ is bounded by

$$\frac{1}{N} \sum_{i=1}^{N} \left\| \tilde{\boldsymbol{\lambda}}_i - \tilde{\boldsymbol{H}}^{-1} \boldsymbol{\lambda}_{0i} \right\|^4 + 2 \left\| \tilde{\boldsymbol{H}}^{-1} \right\|^2 \left(\frac{1}{N} \sum_{i=1}^{N} \|\boldsymbol{\lambda}_{0i}\|^4 \right)^{1/2} \left(\frac{1}{N} \sum_{i=1}^{N} \left\| \tilde{\boldsymbol{\lambda}}_i - \tilde{\boldsymbol{H}}^{-1} \boldsymbol{\lambda}_{0i} \right\|^4 \right)^{1/2}.$$

Since $\tilde{\boldsymbol{H}} \xrightarrow{P} \boldsymbol{H}$, with $\boldsymbol{H}$ is nonsingular, $\tilde{\boldsymbol{H}}^{-1} = O_P(1)$. Moreover, since $\mathrm{E}\left(\|\boldsymbol{\lambda}_{0i}\|^4 \right)$, we also have that $\frac{1}{N} \sum_{i=1}^{N} \|\boldsymbol{\lambda}_{0i}\|^4 = O_P(1)$. Hence, we only need to show that $\frac{1}{N} \sum_{i=1}^{N} \left\| \tilde{\boldsymbol{\lambda}}_i - \tilde{\boldsymbol{H}}^{-1} \boldsymbol{\lambda}_{0i} \right\|^4 \xrightarrow{P} 0$. We use here nearly identical steps to the proof of (A.6), with $\frac{1}{N} \sum_{i=1}^{N} \left\| \tilde{\boldsymbol{\lambda}}_i - \boldsymbol{H}^{-1} \boldsymbol{\lambda}_{0i} \right\|^4$ lower than

$$3^3 \left(\underbrace{\frac{1}{N} \sum_{i=1}^{N} \left\| \frac{1}{T} \tilde{\boldsymbol{H}}' \boldsymbol{F}_0' \boldsymbol{e}_i \right\|^4}_{I_1} + \underbrace{\frac{1}{N} \sum_{i=1}^{N} \left\| \frac{1}{T} \tilde{\boldsymbol{F}}' \left(\boldsymbol{F}_0 - \tilde{\boldsymbol{F}} \tilde{\boldsymbol{H}}^{-1} \right) \boldsymbol{\lambda}_{0i} \right\|^4}_{I_2} + \underbrace{\frac{1}{N} \sum_{i=1}^{N} \left\| \frac{1}{T} \left(\tilde{\boldsymbol{F}} - \boldsymbol{F}_0 \tilde{\boldsymbol{H}} \right)' \boldsymbol{e}_i \right\|^4}_{I_3} \right)$$

based on (A.20) and the c_r inequality. As $\tilde{\boldsymbol{H}} = O_P(1)$ and $\frac{1}{N} \sum_{i=1}^{N} \left\| \frac{1}{\sqrt{T}} \boldsymbol{F}_0' \boldsymbol{e}_i \right\|^2 = O_P(1)$ from Assumption 2 (b), we note that

$$I_1 \leq \frac{1}{T^2} \left\| \tilde{\boldsymbol{H}} \right\|^4 \frac{1}{N} \sum_{i=1}^{N} \left\| \frac{1}{\sqrt{T}} \boldsymbol{F}_0' \boldsymbol{e}_i \right\|^4 \leq \frac{N}{T^2} \left\| \tilde{\boldsymbol{H}} \right\|^4 \left(\frac{1}{N} \sum_{i=1}^{N} \left\| \frac{1}{\sqrt{T}} \boldsymbol{F}_0' \boldsymbol{e}_i \right\|^2 \right)^2 = O_P\left(\frac{N}{T^2} \right).$$

Moreover, because $\frac{1}{T} \tilde{\boldsymbol{F}}' \left(\boldsymbol{F}_0 - \tilde{\boldsymbol{F}} \tilde{\boldsymbol{H}}_0^{-1} \right) = O_P\left(\frac{1}{\delta_{NT}^2} \right)$ and $\frac{1}{N} \sum_{i=1}^{N} \|\boldsymbol{\lambda}_{0i}\|^4 = O_P(1)$ given $\mathrm{E}\left(\|\lambda_{0i}\|^4 \right) \leq C$, we obtain $I_2 = \left\| \frac{1}{T} \tilde{\boldsymbol{F}}' \left(\boldsymbol{F}_0 - \tilde{\boldsymbol{F}} \tilde{\boldsymbol{H}}^{-1} \right) \right\|^4 \frac{1}{N} \sum_{i=1}^{N} \|\boldsymbol{\lambda}_{0i}\|^4 = O_P\left(\frac{1}{\delta_{NT}^8} \right)$. We also have by Cauchy-

Schwarz inequality that

$$I_3 \leq \left(\frac{1}{T}\sum_{s=1}^{T}\left\|\tilde{\boldsymbol{f}}_t - \tilde{\boldsymbol{H}}'\boldsymbol{f}_{0t}\right\|^2\right)^2 \frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{T}\sum_{t=1}^{T}e_{it}^2\right)^2 = \left(\frac{1}{T}\sum_{s=1}^{T}\left\|\tilde{\boldsymbol{f}}_t - \tilde{\boldsymbol{H}}'\boldsymbol{f}_{0t}\right\|^2\right)^2 \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}e_{it}^4 = O_P\left(\frac{1}{\delta_{NT}^4}\right),$$

where the first equality employs Jensen inequality. From the bound for I_1, I_2 and I_3, we deduce $\frac{1}{N}\sum_{i=1}^{N}\left\|\tilde{\boldsymbol{\lambda}}_i - \boldsymbol{H}^{-1}\boldsymbol{\lambda}_{0i}\right\|^4 = O_P\left(\frac{N}{T^2}\right)$. Hence, $\frac{1}{N}\sum_{i=1}^{N}\left\|\tilde{\boldsymbol{\lambda}}_i\tilde{\boldsymbol{\lambda}}_i' - \tilde{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'\tilde{\boldsymbol{H}}'^{-1}\right\|^2 = O_P\left(\frac{N}{T^2}\right)$.

Proof of (A.9) The result follows from the decomposition $\tilde{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'\tilde{\boldsymbol{H}}'^{-1} - \boldsymbol{H}^{-1}\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'\boldsymbol{H}'^{-1} = \tilde{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'\left(\tilde{\boldsymbol{H}}^{-1} - \boldsymbol{H}^{-1}\right)' + \left(\tilde{\boldsymbol{H}}^{-1} - \boldsymbol{H}^{-1}\right)\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'\boldsymbol{H}'^{-1}$, and the c_r inequality. We note that

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

using in particular the fact that $\tilde{\boldsymbol{H}}^{-1} - \boldsymbol{H}^{-1} = o_P(1)$ as $\tilde{\boldsymbol{H}} \xrightarrow{P} \boldsymbol{H}$, 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.10) Define

$$\boldsymbol{A}_1\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right) = \frac{\sqrt{N}}{\alpha N}\sum_{i=1}^{\alpha N}\text{Vech}\left(\check{\boldsymbol{H}}^{-1}\left(\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\check{\boldsymbol{H}}^{-1\prime}\right) \tag{A.21}$$

and

$$\boldsymbol{A}_2\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right) = \frac{\sqrt{N}}{(1-\alpha)N}\sum_{i=\alpha N+1}^{N}\text{Vech}\left(\check{\boldsymbol{H}}\left(\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\check{\boldsymbol{H}}\right). \tag{A.22}$$

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

$$\begin{aligned}
\boldsymbol{S}_1\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right) &= \frac{1}{\alpha}\frac{1}{N}\sum_{i=1}^{N}\left(\text{Vech}\left(\check{\boldsymbol{H}}^{-1}\left(\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\check{\boldsymbol{H}}'^{-1}\right)\text{Vech}\left(\check{\boldsymbol{H}}^{-1}\left(\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\check{\boldsymbol{H}}'^{-1}\right)'\right), \\
\boldsymbol{S}_2\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right) &= \frac{1}{1-\alpha}\frac{1}{N}\sum_{i=1}^{N}\left(\text{Vech}\left(\check{\boldsymbol{H}}^{-1}\left(\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\check{\boldsymbol{H}}'^{-1}\right)\text{Vech}\left(\check{\boldsymbol{H}}^{-1}\left(\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\check{\boldsymbol{H}}'^{-1}\right)'\right),
\end{aligned}$$

the proof proceeds by showing that

$$\operatorname*{plim}_{N\to\infty}\boldsymbol{S}_1\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right) = \lim_{N\to\infty}\text{Var}\left(\boldsymbol{A}_1\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right)\right) \tag{A.23}$$

and

$$\operatorname*{plim}_{N\to\infty}\boldsymbol{S}_2\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right) = \lim_{N\to\infty}\text{Var}\left(\boldsymbol{A}_2\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right)\right). \tag{A.24}$$

We start with (A.23). From Assumption 1 (a) and Assumption 2 (e), $\text{Var}\left(\boldsymbol{A}_1\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right)\right)$ is

$$\frac{1}{\alpha}\frac{1}{\alpha N}\sum_{i=1}^{\alpha N} \text{E}\left(\text{Vech}\left(\check{\boldsymbol{H}}^{-1}\left(\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\check{\boldsymbol{H}}'^{-1}\right)\text{Vech}\left(\check{\boldsymbol{H}}^{-1}\left(\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\check{\boldsymbol{H}}'^{-1}\right)'\right) = \text{E}\left(\boldsymbol{S}_1\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right)\right).$$

In consequence, to show (A.23), we only need to prove that $\boldsymbol{a}'\left(\boldsymbol{S}_1\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right) - \text{E}\left(\boldsymbol{S}_1\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right)\right)\right)\boldsymbol{a}' = o_P(1)$ where $\boldsymbol{a}$ is a $\frac{r_0(r_0+1)}{2}\times 1$-vector of real numbers such that $\boldsymbol{a}'\boldsymbol{a} = 1$. We have $\boldsymbol{a}'\boldsymbol{S}_1\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right)\boldsymbol{a} - \boldsymbol{a}'\text{E}\left(\boldsymbol{S}_1\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right)\right)\boldsymbol{a} = \frac{1}{\alpha N}\sum_{i=1}^{N} s_i$, using Assumption 2 (e), with

$$\begin{aligned} s_i &= \boldsymbol{a}'\text{Vech}\left(\check{\boldsymbol{H}}^{-1}\left(\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\check{\boldsymbol{H}}'^{-1}\right)\text{Vech}\left(\check{\boldsymbol{H}}^{-1}\left(\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\check{\boldsymbol{H}}'^{-1}\right)'\boldsymbol{a} \\ &\quad -\boldsymbol{a}'\text{E}\left(\text{Vech}\left(\check{\boldsymbol{H}}^{-1}\left(\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\check{\boldsymbol{H}}'^{-1}\right)\text{Vech}\left(\check{\boldsymbol{H}}^{-1}\left(\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\check{\boldsymbol{H}}'^{-1}\right)'\right)\boldsymbol{a}. \end{aligned}$$

Since $\text{E}(s_i) = 0$, $N^{-1/2}\sum_{i=1}^{N} s_i = O_P(1)$ if $\text{E}(|s_i|) \leq C$. To show that $\text{E}(|s_i|) \leq C$, we first apply the triangle inequality and the Jensen inequality to have

$$\text{E}(|s_i|) \leq 2\text{E}\left|\boldsymbol{a}'\left(\text{Vech}\left(\check{\boldsymbol{H}}^{-1}\left(\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\check{\boldsymbol{H}}'^{-1}\right)\text{Vech}\left(\check{\boldsymbol{H}}^{-1}\left(\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\check{\boldsymbol{H}}'^{-1}\right)'\right)\boldsymbol{a}\right|,$$

and using the Cauchy Schwarz inequality and the c_r inequality, we find that

$$\text{E}(|s_i|) \leq 2\text{E}\left\|\text{Vech}\left(\check{\boldsymbol{H}}^{-1}\left(\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\check{\boldsymbol{H}}'^{-1}\right)\right\|^2 \leq 4\left\|\check{\boldsymbol{H}}^{-1}\right\|^4 \text{E}\left\|\boldsymbol{\lambda}_{0i}\right\|^4 \leq C.$$

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

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

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

Proof of (A.11) Recall that $\boldsymbol{A}_1\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right) = \frac{\sqrt{N}}{\alpha N}\sum_{i=1}^{\alpha N}\text{Vech}\left(\check{\boldsymbol{H}}^{-1}\left(\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\check{\boldsymbol{H}}'^{-1}\right)$. Since the limit of $\text{Var}\left(\boldsymbol{A}_1\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right)\right)$ is positive definite under Assumption 2 (e), we simply need to show that $\left(\text{Var}\left(\boldsymbol{A}_1\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right)\right)\right)^{-1/2}\boldsymbol{A}_1\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right) \xrightarrow{d} \text{N}(\boldsymbol{0}, \boldsymbol{I})$. Let $\boldsymbol{a}$ be an $\frac{r_0(r_0+1)}{2}\times 1$-vector of real numbers such that $\boldsymbol{a}'\boldsymbol{a} = 1$. Define

$$z_i = \boldsymbol{a}'\left(\text{Var}\left(\boldsymbol{A}_1\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right)\right)\right)^{-1/2}\frac{\sqrt{N}}{\alpha N}\text{Vech}\left(\check{\boldsymbol{H}}^{-1}\left(\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\check{\boldsymbol{H}}'^{-1}\right).$$

It follows that, $\text{E}\left(\sum_{i=1}^{\alpha N} z_i\right) = 0$ and $\text{E}\left(\sum_{i=1}^{\alpha N} z_i^2\right) = 1$. Therefore (A.11) 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

$$\sum_{i=1}^{\alpha N}\text{E}|z_i|^{2+\xi} \leq \left\|\left(\text{Var}\left(\boldsymbol{A}_1\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right)\right)\right)^{-1/2}\right\|^{2+\xi}\left(\frac{\sqrt{N}}{\alpha N}\right)^{2+\xi}\sum_{i=1}^{N}\text{E}\left\|\text{Vech}\left(\check{\boldsymbol{H}}^{-1}\left(\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\check{\boldsymbol{H}}'^{-1}\right)\right\|^{2+\xi}. \tag{A.25}$$

By an application of the c_r inequality, the fact that $\mathrm{E}\left(\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'\right) = \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}$ and the Jensen inequality, we have $\mathrm{E}\left\|\mathrm{Vech}\left(\check{\boldsymbol{H}}^{-1}\left(\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\check{\boldsymbol{H}}'^{-1}\right)\right\|^{2+\xi} \leq \left\|\check{\boldsymbol{H}}^{-1}\right\|^{4+2\xi} 2^{1+\xi} \times 2\mathrm{E}\left\|\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'\right\|^{2+\xi} \leq C$ if $\mathrm{E}\left\|\boldsymbol{\lambda}_{0i}\right\|^{4+2\xi} \leq C$ (Assumption 1 (b)), and deduce that $\sum_{i=1}^{N}\mathrm{E}|z_i|^{2+\xi} = O\left(\left(\frac{\sqrt{N}}{\alpha N}\right)^{2+\xi}\alpha N\right) = O\left(\frac{1}{N^{\xi/2}}\right) = o_P(1)$ given that $\mathrm{Var}\left(\boldsymbol{A}_1\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right)\right) = O(1)$. It similarly follows that

$$\boldsymbol{A}_2\left(\alpha, \boldsymbol{\Lambda}\left(\check{\boldsymbol{H}}'\right)^{-1}\right) \xrightarrow{d} \mathrm{N}\left(\mathbf{0}, \lim_{N\to\infty}\mathrm{Var}\left(\boldsymbol{A}_2\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right)\right)\right). \tag{A.26}$$

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

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

is positive definite, we conclude that (A.11) holds.

Proof of (A.12) Let $\boldsymbol{a}$ be an $\frac{r_0(r_0+1)}{2} \times 1$-vector such that $\boldsymbol{a}'\boldsymbol{a} = 1$. Assumption 1 (b) implies

$$\mathrm{E}\left(\frac{\sqrt{N}}{\alpha N}\sum_{i=1}^{\alpha N}\boldsymbol{a}'\left(\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\boldsymbol{a}\right)^2 = \frac{1}{\alpha}\frac{1}{\alpha N}\sum_{i=1}^{\alpha N}\mathrm{E}\left(\boldsymbol{a}'\left(\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\boldsymbol{a}\right)^2 \leq \frac{1}{\alpha}\frac{1}{\alpha N}\sum_{i=1}^{\alpha N}\mathrm{E}\left(\boldsymbol{a}'\left(\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'\right)\boldsymbol{a}\right)^2.$$

Given that $\mathrm{E}\left(\boldsymbol{a}'\left(\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'\right)\boldsymbol{a}\right)^2 = \left(\mathrm{E}\left(\boldsymbol{a}'\boldsymbol{\lambda}_{0i}\right)^4\right) \leq \mathrm{E}\left\|\boldsymbol{\lambda}_i\right\|^4 \leq C$, we obtain $\mathrm{E}\left(\frac{\sqrt{N}}{\alpha N}\sum_{i=1}^{\alpha N}\boldsymbol{a}'\left(\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\boldsymbol{a}\right)^2 \leq \frac{1}{\alpha}C < \infty$. Thus, $\frac{\sqrt{N}}{\alpha N}\sum_{i=1}^{\alpha N}\boldsymbol{a}'\left(\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\boldsymbol{a} = O_P(1)$, for any $\alpha \in [\alpha_1, \alpha_2]$. By identical arguments, it also holds that $\frac{\sqrt{N}}{N-\alpha N}\sum_{i=\alpha N+1}^{N}\left(\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}' - \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right) = O_P(1)$ for any $\alpha \in [\alpha_1, \alpha_2]$.

A.3 Proof of Lemma 2.1

We first note that the results in this section hold uniformly in α and we start with (13).

Proof of (13) A triangle inequality implies that $\left\|\boldsymbol{A}\left(\alpha, \hat{\boldsymbol{\Lambda}}\right) - \boldsymbol{A}\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right)\right\| \leq J_1 + J_2$, where $J_1 = \left\|\boldsymbol{A}\left(\alpha, \hat{\boldsymbol{\Lambda}}\right) - \boldsymbol{A}\left(\alpha, \boldsymbol{\Lambda}\hat{\boldsymbol{H}}'^{-1}\right)\right\|$ and $J_2 = \left\|\boldsymbol{A}\left(\alpha, \boldsymbol{\Lambda}\hat{\boldsymbol{H}}'^{-1}\right) - \boldsymbol{A}\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right)\right\|$. Applying again the triangle inequality, we have $J_1 \leq J_{11} + J_{12}$ and $J_2 \leq J_{21} + J_{22}$, with

$$\begin{aligned}
J_{11} &= \sqrt{N}\left\|\mathrm{Vech}\left(\frac{1}{\lfloor \alpha N\rfloor}\sum_{i=1}^{\lfloor \alpha N\rfloor}\left(\hat{\boldsymbol{\lambda}}_i\hat{\boldsymbol{\lambda}}_i' - \hat{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'\left(\hat{\boldsymbol{H}}^{-1}\right)'\right)\right)\right\|, \\
J_{12} &= \sqrt{N}\left\|\mathrm{Vech}\left(\frac{1}{N-\lfloor \alpha N\rfloor}\sum_{i=\lfloor \alpha N\rfloor+1}^{N}\left(\hat{\boldsymbol{\lambda}}_i\hat{\boldsymbol{\lambda}}_i' - \hat{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'\left(\hat{\boldsymbol{H}}^{-1}\right)'\right)\right)\right\|, \\
J_{21} &= \sqrt{N}\left\|\mathrm{Vech}\left(\left(\hat{\boldsymbol{H}}^{-1} - \check{\boldsymbol{H}}^{-1}\right)\left(\frac{1}{\lfloor \alpha N\rfloor}\sum_{i=1}^{\lfloor \alpha N\rfloor}\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}' - \frac{1}{N-\lfloor \alpha N\rfloor}\sum_{i=\lfloor \alpha N\rfloor+1}^{N}\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'\right)\left(\hat{\boldsymbol{H}}^{-1}\right)'\right)\right\|
\end{aligned}$$

and

$$J_{22} = \sqrt{N}\left\|\mathrm{Vech}\left(\check{\boldsymbol{H}}^{-1}\left(\frac{1}{\lfloor \alpha N\rfloor}\sum_{i=1}^{\lfloor \alpha N\rfloor}\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}' - \frac{1}{N-\lfloor \alpha N\rfloor}\sum_{i=\lfloor \alpha N\rfloor+1}^{N}\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'\right)\left(\hat{\boldsymbol{H}}^{-1} - \check{\boldsymbol{H}}^{-1}\right)'\right)\right\|.$$

Therefore, to complete the proof of (13), 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 triangle inequality, J_{11} is bounded by

$$
\begin{aligned}
& \sqrt{N}\left\|\frac{1}{\lfloor\alpha N\rfloor}\sum_{i=1}^{\lfloor\alpha N\rfloor}\left(\hat{\boldsymbol{\lambda}}_i-\hat{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\right)\left(\hat{\boldsymbol{\lambda}}_i-\hat{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\right)'\right\|+2\sqrt{N}\left\|\frac{1}{\lfloor\alpha N\rfloor}\sum_{i=1}^{\lfloor\alpha N\rfloor}\hat{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\left(\hat{\boldsymbol{\lambda}}_i-\hat{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\right)'\right\| \\
\leq\;& \sqrt{N}\frac{1}{\lfloor\alpha N\rfloor}\sum_{i=1}^{\lfloor\alpha N\rfloor}\left\|\hat{\boldsymbol{\lambda}}_i-\hat{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\right\|^2+2\left\|\hat{\boldsymbol{H}}^{-1}\right\|\sqrt{N}\left\|\frac{1}{\lfloor\alpha N\rfloor}\sum_{i=1}^{\lfloor\alpha N\rfloor}\hat{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\left(\hat{\boldsymbol{\lambda}}_i-\hat{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\right)'\right\|.
\end{aligned}
$$

Thus, $J_{11} = O_P\left(\sqrt{N}\delta_{NT}^{-2}\right) = o_P(1)$ by an application of (A.4), (A.6), $\hat{\boldsymbol{H}}^{-1} = O_P(1)$ and $\sqrt{N}/T \to 0$. We also have for any $\alpha \in [\alpha_1, \alpha_2]$,

$$
\begin{aligned}
J_{21} \;\leq\;& \left\|\hat{\boldsymbol{H}}^{-1}-\check{\boldsymbol{H}}^{-1}\right\|\left\|\frac{\sqrt{N}}{\lfloor\alpha N\rfloor}\sum_{i=1}^{\lfloor\alpha N\rfloor}\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'-\frac{\sqrt{N}}{N-\lfloor\alpha N\rfloor}\sum_{i=\lfloor\alpha N\rfloor+1}^{N}\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'\right\|\left\|\hat{\boldsymbol{H}}^{-1}\right\| \\
\leq\;& \left\|\hat{\boldsymbol{H}}^{-1}-\check{\boldsymbol{H}}^{-1}\right\|\left(\left\|\frac{\sqrt{N}}{\lfloor\alpha N\rfloor}\sum_{i=1}^{\lfloor\alpha N\rfloor}\left(\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'-\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\right\|+\left\|\frac{\sqrt{N}}{N-\lfloor\alpha N\rfloor}\sum_{i=\lfloor\alpha N\rfloor+1}^{N}\left(\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'-\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}}\right)\right\|\right)\left\|\hat{\boldsymbol{H}}^{-1}\right\|,
\end{aligned}
$$

which is $o_P(1)$ using again $\hat{\boldsymbol{H}}^{-1} - \check{\boldsymbol{H}}^{-1} = o_P(1)$, $\hat{\boldsymbol{H}}^{-1} = O_P(1)$ and Equation (A.12).

Proof of (14) To show the consistency of the variance estimator $\boldsymbol{S}\left(\alpha, \hat{\boldsymbol{\Lambda}}\right)$ of $\boldsymbol{A}\left(\alpha, \hat{\boldsymbol{\Lambda}}\right)$, we need to prove that $\boldsymbol{S}\left(\alpha, \hat{\boldsymbol{\Lambda}}\right) - \boldsymbol{S}\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right) = o_P(1)$. Noting that

$$
\boldsymbol{S}\left(\alpha,\hat{\boldsymbol{\Lambda}}\right)-\boldsymbol{S}\left(\alpha,\boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right)=\left(\boldsymbol{S}\left(\alpha,\hat{\boldsymbol{\Lambda}}\right)-\boldsymbol{S}\left(\alpha,\boldsymbol{\Lambda}\hat{\boldsymbol{H}}'^{-1}\right)\right)+\left(\boldsymbol{S}\left(\alpha,\boldsymbol{\Lambda}\hat{\boldsymbol{H}}'^{-1}\right)-\boldsymbol{S}\left(\alpha,\boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right)\right), \tag{A.27}
$$

the next steps will consist of showing that

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

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

$$
\hat{\boldsymbol{\lambda}}_i\hat{\boldsymbol{\lambda}}_i'-\boldsymbol{I}=\left(\hat{\lambda}_i\hat{\boldsymbol{\lambda}}_i'-\hat{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'\hat{\boldsymbol{H}}'^{-1}\right)+\left(\hat{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'\hat{\boldsymbol{H}}'^{-1}-\boldsymbol{I}\right),
$$

$$
\hat{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'\hat{\boldsymbol{H}}'^{-1}-\boldsymbol{I}=\left(\hat{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'\hat{\boldsymbol{H}}'^{-1}-\hat{\boldsymbol{\lambda}}_i\hat{\boldsymbol{\lambda}}_i'\right)+\left(\hat{\boldsymbol{\lambda}}_i\hat{\boldsymbol{\lambda}}_i'-\boldsymbol{I}\right),
$$

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

$$
\begin{aligned}
\boldsymbol{C}_1 &= \frac{1}{N}\sum_{i=1}^{N}\operatorname{Vech}\left(\hat{\boldsymbol{\lambda}}_i\hat{\boldsymbol{\lambda}}_i'-\boldsymbol{I}\right)\operatorname{Vech}\left(\hat{\boldsymbol{\lambda}}_i\hat{\boldsymbol{\lambda}}_i'-\hat{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'\hat{\boldsymbol{H}}'^{-1}\right)', \\
\boldsymbol{C}_2 &= \frac{1}{N}\sum_{i=1}^{N}\operatorname{Vech}\left(\hat{\boldsymbol{\lambda}}_i\hat{\boldsymbol{\lambda}}_i'-\hat{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'\hat{\boldsymbol{H}}'^{-1}\right)\operatorname{Vech}\left(\hat{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'\hat{\boldsymbol{H}}'^{-1}-\boldsymbol{I}\right)'.
\end{aligned}
$$

From the Cauchy-Schwarz inequality and the c_r inequality,

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

where

$$\|\boldsymbol{C}_1\| \leq 2^{1/2}\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\hat{\boldsymbol{\lambda}}_i\hat{\boldsymbol{\lambda}}_i'\right\|^2+\|\boldsymbol{I}\|^2\right)^{1/2}\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\hat{\boldsymbol{\lambda}}_i\hat{\boldsymbol{\lambda}}_i'-\hat{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'\hat{\boldsymbol{H}}'^{-1}\right\|^2\right)^{1/2}=O_P\left(\sqrt{N}/T\right),$$

and

$$\|\boldsymbol{C}_2\| \leq 2^{1/2}\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\hat{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_i\boldsymbol{\lambda}_i'\hat{\boldsymbol{H}}'^{-1}\right\|^2+\|\boldsymbol{I}\|^2\right)^{1/2}\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\hat{\boldsymbol{\lambda}}_i\hat{\boldsymbol{\lambda}}_i'-\hat{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'\hat{\boldsymbol{H}}'^{-1}\right\|^2\right)^{1/2}=O_P\left(\sqrt{N}/T\right),$$

because $\frac{1}{N}\sum_{i=1}^{N}\left\|\hat{\boldsymbol{\lambda}}_i\hat{\boldsymbol{\lambda}}_i'-\hat{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'\hat{\boldsymbol{H}}'^{-1}\right\|^2=O_P\left(\sqrt{N}/T\right)$, using (A.8). We also use the fact that $\frac{1}{N}\sum_{i=1}^{N}\left\|\hat{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'\hat{\boldsymbol{H}}'^{-1}\right\|^2$ is bounded by $\left\|\hat{\boldsymbol{H}}^{-1}\right\|^4\frac{1}{N}\sum_{i=1}^{N}\|\boldsymbol{\lambda}_{0i}\|^4=O_P(1)$, and $\frac{1}{N}\sum_{i=1}^{N}\left\|\hat{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_i\right\|^2=\left\|\hat{\boldsymbol{H}}^{-1}\right\|^2\frac{1}{N}\sum_{i=1}^{N}\|\boldsymbol{\lambda}_{0i}\|^2=O_P(1)$. Note that by the c_r inequality, we have

$$\frac{1}{N}\sum_{i=1}^{N}\left\|\hat{\boldsymbol{\lambda}}_i\hat{\boldsymbol{\lambda}}_i'\right\|^2\leq 2\frac{1}{N}\sum_{i=1}^{N}\left\|\hat{\boldsymbol{\lambda}}_i\hat{\boldsymbol{\lambda}}_i'-\hat{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'\hat{\boldsymbol{H}}'^{-1}\right\|^2+2\frac{1}{N}\sum_{i=1}^{N}\left\|\hat{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'\hat{\boldsymbol{H}}'^{-1}\right\|^2=O_P(1).$$

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

$$\begin{aligned}&2^{1/2}\left(\frac{1}{\alpha}+\frac{1}{1-\alpha}\right)\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\hat{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'\hat{\boldsymbol{H}}'^{-1}-\check{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'\check{\boldsymbol{H}}'^{-1}\right\|^2\right)^{1/2}\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\hat{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\right\|^4+\|\boldsymbol{I}\|^4\right)^{1/2}\\&+2^{1/2}\left(\frac{1}{\alpha}+\frac{1}{1-\alpha}\right)\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\hat{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'\hat{\boldsymbol{H}}'^{-1}-\check{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'\check{\boldsymbol{H}}'^{-1}\right\|^2\right)^{1/2}\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\check{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\right\|^4+\|\boldsymbol{I}\|^4\right)^{1/2}.\end{aligned}$$

Since $\frac{1}{N}\sum_{i=1}^{N}\left\|\hat{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'\hat{\boldsymbol{H}}'^{-1}-\check{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_{0i}\boldsymbol{\lambda}_{0i}'\check{\boldsymbol{H}}'^{-1}\right\|^2=o_P(1)$ from (A.9), and also that $\frac{1}{N}\sum_{i=1}^{N}\left\|\hat{\boldsymbol{H}}^{-1}\boldsymbol{\lambda}_i\right\|^4\leq\left\|\hat{\boldsymbol{H}}^{-1}\right\|^4\frac{1}{N}\sum_{i=1}^{N}\|\boldsymbol{\lambda}_{0i}\|^4=O_P(1)$, we obtain the second needed result, which is

$$\left\|\boldsymbol{S}\left(\alpha,\boldsymbol{\Lambda}\hat{\boldsymbol{H}}'^{-1}\right)-\boldsymbol{S}\left(\alpha,\boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right)\right\|=o_P(1).$$

A.4 Proof of Theorem 1

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

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

and

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

The proof uses the auxiliary results that (a) $\boldsymbol{A}\left(\alpha, \hat{\boldsymbol{\Lambda}}\right) - \boldsymbol{A}\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right) = o_P(1)$, (b) $\boldsymbol{A}\left(\alpha, \boldsymbol{\Lambda}\left(\check{\boldsymbol{H}}'\right)^{-1}\right) = O_P(1)$ and $\boldsymbol{A}\left(\alpha, \hat{\boldsymbol{\Lambda}}\right) = O_P(1)$ and (c) $\left\| \boldsymbol{S}\left(\alpha, \hat{\boldsymbol{\Lambda}}\right)^{-1} - \boldsymbol{S}\left(\alpha, \boldsymbol{\Lambda}\left(\check{\boldsymbol{H}}'\right)^{-1}\right)^{-1}\right\| = o_P(1)$ uniformly in α. We note that (a) follows from (13). Second, note that $\left\| \boldsymbol{A}\left(\alpha, \boldsymbol{\Lambda}\left(\check{\boldsymbol{H}}'\right)^{-1}\right)\right\|$ is equal to

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

uniformly in α given Lemma A.2 (A.11). We also have uniformly in α, that

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

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

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

where $\boldsymbol{S}\left(\alpha, \hat{\boldsymbol{\Lambda}}\right) - \boldsymbol{S}\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right) \xrightarrow{P} \boldsymbol{0}$ by Lemma 2.1 (14) and the limit in probability of $\boldsymbol{S}\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right)$ is positive definite from Assumption 2 (e). Hence, $\left\| \boldsymbol{S}\left(\alpha, \hat{\boldsymbol{\Lambda}}\right) - \boldsymbol{S}\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right)\right\| = o_P(1)$, $\left\| \boldsymbol{S}\left(\alpha, \hat{\boldsymbol{\Lambda}}\right)^{-1}\right\| = O_P(1)$, $\left\| \boldsymbol{S}\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right)^{-1}\right\| = O_P(1)$. Thus, we can see that $\left\| \boldsymbol{S}\left(\alpha, \hat{\boldsymbol{\Lambda}}\right)^{-1} - \boldsymbol{S}\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right)^{-1}\right\| = o_P(1)$, which is result (c). From (a), (b) and (c), $\left| LM_N\left(\alpha, \hat{\boldsymbol{\Lambda}}\right) - LM_N\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right)\right| = o_P(1)$. Because $LM_N\left(\alpha, \boldsymbol{\Lambda}\check{\boldsymbol{H}}'^{-1}\right) \xrightarrow{d} \chi^2\left(\frac{r_0(r_0+1)}{2}\right)$ from Lemma A.2 (A.11), the result follows, applying the asymptotic equivalence lemma.

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 (14) are satisfied in the context of the one-level representation of the two-level alternative. This is the case if Assumption 3 complemented Assumptions 1 and 2. The analogue of Lemma A.1 follows with $\hat{\boldsymbol{\Xi}}^{-1}$ and $\check{\boldsymbol{\Xi}}^{-1}$ in the place $\hat{\boldsymbol{H}}^{-1}$ and $\check{\boldsymbol{H}}^{-1}$, respectively, using the same the steps as in the proof of Lemma A.1.

A.6 Proof of Theorem 2

Recall the matrix of factor loadings $\boldsymbol{\Phi} = [\boldsymbol{\phi}_1 \cdots \boldsymbol{\phi}_N]' : N \times r$ of the one-level representation under the two-level alternative. From Bai (2003), the PCM estimates $\hat{\boldsymbol{\Lambda}}$ converge to a rotation $\boldsymbol{\Phi}\hat{\boldsymbol{\Xi}}'^{-1}$

with $\hat{\boldsymbol{\Xi}}$ an $r \times r$ rotation matrix. Thus, we can write

$$\frac{1}{N_1}\sum_{i=1}^{N_1}\hat{\boldsymbol{\lambda}}_i\hat{\boldsymbol{\lambda}}_i' - \frac{1}{N-N_1}\sum_{i=N_1+1}^{N}\hat{\boldsymbol{\lambda}}_i\hat{\boldsymbol{\lambda}}_i' = \boldsymbol{R}_1 + \boldsymbol{R}_2 - \boldsymbol{R}_3, \tag{A.28}$$

where

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

Given that

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

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,

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

In addition, $\hat{\boldsymbol{\Xi}}^{-1} = \check{\boldsymbol{\Xi}}^{-1} + o_P(1)$. Using Assumption 3 (b) and the same steps as in Lemma A.2 (A.12), $\frac{1}{N_1}\sum_{i=1}^{N_1}\boldsymbol{\phi}_i\boldsymbol{\phi}_i' = O_P(1)$ and $\frac{1}{N-N_1}\sum_{i=N_1+1}^{N}\boldsymbol{\phi}_i\boldsymbol{\phi}_i' = O_P(1)$. More precisely, $\boldsymbol{R}_1 = \boldsymbol{R}_0 + o_P(1)$, where

$$\boldsymbol{R}_0 = \check{\boldsymbol{\Xi}}^{-1}\operatorname*{plim}_{N\to\infty}\left(\frac{1}{N_1}\sum_{i=1}^{N_1}\boldsymbol{\phi}_i\boldsymbol{\phi}_i' - \frac{1}{N-N_1}\sum_{i=N_1+1}^{N}\boldsymbol{\phi}_i\boldsymbol{\phi}_i'\right)\check{\boldsymbol{\Xi}}'^{-1} = \check{\boldsymbol{\Xi}}^{-1}\begin{bmatrix}\mathbf{0} & \boldsymbol{\Sigma}_{0D} & -\boldsymbol{\Sigma}_{0E}\\ \boldsymbol{\Sigma}_{0D}' & \boldsymbol{\Sigma}_{DD} & \mathbf{0}\\ -\boldsymbol{\Sigma}_{0E}' & \mathbf{0} & -\boldsymbol{\Sigma}_{EE}\end{bmatrix}\check{\boldsymbol{\Xi}}'^{-1} \tag{A.29}$$

is different from $\mathbf{0}$ since the second term is nonzero as $\boldsymbol{\Sigma}_{DD} \neq \mathbf{0}$ or $\boldsymbol{\Sigma}_{EE} \neq \mathbf{0}$, and the rows of $\check{\boldsymbol{\Xi}}'^{-1}$ are linearly independent. Hence,

$$\frac{1}{N_1}\sum_{i=1}^{N_1}\hat{\boldsymbol{\lambda}}_i\hat{\boldsymbol{\lambda}}_i' - \frac{1}{N-N_1}\sum_{i=N_1+1}^{N}\hat{\boldsymbol{\lambda}}_i\hat{\boldsymbol{\lambda}}_i' = \boldsymbol{R}_0 + o_P(1), \tag{A.30}$$

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

$$\left\|\boldsymbol{S}\left(\alpha, \hat{\boldsymbol{\Lambda}}\right) - \boldsymbol{S}\left(\alpha, \boldsymbol{\Phi}\check{\boldsymbol{\Xi}}'^{-1}\right)\right\| = o_P(1), \tag{A.31}$$

given Lemma A.3 (A.19), and the positive definiteness of the limit in probability $\boldsymbol{S}_0$ of $\boldsymbol{S}\left(\alpha, \boldsymbol{\Phi}\check{\boldsymbol{\Xi}}'^{-1}\right)$ given Assumption 3 (i).

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 (2008). Large dimensional factor analysis. *Foundations and Trends(R) in Econometrics 3*(2), 89–163.

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.).

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. A. (2019). Model Selection in Factor-Augmented Regressions with Estimated Factors. Working Papers 1391, Queen's University, Department of Economics.

Flood, R. P. and A. Rose (2010). Inflation targeting and business cycle synchronization. *Journal of International Money and Finance 29*(4), 704–727.

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.

Hirata, H., A. M. Kose, and C. Otrok (2013). *Regionalization vs. Globalization? Global Interdependence, Decoupling, and Recoupling.* Mit Press.

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.

Mumtaz, H., S. Simonelli, and P. Surico (2011). International Comovements, Business Cycle and Inflation: a Historical Perspective. *Review of Economic Dynamics 14*(1), 176–198.

Onatski, A. (2010). Determining the number of factors from empirical distribution of eigenvalues. *The Review of Economics and Statistics 92*(4), 1004–1016.

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.

Wälti, S. (2012). The myth of decoupling. *Applied Economics 44*(26), 3407–3419.

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.