12-2020

Testing for structural changes in factor models via a nonparametric regression

Liangjun SU
Singapore Management University, ljsu@smu.edu.sg

Xia WANG
Sun Yat-sen University

Testing for Structural Changes in Factor Models via a Nonparametric Regression*

Liangjun Su[a] and Xia Wang[b]

[a] School of Economics, Singapore Management University, Singapore

[b] Lingnan (University) College, Sun Yat-Sen University, China

September 10, 2019

Abstract

We propose a model-free test for structural changes in factor models. The basic idea is to regress the data on the commonly estimated factors by local smoothing and compare the fitted values of time-varying factor loadings with those of the time-invariant factor loadings estimated by the principal component analysis. By construction, the test is powerful against both smooth structural changes and sudden structural breaks with possibly unknown number of breaks and unknown break dates in the factor loadings. No restrictions on the form of alternatives or trimming of the boundary regions near the beginning or ending of the sample period is required for the test. The test has power to detect the usual nonparametric rate of local alternatives. Monte Carlo studies demonstrate excellent power of the test in detecting both smooth and sudden structural changes in the factor loadings. In an application to U.S. asset returns, we find significant evidence against time-invariant factor loadings.

JEL Classification: C12, C14, C33, C38.

Key Words: Factor model, Test, Local smoothing, Structural change, Local power

*The authors sincerely thank the co-editor Patrik Guggenberger and three anonymous referees for their constructive comments and suggestions on the early versions of the paper. Su gratefully acknowledges the Singapore Ministry of Education for Tier-2 Academic Research Fund (AcRF) under grant number MOE2012-T2-2-021 and the funding support provided by the Lee Kong Chian Fund for Excellence. Wang acknowledges financial supports from the National Science Foundation of China (No.71873151, No. 71401160). Please address correspondence to: Xia Wang, Lingnan (University) College, Sun Yat-Sen University, Guangzhou 510275, China, Phone +86 8411 1191. Email: *wangxia25@mail.sysu.edu.cn*.

1 Introduction

Factor models are extensively studied in the economics and finance literature. Since the datasets under analysis usually span a large number of time periods, the data generating processes for the underlying variables may undergo significant structural changes during the sampling period. However, the factor loadings, which capture the relationship between economic/financial variables and the latent common factors, are usually assumed to be time-invariant by most of the existing literature (e.g., Stock and Watson 2002, Bai and Ng 2002, Bai 2003). This assumption may stem from Stock and Watson's (2002, 2009) arguments that when the factor loadings undergo "small" instabilities, the estimated factors by the principal component analysis (PCA) are still consistent. In fact, since macroeconomic datasets usually have a long time span, it is difficult to assume that the factor loadings are time-invariant or only undergo asymptotically negligible changes during the sampling period. The driving forces such as institutional switching, economic transition, preference changes and technological progress, may influence the relationship between economic/financial variables and the common factors significantly. If the assumption of time-invariant factor loadings fails, the estimated common factors can be inconsistent and the inference and forecasting based on such an assumption may lead to misleading conclusions.

Recently, many researchers study the issue of structural changes in factor models. Stock and Watson (2009) examine the forecasting reliability when there exist structural breaks in the factor loadings. Breitung and Eickmeier (2011) propose LR, LM and Wald statistics to test structural breaks in factor loadings. Chen et al. (2014) propose a two-stage procedure to detect big breaks in factor loadings by testing the parameter stability in a regression of one of the factors estimated by PCA on the remaining estimated factors. Corradi and Swanson (2014) propose a test to check structural stability of both factor loadings and factor augmented forecasting regression coefficients. Han and Inoue (2015) propose a test for structural breaks of factor loadings by checking whether the second moments of the estimated factors exhibit a structural change. Yamamoto and Tanaka (2015) propose a modified version of Breitung and Eickmeier's (2011) test to ensure the robustness to the non-monotonic power problem. Cheng et al. (2016) consider the case in which both the factor loadings and the number of factors may change simultaneously. Baltagi et al. (2016) propose a sequential procedure to test multiple structural breaks by testing the null of l change points versus the alternative of $l+1$ change points. These studies provide appropriate econometric tools to examine the possible structural breaks in factor loadings. However, all these tests are proposed to check sudden structural breaks. In fact, such driving forces of structural changes as preference changes, technological progress, and policy changes usually take effect gradually in a long time. Even some abrupt policy changes also take a period to take effect. Hence, it seems more realistic to assume smooth changes rather than sudden breaks in such scenarios. In addition, there is a growing literature on the time-varying factor models; see Stock and Watson (2002), Banerjee et al. (2008), Del Negro and Otrok (2009), Bates et al. (2013), Eickmeier et al. (2015), and Mikkelsen et al. (2015). All these papers assume that the time-varying factor loadings follow a random walk process or a vector autoregressive process and only consider the estimation problem. Recently, Su and Wang (2017) model the time-varying factor loadings as piecewise smooth functions of scaled time and estimate the time-varying factor loadings and factors by the local version of PCA. They show that these estimators are consistent up to a rotation matrix that is *time-dependent*. This is in sharp contrast to the fact that the rotation matrix for the conventional PCA is *time-independent*. For this reason, they cannot compare the factor loadings estimates from the conventional PCA with those from the local PCA directly. Instead, they propose an L_2-distance-based test statistic to contrast the estimators of the *common components* under the null hypothesis of no structural changes in the factor loadings and the alternative hypothesis respectively.

In this paper, we propose a simple nonparametric test for the null hypothesis of no structural changes in large dimensional factor models. Specifically, the test is preceded by two steps. First, we apply the conventional PCA to the data $\{X_{it},\ i = 1, 2, ..., N,\ t = 1, 2, ..., T\}$ to obtain the estimators of the common factors and time-invariant factor loadings under the null, which are consistent up to a time-independent rotation matrix under the null and inconsistent in general otherwise. Second, we regress $\{X_{it}\}$ on the estimated common factors by local smoothing to obtain the estimators of time-varying factor loadings. The test statistic is then constructed by measuring the squared-L_2-distance between these two estimators of factor loadings. The intuition works as follows. If the null hypothesis is true, both the first-stage PCA estimator of the factor loadings and the second-stage local PCA estimator of the factor loadings are consistent with the same true time-invariant factor loadings up to a common rotation matrix and thus the L_2-distance between them would be small. On the other hand, under the alternative hypothesis of structural changes in the factor loadings, the two estimators of the factor loadings will differ substantially from each other, which gives the power for the resulting L_2-distance-based test statistic.

The test complements the nonparametric test of Su and Wang (2017) and generally outperforms existing parametric tests. It has a number of appealing features. First, like the test of Su and Wang (2017), the new test does not require one to impose stringent restrictions on the form of alternatives. In particular, our test is powerful against a large class of smooth structural changes as well as one or multiple sudden structural breaks in factor loadings. For the sudden structural breaks, we don't need to know the break dates or the number of breaks. This is in contrast to existing tests for the stability of factor loadings, all of which focus on the sudden structural breaks, especially the single structural break. Simulation studies demonstrate excellent power of our test in detecting various forms of structural changes such as the single structural break, the multiple structural breaks, and the smooth structural changes. Second, our test does not require trimming of the boundary regions near the starting or ending of the sampling period and has excellent power in detecting breaks that occur near the end of the sample. In contrast, all existing parametric tests for unknown break date such as the supremum-type tests of Breitung and Eickmeier (2011), Chen et al. (2014), Han and Inoue (2015) and Cheng et al. (2016) rely on a prespecified trimming parameter and hence would miss the possible structural changes in the boundary regions. Third, the new test shares some common features as the test of Su and Wang (2017): both tests are of nonparametric nature, have the asymptotic normal null distribution, and can detect local alternatives that converge to the null at the same rate that is faster than the usual $T^{-1/2}$-rate detected by existing parametric tests. The superb performance of our nonparametric test and that of Su and Wang (2017) is essentially because they explore the information from both the cross-sectional and time dimensions effectively while existing parametric tests mainly rely on the information along the time dimension. To simplify the derivation, our asymptotic theory requires that the error term be the m.d.s. over the time but it allows for cross-sectional dependence. One could possibly allow for both serial and cross-sectional dependence. In this case, the asymptotic variance of our test statistic will involve double summations along both the cross-sectional and time dimensions and we are not aware how to estimate it consistently.

After the first version of the paper was circulated, we found that Mikkelsen et al. (2015) also consider time-varying factor models. They model the time-varying factor loadings as a vector autoregressive process and propose a two-step maximum likelihood procedure to estimate the parameters in the models. Nevertheless, they do not consider a specification test for the time-invariant factor models.

The rest of this paper is organized as follows. In Section 2, we state the hypotheses of interest and construct the test statistic. In Section 3, we investigate the asymptotic properties of our test. In Section 4, we study the finite sample performance of the test via simulations. Section 5 provides an empirical application. Section 6 concludes. The proofs of the main results and some additional simulation results are relegated to the Online Supplement.

NOTATION. For an $m \times n$ real matrix A, we denote its transpose as A', its Frobenius norm as $\|A\|$ ($\equiv [\mathrm{tr}(AA')]^{1/2}$), and its spectral norm as $\|A\|_{\mathrm{sp}}$ ($\equiv \sqrt{\mu_1(A'A)}$), where $\equiv$ means "is defined as" and $\mu_s(\cdot)$ denotes the sth largest eigenvalue of a real symmetric matrix by counting eigenvalues of multiplicity multiple times. The operator $\stackrel{P}{\to}$ denotes convergence in probability, $\stackrel{d}{\to}$ convergence in distribution, and plim probability limit. We use $(N,T) \to \infty$ to denote that N and T pass to infinity jointly. Let $N \wedge T = \min(N,T)$. Let $C \in (0,\infty)$ denote a generic positive constant that may vary from case to case.

2 Hypotheses and Test Statistic

In this section, we introduce the hypotheses and test statistic.

2.1 Hypotheses

Let $\{X_{it}, i = 1,2,\ldots,N;\ t = 1,2,\ldots,T\}$ be an N-dimensional time series with T observations. The index i represents the ith cross section unit in panel data set or the ith random variable in a multiple time series data set. We assume that X_{it} is generated via the following factor model

$$X_{it} = \lambda_{it}' F_t + e_{it}, \tag{2.1}$$

where F_t is an $R \times 1$ vector of common factors, λ_{it} is an $R \times 1$ vector of factor loadings that can admit sudden and/or smooth structural changes over time, and e_{it} is the idiosyncratic error term.

The null hypothesis of no structural change in the above factor model could be written as:

$$\mathbb{H}_0 : \lambda_{it} = \lambda_{i0} \text{ for } i = 1,2,\ldots,N \text{ and } t = 1,2,\ldots,T.$$

The alternative hypothesis is

$$\mathbb{H}_1 : \lambda_{it} \neq \lambda_{i0} \text{ for some non-negligible values of } (i,t).$$

Apparently, under $\mathbb{H}_0$, λ_{it} is time-invariant and model (2.1) degenerates to the factor model with time-invariant factor loadings. This model has been elaboratively studied by Stock and Watson (2002), Bai and Ng (2002), and Bai (2003), among others. Nevertheless, it is well known that factor models may exhibit structural changes over time. For this reason, much recent research has been focusing on testing for structural changes in factor models. See, e.g., Breitung and Eickmeier (2011), Chen et al. (2014), Han and Inoue (2015), and Cheng et al. (2016). These authors mainly focus on testing the existence of a single structural break in the factor loadings by using some supremum-type test statistics. However, usually no prior information about the structural change alternative is available in practice. It is extremely restrictive to assume only a single sudden structural break in factor loadings. Most recently, Baltagi et al. (2016) provide a sequential procedure to detect multiple structural changes, which is also a special case of our alternative hypothesis.

To capture a wide range of alternatives, we consider a nonparametric local smoothing approach. More precisely, we follow the nonparametric literature on time-varying models (see, e.g., Cai 2007, Robinson 2012 and Chen et al. 2012) and model λ_{it} as a nonrandom function of t/T :

$$\lambda_{it} = \lambda_i(t/T),$$

where $\lambda_i(\cdot)$ is an unknown piece-wise smooth function on $(0,1]$ for each i with a finite number of discontinuity points. By allowing $\lambda_i(\cdot)$ to have a finite number of discontinuities, our alternative covers both sudden structural breaks and smooth structural changes. A special case is:

$$\lambda_i(t/T) = \begin{cases} \lambda_{i(1)} & \text{if } t < T_1 \\ \lambda_{i(2)} & \text{if } t \geq T_1 \end{cases}$$

for some $T_1 \in (1, T)$. This is the factor model with a single structural break at the common break date T_1 for all individuals, and is the alternative considered by Breitung and Eickmeier (2011), Chen et al. (2014), Han and Inoue (2015), and Cheng et al. (2016). Apparently, this is a very restrictive alternative. In contrast, our model under the alternative allows for multiple structural breaks, with possibly unknown break dates or unknown number of breaks. More importantly, by assuming $\lambda_i(\cdot)$ to be a piece-wise smooth function, we allow for smooth structural changes in the factor loadings. This type of alternative appears more reasonable and realistic than the single structural break alternative given the fact that the driving forces of structural changes such as preference changes, technological progress and policy modifications take effect gradually over a long horizon.

2.2 Test Statistic

Under the null hypothesis, we can follow Bai and Ng (2002) and Bai (2003) and apply the method of PCA to estimate the following model

$$X_{it} = \lambda_{i0}' F_t + e_{it}^{\dagger}, \tag{2.2}$$

where $e_{it}^{\dagger} = e_{it}$ under $\mathbb{H}_0$ and the two are distinct under $\mathbb{H}_1$.

Let $X_t \equiv (X_{1t}, \ldots, X_{Nt})'$, $e_t \equiv (e_{1t}, \ldots, e_{Nt})'$, $e_t^{\dagger} \equiv (e_{1t}^{\dagger}, ..., e_{Nt}^{\dagger})'$, $F \equiv (F_1, \ldots, F_T)'$, and $\Lambda_0 \equiv (\lambda_{10}, ..., \lambda_{N0})'$. Let $X \equiv (X_1, ..., X_T)'$, $e \equiv (e_1, \ldots, e_T)'$, $e^{\dagger} \equiv (e_1^{\dagger}, ..., e_T^{\dagger})'$. Then we can rewrite (2.2) in matrix form

$$X = F\Lambda_0' + e^{\dagger}.$$

The PCA method solves the following minimization problem:

$$\min_{F,\Lambda} \operatorname{tr} (X - F\Lambda_0') (X - F\Lambda_0')' = \sum_{i=1}^{N} \sum_{t=1}^{T} \left(X_{it} - \lambda_{i0}' F_t \right)^2$$

under certain identification restrictions. In this paper, we follow Bai (2003) and consider the following identification restrictions:

$$T^{-1}F'F = \mathbb{I}_R \text{ and } \Lambda_0'\Lambda_0 \text{ is a diagonal matrix.}$$

Let $\tilde{F}_t$ and $\tilde{\lambda}_{i0}$ be the principal component estimators of F_t and λ_{i0} respectively under the above identification restrictions. Let $\tilde{F} = (\tilde{F}_1, ..., \tilde{F}_T)'$ and $\tilde{\Lambda}_0 = (\tilde{\lambda}_{10}, ..., \tilde{\lambda}_{N0})'$. It is well known that $\tilde{F}$ is $\sqrt{T}$ times eigenvectors corresponding to the R largest eigenvalues of the $T \times T$ matrix XX', and $\tilde{\Lambda}_0' = (\tilde{F}'\tilde{F})^{-1}\tilde{F}'X = T^{-1}\tilde{F}'X$.

After obtaining the restricted estimators $\tilde{F}_t$ and $\tilde{\lambda}_{i0}$ of F_t and λ_{i0}, we now consider the following nonparametric regression model:

$$X_{it} = \lambda_i \left(\frac{t}{T}\right)' \tilde{F}_t + e_{it}^{\ddagger}, \tag{2.3}$$

where $e_{it}^{\ddagger}$ is another error term that takes into account the estimation error introduced by replacing F_t with $\tilde{F}_t$. The intuition of our test goes as follows: if $\mathbb{H}_0$ is true, then any nonparametric estimate of $\lambda_i(\cdot)$ in (2.3) should not differ much from the restricted estimate $\tilde{\lambda}_{i0}$. However, if $\mathbb{H}_0$ is false, a typical nonparametric estimate of $\lambda_i(\cdot)$ can deviate a lot from the restricted estimate $\tilde{\lambda}_{i0}$. Hence, we can test $\mathbb{H}_0$ by measuring the distance between a typical nonparametric estimate of $\lambda_i(\cdot)$ and the restricted estimate $\tilde{\lambda}_{i0}$.

In this paper, we consider the simple local constant estimate of $\lambda_i\left(\frac{t}{T}\right)$ in (2.3). Let h be the bandwidth and $K(\cdot)$ be a kernel function with compact support $[-1,1]$. To avoid the boundary bias problem, we follow Hong and Li (2005) and Li and Racine (2007, p.31) and apply the following boundary kernel:

$$k_{h,tr} = h^{-1}K_r\left(\frac{t-r}{Th}\right) = \begin{cases} h^{-1}K\left(\frac{t-r}{Th}\right) / \int_{-r/(Th)}^{1} K(u)du, & \text{if } r \in [0, Th) \\ h^{-1}K\left(\frac{t-r}{Th}\right), & \text{if } r \in [Th, T-Th] \\ h^{-1}K\left(\frac{t-r}{Th}\right) / \int_{-1}^{(1-r/T)/h} K(u)du, & \text{if } r \in (T-Th, T] \end{cases}.$$

We note that $K_r(\cdot)$ equals to $K(\cdot)$ in the interior region but not in the boundary regions. The local constant estimator of $\lambda_i\left(\frac{t}{T}\right)$ is given by:

$$\hat{\lambda}_{it} = \hat{\lambda}_i\left(\frac{t}{T}\right) = \left(\frac{1}{T}\sum_{s=1}^{T} k_{h,st}\tilde{F}_s\tilde{F}_s'\right)^{-1} \frac{1}{T}\sum_{s=1}^{T} k_{h,st}\tilde{F}_s X_{is}. \tag{2.4}$$

Under $\mathbb{H}_0$, we have $\lambda_{it} = \lambda_{i0}$ for all t. $\hat{\lambda}_{it}$ will be close to $\tilde{\lambda}_{i0}$ for each t. Under $\mathbb{H}_1$, λ_{it} is not a constant over time and we would expect large deviations of $\hat{\lambda}_{it}$ from $\tilde{\lambda}_{i0}$ for some t. Therefore, we could test $\mathbb{H}_0$ by measuring the squared distance between $\hat{\lambda}_{it}$ and $\tilde{\lambda}_{i0}$:

$$\hat{M} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\hat{\lambda}_{it} - \tilde{\lambda}_{i0}\right)'\left(\hat{\lambda}_{it} - \tilde{\lambda}_{i0}\right). \tag{2.5}$$

The test statistic is a standardized version of $\hat{M}$ in (2.5):

$$\widehat{SM}_{NT} = \hat{\mathbb{V}}_{NT}^{-1/2}\left(TN^{1/2}h^{1/2}\hat{M} - \hat{\mathbb{B}}_{NT}\right), \tag{2.6}$$

where the centering factor $\hat{\mathbb{B}}_{NT}$ and the scaling factor $\hat{\mathbb{V}}_{NT}$ are defined as follows:

$$\begin{aligned} \hat{\mathbb{B}}_{NT} &= \frac{h^{1/2}}{T^2N^{1/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\tilde{F}_s'\left(k_{h,st}S_{Tt}^{-1} - \mathbb{I}_R\right)\left(k_{h,st}S_{Tt}^{-1} - \mathbb{I}_R\right)\tilde{F}_s\tilde{e}_{is}^2, \\ \hat{\mathbb{V}}_{NT} &= 2T^{-2}N^{-1}h^{-1}\sum_{1\le s\ne r\le T}\bar{K}\left(\frac{s-r}{Th}\right)^2 \tilde{F}_r'\tilde{\Sigma}_F^{-1}\tilde{\Sigma}_F^{-1}\tilde{F}_s\tilde{F}_s'\tilde{\Sigma}_F^{-1}\tilde{\Sigma}_F^{-1}\tilde{F}_r\left(\tilde{e}_r'\tilde{e}_s\right)^2, \end{aligned}$$

with $S_{Tt} = \frac{1}{T}\sum_{r=1}^{T} k_{h,rt}\tilde{F}_r\tilde{F}_r'$, $\tilde{\Sigma}_F = T^{-1}\sum_{t=1}^{T}\tilde{F}_t\tilde{F}_t'$, $\tilde{e}_{is} = X_{is} - \tilde{\lambda}_{i0}'\tilde{F}_s$, $\tilde{e}_t = (\tilde{e}_{1t}, \ldots, \tilde{e}_{Nt})'$, and $\bar{K}(u) = \int_{-1}^{1} K(v) K(u-v)\, dv$ being the two-fold convolution kernel of $K(\cdot)$. For example, if we use the Epanechnikov kernel $K(u) = 0.75(1-u^2)\mathbf{1}(|u| \le 1)$, then $\bar{K}(u) = (\frac{3}{5} - \frac{3}{4}u^2 + \frac{3}{8}|u|^3 - \frac{3}{160}|u|^5)\mathbf{1}(|u| \le 2)$, where $\mathbf{1}(\cdot)$ is the usual indicator function.

3 Asymptotic Properties of the Test Statistic

In this section, we will establish the asymptotic null distribution of our test and study its asymptotic local power property. In addition, we also propose a bootstrap procedure to improve the finite sample performance of the test and establish its asymptotic validity.

3.1 Assumptions

Let $\gamma_N(s,t) = N^{-1}E(e_s'e_t)$, $\xi_{st} = N^{-1}[e_s'e_t - E(e_s'e_t)]$, $\gamma_{N,FF}(s,t) = N^{-1}E(F_se_s'e_tF_t')$ and $\tau_{ij,st} = E(e_{it}e_{js}F_t'F_s)$. We use $\max_i$, $\max_t$, $\max_{i,t}$ and $\max_{s,t}$ to denote $\max_{1\le i\le N}$, $\max_{1\le t\le T}$, $\max_{1\le i\le N}\max_{1\le t\le T}$ and $\max_{1\le s,t\le T}$, respectively. Throughout, we make the following assumptions.

Assumption A.1 [Factors]

(i) $E(F_tF_t') = \Sigma_F$ $\forall t$ for some $R\times R$ positive definite matrix Σ_F.

(ii) $\max_t E\|F_t\|^{8+\sigma} < \infty$ for some $\sigma > 0$.

Assumption A.2 [Factor Loadings]

(i) λ_{i0} are nonrandom such that $\max_{1\le i\le N}\|\lambda_{i0}\| \le C$.

(ii) $N^{-1}\Lambda_0'\Lambda_0 = N^{-1}\sum_{i=1}^{N}\lambda_{i0}\lambda_{i0}' \to \Sigma_{\Lambda_0}$ for some $R\times R$ positive definite matrix Σ_{Λ_0}.

(iii) The eigenvalues of the $R\times R$ matrix $\Sigma_F\Sigma_{\Lambda_0}$ are distinct.

Assumption A.3 [Error term]

(i) $E(e_{it}) = 0$, $\max_{i,t} E|e_{it}|^{8+\sigma} \le C$ and $\max_{i,t} E\|F_te_{it}\|^{8+4\sigma} \le C$ for some $\sigma > 0$.

(ii) For each $i = 1, 2, \ldots, N$, the process $\{e_{it}, t = 1, 2, \ldots\}$ is a martingale difference sequence (m.d.s.) with respect to $\mathcal{F}_{NT,t}$: $E(e_t|\mathcal{F}_{NT,t-1}) = 0$, where $\mathcal{F}_{NT,t-1}$ is the σ-field generated from $(F_t, F_{t-1}, \ldots, e_{t-1}, e_{t-2}, \ldots)$.

(iii) For each $i = 1, 2, \ldots, N$, the process $\{(e_{it}, F_t), t = 1, 2, \ldots\}$ is strong mixing with mixing coefficients $\alpha_i(\cdot)$. $\alpha(\cdot) \equiv \max_i \alpha_i(\cdot)$ satisfies $\sum_{s=1}^{\infty}\alpha(s)^{\delta/(2+\delta)} \le C$ for some $\delta > 0$. In addition, there exists an integer $T_0 \in [1, T)$ such that $T^{-2}\max(T_0^4, T_0^3h^{-1}, T_0^2h^{-2}) \to 0$ and $N^2Th^2\alpha(T_0)^{\delta/(1+\delta)} \to 0$ as $(N,T) \to \infty$.

(iv) $\max_t \sum_{s=1}^{T}|\gamma_N(s,t)| \le C$, $\max_{s,t} E\left|N^{1/2}\xi_{st}\right|^4 \le C$, $\max_t E|N^{-1/2}\sum_{i=1}^{N}[e_{it}^2 - E(e_{it}^2)]|^4 \le C$.

(v) $\max_t \sum_{s=1}^{T}\left|\gamma_{N,FF}(s,t)\right| \le C$, $\max_{t\ne r} E\|N^{-1/2}F_te_t'e_rF_r'\|^4 \le C$, and $N^{-1}T^{-1}\sum_{i,j=1}^{N}\sum_{s,t=1}^{T}|\tau_{ij,st}| \le C$.

(vi) $\|e\|_{\text{sp}} = O_P(N^{1/2} + T^{1/2})$.

Assumption A.4 [Kernel function and Bandwidth]

(i) The kernel function $K : \mathbb{R} \to \mathbb{R}^+$ is symmetric and continuously differentiable probability density function with compact support $[-1, 1]$.

(ii) As $(N,T) \to \infty$, $h \to 0$, $Th^2 \to \infty$, $Th^2/N^3 \to 0$, $Nh^2/T \to 0$, $Th(\ln T)^{-2} \to \infty$, $Nh^2(\ln T)^{-4} \to \infty$, $T^2N^{-1}h^3(\ln T)^{-6} \to \infty$.

Assumption A.1 imposes some conditions on the latent common factors. We follow Stock and Watson (2002), Breitung and Eickmeier (2011), Chen et al. (2014), Han and Inoue (2015) and Su and Wang (2017) and assume that $E\left(F_tF_t'\right) = \Sigma_F$ is homogeneous over t. Motta et al. (2011) make a much stronger assumption: $F_t \sim i.i.d.(0, \Sigma_F)$ with Σ_F being diagonal and positive definite. Assumption A.1(i) greatly facilitates the derivation of the asymptotic results and can be regarded as an identification condition. Since the latent common factors and the factor loadings are not separately identifiable, it is difficult to distinguish the structural changes on factor loadings with those on common factors. This explains why researchers have frequently made some normalization restrictions like $E(F_tF_t') = \Sigma_F$ in the literature. Otherwise, even if there is no structural change on the second moment of F_t such that $E(F_tF_t') = \Sigma_F$ is satisfied for all t and the factor loadings λ_i are constant over time, we can always write

$$\lambda_i'F_t = \lambda_i'Q\left(t/T\right)^{-1}Q\left(t/T\right)F_t = \lambda_{it}^{*\prime}F_t^* \text{ for any nonsingular matrix } Q\left(t/T\right),$$

where $F_t^* = Q\left(t/T\right)F_t$ and $\lambda_{it}^* = [Q\left(t/T\right)^{-1}]'\lambda_i$. That is, $\lambda_i'F_t$ can be equivalently rewritten as the inner product of a time-varying factor loading λ_{it}^* with a factor F_t^* that has time-varying second moment. [Note that $E(F_t^*F_t^{*\prime}) = Q(t/T)\Sigma_FQ(t/T)'$]. The constancy of the second moment of the factor in Assumption A1(i) aims to rule out this situation even though it is still not sufficient to identify the factors and factor loadings under the alternative.

Assumption A.2 ensures that each factor has a nontrivial contribution to the variance of X_t. Following Bai (2003) and Breitung and Eickmeier (2011), we assume that the factor loadings are nonrandom for simplicity.

A.3 imposes moment conditions on the errors and their interactions with the factors and factor loadings. A.3(i) and (iv) correspond to Assumptions C.1 and C.5 in Bai (2003). A.3(ii) assumes that the process $\{e_{it}, t = 1, 2, ...\}$ is an m.d.s. with respect to the filter $\{\mathcal{F}_{NT,t}\}$ and it allows for cross-sectional dependence among the error terms. This assumption is essential for the establishment of the asymptotic distribution of our test statistic under the null hypothesis and a sequence of Pitman local alternatives. It is possible to allow for both serial dependence and cross-sectional dependence in the error terms. But that will substantially complicate the asymptotic analysis and we are not sure how to estimate the asymptotic variance of our raw test statistic in this case. A.3(iii) requires the process $\{(e_{it}, F_t), t = 1, 2, ...\}$ to be strong mixing with some algebraic mixing rate. With more complicated notation, one can allow different individual time series to have different mixing rates and then relax the summability mixing condition to $\limsup_N \frac{1}{N}\sum_{i=1}^{N}\sum_{s=1}^{\infty}\alpha_i\left(s\right)^{\delta/(1+\delta)} \leq C < \infty$. If the processes are strong mixing with a geometric rate (e.g., $\alpha\left(s\right) = \rho^s$ for some $\rho \in [0, 1)$), then the conditions on $\alpha\left(\cdot\right)$ can be all met by specifying $T_0 = \lfloor C_0 \ln T \rfloor$ for some sufficiently large positive constant C_0. A.3(iv) and (v) control the cross-sectional dependence among $\{e_{it}, i = 1, 2, \ldots, N\}$ and $\{F_te_{it}, i = 1, 2, \ldots, N\}$, respectively. A.3(vi) is widely assumed in the factor literature; see, e.g., Moon and Weidner (2015), Su and Wang (2017), and Ma and Su (2018).

A.4 imposes regularity conditions on the kernel and bandwidth. The familiar positive bounded kernels, such as the Epanechnikov, Quartic and Uniform kernels, are allowed. However, it rules out the Gaussian kernel, which has unbounded support. We allow the choice of a wide range of admissible rates for bandwidth h. For example, if N and T are the same order of magnitude as in many applications, one can specify $h \propto T^{-a}$ for $0 < a < 1/3$. Thus the optimal rate of bandwidth $(T^{-1/5})$ in terms of minimizing the mean squared error of the nonparametric estimation for $\lambda_i(\cdot)$ would satisfy A4(ii) in this case even though it is typically not the optimal bandwidth for our test. Moreover, Assumption A.4 also allows for a wide range of admissible relative magnitudes of N and T. One can specify $h \propto T^{-a}$ and $N \propto T^{b}$ for $\max\{2a, \frac{1}{3}(1-2a)\} < b < \min\{1+2a, 2-3a\}$. For example, if $a = \frac{1}{5}$, then $\frac{2}{5} < b < \frac{7}{5}$. This includes the most common scenario in applications where N and T pass to infinity at the same rate.

The nonparametric regression we used in the second step is the time-varying coefficient time series model given by Cai (2007). Following the analysis in Cai (2007), we can show that the asymptotic bias of the estimator of $\lambda_i(t/T)$ is $O(h^2)$ and the asymptotic variance is $O(T^{-1}h^{-1})$. Therefore, a popular rule-of thumb procedure is to choose $h = c_h \sigma_s T^{-1/5}$, where σ_s is the sample standard deviation of the smooth variable and c_h is a constant depending on the kernel in use. For the Epanechnikov kernel, $c_h = 2.35$. Here, the smooth variable $\{t/T\}_{t=1}^{T}$ behaves like a uniform random variable on $[0,1]$ and thus one can set $\sigma_s = 1/\sqrt{12}$. Therefore, we use the bandwidth $h = (2.35/\sqrt{12})T^{-1/5}$ as the benchmark bandwidth in our simulations and check the effect of different bandwidth sequences by setting $h = c(2.35/\sqrt{12})T^{-1/5}$ for $c = 0.5, 1.5$ in the online supplement.

In practice, one could also consider a data-driven bandwidth using the leave-one-out cross-validation (CV) method. That is, we can choose h as

$$h = \arg \min_{c_1 n^{-\gamma} \leq h \leq c_2 n^{-\gamma}} \sum_{i=1}^{N} \sum_{t=1}^{T} [X_{it} - \lambda_{i,-}\left(\frac{t}{T}\right) \tilde{F}_t]^2$$

where $\lambda_{i,-}\left(\frac{t}{T}\right) = \left(\frac{1}{T}\sum_{s\neq t} k_{h,st}\tilde{F}_s\tilde{F}_s'\right)^{-1} \frac{1}{T}\sum_{s\neq t} k_{h,st}\tilde{F}_s X_{is}$ is the leave-one-out estimator, $\gamma = 1/5$ and $0 < c_1 < c_2 < \infty$ are two pre-specified constants. Although the above cross-validated bandwidth is asymptotically optimal for the estimation of the time-varying nonparametric regression model in terms of mean squared error, it is not optimal for our test. For testing problems, the essential concern is the Type I and Type II errors. Based on the Edgeworth expansion of the asymptotic distribution of a test statistic in a different but related nonparametric context, Gao and Gijbels (2008) show that the choice of h affects both the Type I and Type II errors, and usually there exists a tradeoff between these two. A sensible optimal rule is to choose h to maximize the power of a test given a significant level. Gao and Gijbels (2008) derive the leading terms of the size and power functions of their test and then choose a bandwidth to maximize the power under a class of local alternatives with a controlled significance level. Unfortunately, Gao and Gijbels's (2008) results cannot be directly applied to our test, because the higher order terms of size and power functions depend on the form of test statistic, the DGP, the kernel and the bandwidth, among other things. In another different but related context, Sun, Phillips and Jin (2008) also consider a data-driven bandwidth by minimizing a weighted average of the Type I and Type II errors of a test. It is possible to extend these approaches to our test, but the analytical expressions for the leading terms of the size and power functions or the two type errors of our test are rather involved and is beyond the scope of the present paper. We will pursue this important issue in a subsequent study.

3.2 Asymptotic Null Distribution

Under the above regularity conditions, we now state the asymptotic distribution of $\widehat{SM}_{NT}$ under $\mathbb{H}_0$.

Theorem 3.1 *Suppose Assumptions A.1-A.4 hold. Then* $\widehat{SM}_{NT} \xrightarrow{d} N(0,1)$ *under* $\mathbb{H}_0$.

Remark 1. The test statistic is based on a sample quadratic form, which measures the distance between the local smoothing estimator $\hat{\lambda}_{it}$ and the principal component estimator $\tilde{\lambda}_{i0}$. Under $\mathbb{H}_0$, $\tilde{\lambda}_{i0}$ converges to the true factor loadings coupled with an unknown rotation matrix with a faster rate than that of the local smoothing estimator $\hat{\lambda}_{it}$. As a result, the limiting behavior of $\widehat{SM}_{NT}$ is solely determined by $\hat{\lambda}_{it}$. In particular, by subtracting the bias term, the quadratic form statistic yields a dominant degenerate U-statistic, which determines the asymptotic distribution of our test. Since a large value of $\hat{M}$ is in favor of the alternative, our test is a one-sided test.

Remark 2. The test is asymptotically pivotal and has a convenient asymptotic $N(0,1)$ distribution under $\mathbb{H}_0$. Consequently, we can compare our test statistic with the one-sided $N(0,1)$ critical value z_α at the significance level α, and reject $\mathbb{H}_0$ when $\widehat{SM}_{NT} > z_\alpha$. In contrast, the limiting distributions of the existing tests for structural changes with unknown break date, namely the supremum-type tests of Breitung and Eickmeier (2011), Chen et al. (2014), Han and Inoue (2015), rely on a tied-down Bessel process, which depends on a prespecified trimming parameter and the degree of freedom. As a result, one should either simulate or refer to Andrews' (1993) tabulated critical values.

3.3 Asymptotic Local Power

To gain more insight into the asymptotic power property of the test $\widehat{SM}_{NT}$, we now consider a class of local alternatives as follows:

$$\mathbb{H}_1(a_{NT}) : \lambda_{it} = \lambda_{i0} + a_{NT} g_i\left(\frac{t}{T}\right) \text{ for each } i \text{ and } t,$$

where $a_{NT} \to 0$ as $(N,T) \to \infty$. a_{NT} controls the speed at which the local alternative converges to the null hypothesis, and $g_i\left(\frac{t}{T}\right)$ is a piecewise smooth function with a finite number of discontinuity points. Noting that $\lambda_{i0} + a_{NT} g_i\left(\frac{t}{T}\right) = (\lambda_{i0} + c_{i,NT}) + a_{NT}[g_i\left(\frac{t}{T}\right) - c_{i,NT}/a_{NT}]$ for any $c_{i,NT} \in \mathbb{R}^R$, below we will assume that

$$\int_0^1 g_i(u)\, du = 0$$

for location normalization purpose. It turns out this normalization will greatly simplify the local asymptotic power analysis. With such a normalization, both λ_{i0} and $g_i(\cdot)$ can be dependent on the sample sizes N and T. But for notational simplicity, we continue to write them as λ_{i0} and $g_i(\cdot)$ instead of $\lambda_{i0,NT}$ and $g_{i,NT}(\cdot)$.

To study the asymptotic power property of $\widehat{SM}_{NT}$, we add the following assumption:

Assumption A.5

(i) For each $i = 1, 2, \ldots, N$, $g_i(\cdot)$ is piecewise continuous with a finite number of discontinuous points on (0,1]. $\max_{1\le i\le N} \sup_u |g_i(u)| \le C$.

(ii) $\max_{1\le r\le T} \left\| \frac{1}{NT} \sum_{s=1}^{T} \sum_{i=1}^{N} k_{h,sr} F_s e_{is} g_{ir}' \right\| = O_P\left((NTh/\ln(NT))^{-1/2}\right)$ where $g_{ir} = g_i(r/T)$.

A.5(i) allows for both sudden breaks and smooth changes under the local alternative. A.5(ii) can be verified as in Su et al. (2015).

The following theorem studies the asymptotic local power property of $\widehat{SM}_{NT}$.

Theorem 3.2 *Suppose that Assumptions A.1-A.5 hold. Then under* $\mathbb{H}_1(a_{NT})$ *with* $a_{NT} = T^{-1/2}N^{-1/4}h^{-1/4}$,

$$P\left[\widehat{SM}_{NT} \ge z_\alpha | \mathbb{H}_1(a_{NT})\right] \to 1 - \Phi(z_\alpha - \pi_0)$$

as $(N,T) \to \infty$, *where* $\Phi(\cdot)$ *is the standard normal CDF,* z_α *is the one-sided normal critical value at the significance level* α, *and* $\pi_0 = \lim_{(N,T)} \Pi_{1NT}/\mathbb{V}_{NT}^{1/2}$ *with*

$$\begin{aligned}\Pi_{1NT} &= tr\left(\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} Q_0 g_i\left(\frac{t}{T}\right) g_i\left(\frac{t}{T}\right)' Q_0'\right) \text{ and} \\ \mathbb{V}_{NT} &= 2T^{-2}N^{-1}h^{-1}\sum_{1\le s\neq r\le T}\bar{K}\left(\frac{s-r}{Th}\right)^2 E\left[F_r'\Sigma_F^{-1}Q_0'Q_0\Sigma_F^{-1}F_sF_s'\Sigma_F^{-1}Q_0'Q_0\Sigma_F^{-1}F_r(e_r'e_s)^2\right],\end{aligned}$$

where $Q_0 = V_0^{1/2}\Upsilon_0'\Sigma_{\Lambda_0}^{-1/2}$, V_0 *is an* $R\times R$ *diagonal matrix containing the* R *largest eigenvalues of* $\Sigma_{\Lambda_0}^{1/2}\Sigma_F\Sigma_{\Lambda_0}^{1/2}$ *in decreasing order, and* Υ_0 *is the corresponding eigenvector matrix such that* $\Upsilon_0'\Upsilon_0 = \mathbb{I}_R$.

Remark 3. Theorem 3.2 shows that the test has nontrivial power against $\mathbb{H}_1(a_{NT})$ with $a_{NT} = T^{-1/2}N^{-1/4}h^{-1/4}$. Although the test shares the same convergence rate with Su and Wang (2017) theoretically, the test has better finite sample performance for most cases (see Section 4). We conjecture that this is mainly due to the fact that our test focuses exclusively on the time variation in factor loadings while Su and Wang's (2017) test is based on the contrast between the estimates of the common components under the null and alternative hypotheses, respectively. Note that Assumption A.5 allows the presence of a finite number of unknown discontinuity points in factor loadings. As a result, the test is powerful in detecting smooth structural changes as well as sudden structural breaks, with possibly unknown break dates or unknown number of breaks in the factor loadings. In addition, for the sample size (N,T) sufficiently large, $\widehat{SM}_{NT}$ can detect any fixed structural changes that occur close to the starting and ending points of the sampling period, because no trimming is required for out test. This is rather appealing because no prior information about the alternative is available in practice. It avoids blind searches of possible alternatives of structural changes. In contrast, the tests of Breitung and Eickmeier (2011), Chen et al. (2014), Han and Inoue (2015), Yamamoto and Tanaka (2015) and Cheng et al. (2016), all rely on a prespecified trimming parameter ϵ to trim out the first and last ϵT observations in the sample and hence would miss the possible structural changes in the boundary regions.

Remark 4. To ensure our test to have non-trivial power against the $T^{-1/2}N^{-1/4}h^{-1/4}$-rate local alternatives, we need $\pi_0 > 0$, which would require $\Pi_0 \equiv \lim_{(N,T)} \Pi_{NT} > 0$ as one can show that the limit of $\mathbb{V}_{NT}$ is bounded away from 0. This requires that the factor loadings should not be time-varying only for an asymptotically negligible set of individuals or time periods. Let $\mathcal{N} = \{1, 2, ..., N\}$ and $\mathcal{T} = \{1, 2, ..., T\}$. Let $|\cdot|$ denote the cardinality of a set $\cdot$. Define the following subsets of $\mathcal{N}$ and $\mathcal{T}$:

$$\mathcal{S}_N = \{i \in \mathcal{N} : \lambda_{it} = \lambda_{i0} \text{ for all } t\} \text{ and } \mathcal{S}_{Ti} = \{t \in \mathcal{T} : g_{it} = 0\}.$$

Let $\mathcal{S}_N^c = \mathcal{N}\backslash\mathcal{S}_N$, the complement of $\mathcal{S}_N$ relative to $\mathcal{N}$. Define $\mathcal{S}_{Ti}^c = \mathcal{T}\backslash\mathcal{S}_{Ti}$ analogously. It is easy to verify that if either

$$|\mathcal{S}_N^c|/N = o(1) \text{ or } \max_{1\le i\le N}|\mathcal{S}_{Ti}^c|/T = o(1),$$

then $\Pi_0 = 0$, and our test does not have power against the $T^{-1/2}N^{-1/4}h^{-1/4}$-rate local alternatives in this case. Similar phenomenon occurs in Su and Chen's (2013) test for slope homogeneity and Su and Wang's (2017) test for structural changes in factor loadings. In general, as long as a fixed proportional of individuals

$\mathcal{N}$ either undergo abrupt structural change (one break or multiple breaks), or have a fixed proportion of T periods of smooth structural changes, $\Pi_0 > 0$ and our test has asymptotic power to detect them.

Remark 5. We note that, by assuming $a_i(\cdot)$ to be a piecewise smooth function with a finite number of discontinuity points, we allow various types of local alternatives, including the one-time structural breaks, the multiple abrupt changes, and smooth structural changes. The case of one-time structural breaks overlaps with the alternative hypothesis considered by some parametric tests given by Breitung and Eickmeier (2011), Chen et al. (2014) and Han and Inoue (2015). To avoid the comparison between two large dimensional factor loadings matrix estimates, the previous parametric tests all reduce the infinite dimensional problem to a finite dimensional one in different ways. For example, Breitung and Eickmeier (2011) propose three test statistics based on certain time series regressions for each cross sectional unit i; Chen et al. (2014) run the regression of one estimated factor on the remaining ones and then test for the structural changes in such a linear regression by constructing the sup-Wald and sup-LM statistics of Andrews (1993); Han and Inoue (2015) construct their sup-Wald and sup-LM statistics by comparing the pre- and post-break subsample second moments of the estimated factors. In any case, the test statistics have the same asymptotic distribution and the same convergence rate as the conventional sup-Wald statistic of Andrews (1993). As a result, they could only detect local alternatives that converge to the null at the rate $T^{-1/2}$, which is slower than the rate a_{NT} by noticing that $Nh \to \infty$ under our assumptions.

Moreover, we want to mention that, for some types of structural changes that are not identifiable under the alternative, all of existing tests including our test will have no power to detect them. For example, consider a specific structural change process for the factor loadings: $\lambda_{it} = Q(t/T)\lambda_{i0}$, where $Q(u)$ is an $R \times R$ orthogonal matrix for all $u \in [0,1]$ with $Q(0) = I_R$. Then $E(X_t X_t') = \Lambda_t E(F_t F_t')\Lambda_t' + \Sigma_e = \Lambda_t \Lambda_t' + \Sigma_e = \Lambda_0 Q(t/T)' Q(t/T)\Lambda_0' + \Sigma_e = \Lambda_0\Lambda_0' + \Sigma_e$ for $t = 1, \cdots, T$. Thus, the time path $\{Q(u)\}_{u\in(0,1]}$ is entirely unidentified, and no structural break test can have nontrivial power against this specific class of time-varying alternatives. This occurs mainly because the orthogonal matrix $Q(t/T)$ is not heterogeneous across i and we can rewrite $\lambda_{it}' F_t$ as $\lambda_{i0}' F_t^*$ with $F_t^* = Q(t/T)' F_t$. In this case, the conventional PCA estimator of the factor is consistent with a rotational version of F_t^* instead of F_t. However, if the orthogonal function $Q(u)$ exhibits individual heterogeneity, say, $\lambda_{it} = Q_i(t/T)\lambda_{i0}$, we cannot associate $Q_i(t/T)$ as a part of the factor any more and our test still has power to detect such deviations from the null hypothesis.

Remark 6. The exact number R of common factors is typically unknown in practice and one should determine the number of common factors before estimating and testing. This is not actually a concern under the null hypothesis because many popular methods such as those of Bai and Ng (2002), Ahn and Horenstein (2013) and Onatski (2009, 2010) could estimate the number of common factors consistently. Unfortunately, these methods typically break down under the alternative. One exception is Su and Wang's (2017) local-PCA-based information criterion that proves to work under both the null and alternative hypotheses. So we recommend the use of Su and Wang's (2017) local-PCA-based information criterion to determine the number of factors. Of course, in many applications, applied researchers may have a strong prior on the reasonable number of factors to be included into the model (say, $R \leq 4$), and one can also conduct our nonparametric test for each of these prior values. The presence of smooth structural changes can typically cause the rejection of the null. In any case, as a referee remarks, a model with few factors and time-varying loadings can be a more parsimonious, useful, and interpretable model than a conventional factor model with a very large number of factors and constant factor loadings. So it is worthwhile to explore the time-varying factor models as advocated by Su and Wang (2017) once one rejects the null.

3.4 A Bootstrap Version of the Test

Since kernel-based nonparametric tests may not perform well in finite samples and they can be sensitive to the choice of bandwidth, we propose a bootstrap procedure to improve the finite sample performance of our test.

As mentioned in Su and Wang (2017), the wild bootstrap works well if the error terms $\{e_{it}\}$ do not exhibit cross-sectional dependence or only exhibit fairly weak cross-sectional dependence, but it tends to be oversized in the presence of moderate or strong cross-sectional dependence in the error terms. Hence, we follow Su and Wang (2017) and propose a modified parametric bootstrap procedure that tries to mimic the cross-sectional dependence in $\{e_{it}\}$. Let $e_t = (e_{1t}, ..., e_{Nt})'$, $\tilde{e}_t = (\tilde{e}_{1t}, ..., \tilde{e}_{Nt})'$, $\Sigma_e = \text{Var}(e_t) = \{\sigma_{e,ij}\}$, and $\tilde{\Sigma}^0 = T^{-1} \sum_{t=1}^{T} \tilde{e}_t \tilde{e}_t'$. Let $\tilde{\sigma}_{ij}^0$ denote the (i, j)th element of $\tilde{\Sigma}^0$. Define the shrinkage version of $\tilde{\Sigma}^0$ as $\tilde{\Sigma}$ whose (i, j)th element is given by

$$\tilde{\sigma}_{ij} = \tilde{\sigma}_{ij}^0 (1 - \epsilon)^{|j-i|} \text{ for } i, j = 1, ..., N,$$

where ϵ is a small positive number (e.g., 0.01) to ensure the maximum absolute column/row sum norm of $\tilde{\Sigma}$ to be stochastically bounded provided $\max_{i,j} \left|\tilde{\sigma}_{ij}^0\right|$ is. By construction, $\tilde{\Sigma}$ is also symmetric and positive semi-definite. The stochastic boundedness of $\max_{i,j} \left|\tilde{\sigma}_{ij}^0\right|$ is sufficient but not necessary for the justification of the asymptotic validity of our bootstrap procedure below:

1. Estimate the restricted model $X_{it} = \lambda_{i0}' F_t + e_{it}$ to obtain the principal component estimates $\tilde{\lambda}_{i0}$ and $\tilde{F}_t$ and the corresponding residuals $\tilde{e}_{it} = X_{it} - \tilde{\lambda}_{i0}' \tilde{F}_t$. Obtain the nonparametric kernel estimates $\hat{\lambda}_{it}$ and calculate the test statistic $\widehat{SM}_{NT}$ as in Section 2.2.

2. For $i = 1, 2, ..., N$ and $t = 1, 2, \ldots, T$, obtain the bootstrap error $e_t^* = \tilde{\Sigma}^{1/2} \varsigma_t$, where $\varsigma_t = (\varsigma_{1t}, \ldots, \varsigma_{Nt})'$ with ς_{it} being *i.i.d.* $N(0, 1)$ across i and t. Generate $X_{it}^* = \tilde{\lambda}_{i0}' \tilde{F}_t + e_{it}^*$ for $i = 1, 2, \ldots, N$ and $t = 1, 2, \ldots, T$.

3. Use $\{X_{it}^*\}$ to run the restricted model to obtain the bootstrap versions $\{\tilde{\lambda}_{i0}^*, \tilde{F}_t^*\}$ of $\{\tilde{\lambda}_{i0}, \tilde{F}_t\}$ respectively. Run X_{it}^* on $\tilde{F}_t^*$ to obtain the local constant estimate of $\hat{\lambda}_{it}$. Calculate the bootstrap test statistic $\widehat{SM}_{NT}^*$, the bootstrap version of $\widehat{SM}_{NT}$.

4. Repeat steps 2 and 3 for B times and index the bootstrap test statistics as $\{\widehat{SM}_{NT,l}^*\}_{l=1}^B$. The bootstrap p-value is calculated by $p^* \equiv B^{-1} \sum_{l=1}^{B} \mathbf{1}(\widehat{SM}_{NT,l}^* > \widehat{SM}_{NT})$.

The following theorem establishes the asymptotic validity of the above bootstrap method.

Theorem 3.3 *Suppose that Assumptions A.1-A.4 hold. Suppose that (i)* $\max_{i,j} |\tilde{\sigma}_{ij}^0| = O_P(\xi_{NT})$ *with* $\xi_{NT} = o(T^{1/2})$, *(ii)* $\frac{1}{T} \sum_{t=1}^{T} \left\|\tilde{F}_t\right\|^8 = O_P(1)$ *and (iii)* $\frac{1}{N} \sum_{i=1}^{N} \left\|\tilde{\lambda}_{i0}\right\|^8 = O_P(1)$. *Then* $\widehat{SM}_{NT}^* \xrightarrow{d^*} N(0, 1)$ *in probability, where* $\xrightarrow{d^*}$ *denotes weak convergence under the bootstrap probability measure conditional on the observed sample* $\mathcal{W}_{NT} \equiv \{X_{it}, i = 1, ..., N, t = 1, ..., T\}$.

Remark 7. Theorem 3.3 shows that the modified parametric bootstrap provides an asymptotic valid approximation to the limit null distribution of $\widehat{SM}_{NT}$. This holds as long as we generate the bootstrap data by imposing the null hypothesis. If the null hypothesis does not hold in the observed sample, then we expect $\widehat{SM}_{NT}$ to explode at the rate $T^{1/2}N^{1/4}h^{1/4}$, which delivers the consistency of the bootstrap-based test $\widehat{SM}^{*}_{NT}$.

Remark 8. Theorem 3.3 only establishes the first-order asymptotic validity of the bootstrap procedure. We cannot expect that the bootstrap delivers a second order asymptotic refinement relative to the asymptotic normal approximation. Note that the justification of Theorem 3.3 does not require $\tilde{\Sigma}$ to be consistent with the $N \times N$ variance-covariance matrix Σ_e of e_t in terms of spectral norm. In fact, due to the normalization nature of our test statistic, one does not need to mimic the exact structure in Σ_e. Even so, it is desirable to generate the bootstrap errors $\{e_t^*\}$ that share the variance-covariance structure as $\{e_t\}$ asymptotically. In principle, we can follow Fan et al. (2013, FLM hereafter) to obtain a consistent estimate of Σ_e in terms of spectral norm under some additional conditions. Let $\hat{\theta}_{ij} = \frac{1}{T}\sum_{t=1}^{T}[\tilde{e}_{it}\tilde{e}_{jt} - \hat{\sigma}_{ij}^0]^2$. Define $\hat{\Sigma}^{\mathcal{T}} = \left\{\hat{\sigma}_{ij}^{\mathcal{T}}\right\}$ with

$$\hat{\sigma}_{ij}^{\mathcal{T}} = \hat{\sigma}_{ij}^0 \mathbf{1}\,(i=j) + s_{ij}(\hat{\sigma}_{ij}^0)\mathbf{1}\,(i \neq j)\,,$$

where $s_{ij}(z) \equiv sgn(z)(|z| - \tau_{ij})_+$ is the soft thresholding function, $\tau_{ij} = C_0((N \wedge T)^{-1}\log T)^{1/2}(\hat{\theta}_{ij})^{1/2}$, and C_0 is a positive constant. Following the asymptotic analysis in FLM, if

$$\max_{1\le i\le N}\frac{1}{T}\sum_{t=1}^{T}(\tilde{e}_{it}-e_{it})^2 = O_P((N\wedge T)^{-1}\log T) \text{ and } \max_{1\le i\le N}\max_{1\le t\le T}|\tilde{e}_{it}-e_{it}| = o_P(1)\,, \tag{3.1}$$

then we can readily show that $\left\|\hat{\Sigma}^{\mathcal{T}} - \Sigma_e\right\|_{\mathrm{sp}} = O_P([(N\wedge T)^{-1}\log T]^{(1-\gamma_0/2)}) = o_P(1)$ provided that there exists some $\gamma_0 \in [0,1)$ such that

$$\max_{1\le i\le N}\sum_{i=1}^{N}|\sigma_{e,ij}|^{\gamma_0} \le C \text{ for some } C < \infty.$$

The last condition strengthens the typical weak cross-sectional dependence condition $\max_{1\le i\le N}\sum_{i=1}^{N}|\sigma_{e,ij}| = O(1)$ and can be met if e_{it} satisfies certain m-dependence condition cross-sectionally or the correlation between e_{it} and e_{jt} shrinks to zero sufficiently fast as the "distance" between i and j, perhaps after re-ordering the cross-sectional units, increases. The fundamental problem is that we cannot verify the two conditions in (3.1) under the global alternative despite the fact that they can verified under the local alternatives. For this reason, we do not generate e_t^* as $(\hat{\Sigma}^{\mathcal{T}})^{1/2}\varsigma_t$ in our bootstrap procedure. Even if we generate the bootstrap errors from $(\hat{\Sigma}^{\mathcal{T}})^{1/2}\varsigma_t$ and restrict our attention to the local alternatives, we are not sure whether the bootstrap inference can achieve any refinement over the inference based on the asymptotic normal distribution. In fact, to the best of our knowledge, there is no formal study on the bootstrap refinement in the factor literature even for the inference on single factors or factor loadings. Our nonparametric test is involved with the contrast of the factor loadings estimates for all cross-sectional units under the null and alternative. A formal higher order refinement study that involves Edgeworth expansions would become much harder and thus be beyond the scope of this paper.

4 Monte Carlo Simulations

In this section, we study the finite sample performance of the test through Monte Carlo simulations. We also compare our test with the parametric tests of Breitung and Eickmeier (2011), Chen et al. (2014) and Han and Inoue (2015) for a single structural break with an unknown break date in the factor loadings and the nonparametric test of Su and Wang (2017) that also allows for both single or multiple abrupt breaks and smooth changes under the alternative.

4.1 Data Generating Processes

We generate data under the framework of large factor models with $R = 2$ common factors:

$$X_{it} = \lambda_{it}' F_t + e_{it},$$

where $i = 1, ..., N$, $t = 1, ..., T$, $F_t \equiv \ (F_{1,t}, F_{2,t})'$, with $F_{1,t} = 0.6F_{1,t-1} + u_{1t}$, $u_{1t} \sim i.i.d.$ $\mathrm{N}(0, 1 - 0.6^2)$; $F_{2,t} = 0.3F_{2,t-1} + u_{2t}$, $u_{2t} \sim i.i.d.$ $\mathrm{N}(0, 1 - 0.3^2)$.

To examine the size and power, we consider the following setups for the factor loading $\lambda_{it} \equiv \ (\lambda_{it,1}, \lambda_{it,2})'$ and the error term e_{it}:

DGP.S1: $\lambda_{it} = \lambda_{i0} \sim i.i.d.$ $\mathrm{N}(0, \mathbb{I}_2)$, $e_{it} \sim i.i.d.$ $\mathrm{N}(0, 1)$;

DGP.S2: $\lambda_{it} = \lambda_{i0} \sim i.i.d.$ $\mathrm{N}(0, \mathbb{I}_2)$, $e_{it} = \sigma_i v_{it}$, where $\sigma_i \sim i.i.d. U(0.5, 1.5)$, $v_{it} \sim i.i.d.$ $\mathrm{N}(0, 1)$;

DGP.S3: $\lambda_{it} = \lambda_{i0} \sim i.i.d.$ $\mathrm{N}(0, \mathbb{I}_2)$, $e_{it} = \sigma_i(F_t)v_{it}$, where $\sigma_i^2 = (0.2 + \delta_i) + 0.1F_{1t}^2 + 0.2F_{2t}^2$, $\delta_i \sim i.i.d.$ $\mathrm{U}(-0.1, 0.3)$, $v_{it} \sim i.i.d.$ $\mathrm{N}(0, 1)$;

DGP.S4: $\lambda_{it} = \lambda_{i0} \sim i.i.d.$ $\mathrm{N}(0, \mathbb{I}_2)$, $e_{\cdot t} = (e_{1t}, ...e_{Nt})' \sim i.i.d.$ $\mathrm{N}(0, \Sigma_e)$;

DGP.P1: $\lambda_{it,k} = \begin{cases} \lambda_{i0,k}, & \text{for } t = 1, 2, \ldots, T/2 \\ \lambda_{i0,k} + 0.2, & \text{for } t = T/2 + 1, \ldots, T \end{cases}$, $\lambda_{i0,k} \sim i.i.d.$ $\mathrm{N}(1, 1)$ for $k = 1, 2$, $e_{it} = \sigma_i v_{it}$, where $\sigma_i \sim i.i.d.$ $\mathrm{U}(0.5, 1.5)$, $v_{it} \sim i.i.d.$ $\mathrm{N}(0, 1)$;

DGP.P2: $\lambda_{it,1} = \begin{cases} \lambda_{i0,1}, & \text{for } \ 0.1T < t \leq 0.2T \ \text{ or } 0.7T < t \leq 0.8T \\ \lambda_{i0,1} + 0.5, & \text{for } \ 0.4T < t \leq 0.5T \\ \lambda_{i0,1} - 0.5, & \text{otherwise} \end{cases}$, $\lambda_{it,2} = \lambda_{i0,2} \sim i.i.d.$ $\mathrm{N}(0, 1)$, $e_{it} \sim i.i.d.$ $\mathrm{N}(0, 1)$;

DGP.P3: $\lambda_{it,1} = \lambda_{i0,1} \sim i.i.d.$ $\mathrm{N}(0, 1)$, $\lambda_{it,2} = 0.5G\left(10t/T; 2, 5i/N + 2\right)$, $e_{it} \sim i.i.d.$ $\mathrm{N}(0, 1)$;

DGP.P4: $\lambda_{it,1} = \mu_i + 0.5G(10t/T; 0.1, (1, 3, 7, 9)')$, $\mu_i \sim i.i.d.$ $\mathrm{N}(0, 1)$, $\lambda_{it,2} = \lambda_{i0,2} \sim i.i.d.$ $\mathrm{N}(0, 1)$, $e_{it} \sim i.i.d.$ $\mathrm{N}(0, 1)$;

DGP.P5: $\lambda_{it,k} = \begin{cases} \lambda_{i0,k}, & \text{for } t = 1, 2, \ldots, T/2 \\ \lambda_{i0,k} + 0.2, & \text{for } t = T/2 + 1, \ldots, T \end{cases}$, $\lambda_{i0,k} \sim i.i.d.$ $\mathrm{N}(1, 1)$ for $k = 1, 2$, $e_{\cdot t} = (e_{1t}, ...e_{Nt})' \sim i.i.d.$ $\mathrm{N}(0, \Sigma_e)$;

DGP.P6: $\lambda_{it,1} = \begin{cases} \lambda_{i0,1}, & \text{for } \ 0.1T < t \leq 0.2T \ \text{ or } 0.7T < t \leq 0.8T \\ \lambda_{i0,1} + 0.5, & \text{for } \ 0.4T < t \leq 0.5T \\ \lambda_{i0,1} - 0.5, & \text{otherwise} \end{cases}$, $\lambda_{it,2} = \lambda_{i0,2} \sim i.i.d.$ $\mathrm{N}(0, 1)$, $e_{\cdot t} = (e_{1t}, ...e_{Nt})' \sim i.i.d.$ $\mathrm{N}(0, \Sigma_e)$;

DGP.P7: $\lambda_{it,1} = \lambda_{i0,1} \sim i.i.d.$ $\mathrm{N}(0, 1)$, $\lambda_{it,2} = 0.5G\left(10t/T; 2, 5i/N + 2\right)$, $e_{\cdot t} = (e_{1t}, ...e_{Nt})' \sim i.i.d.$ $\mathrm{N}(0, \Sigma_e)$;

DGP.P8: $\lambda_{it,1} = \begin{cases} \lambda_{i0,1}, & \text{for } 0.1T < t \leq 0.2T \text{ or } 0.7T < t \leq 0.8T \\ \lambda_{i0,1} + \nu_i, & \text{for } 0.4T < t \leq 0.5T \\ \lambda_{i0,1} - \nu_i, & \text{otherwise} \end{cases}$, $\lambda_{it,2} = \lambda_{i0,2} \sim i.i.d.N(0,1)$, $e_{it} \sim i.i.d.N(0,1)$, $\nu_i \sim U[0,1]$;

where $\Sigma_e = (c_{ij})_{i,j=1,\ldots,N}$ with $c_{ij} = 0.5^{|i-j|}$, $G(z;\kappa,\boldsymbol{\gamma}) = \{1 + \exp[-\kappa \prod_{l=1}^{p} (z - \gamma_l)]\}^{-1}$ denotes the Logistic function with tuning parameter κ and location parameter $\boldsymbol{\gamma} = (\gamma_1, ..., \gamma_p)'$.

DGP.S1-S4 satisfy the null hypothesis of time-invariant factor loadings and are used to study the size of our test. Specifically, DGP.S2 - S4 examine the performance of the test under heteroskedasticity, conditional heteroskedasticity and cross-sectional dependence respectively. DGP.P1-P8 describe various time-varying factor loadings. Among them, DGP.P1-P2 have a single sudden structural break and multiple sudden structural breaks, respectively. DGP.P3-P4 describe two kinds of smooth structural changes. In particular, the factor loadings generated by DGP.P3 are monotonic function while the factor loadings given by DGP.P4 are smooth transition functions with multiple regime shifts. DGP.P5-P7 parallel DGP.P1-P3 but allow for cross-sectional dependence. DGP.P8 has heterogenous time variation. Some path plots for DGP.P1-P4 are shown in the online supplement.

To examine the asymptotic local power property, we also consider the local alternative: $\lambda_{it} = \lambda_{i0} + ca_{NT}g_i\left(\frac{t}{T}\right)$, where $c = 1, 2, 4$, $a_{NT} = T^{-1/2}N^{-1/4}h^{-1/4}$, $\lambda_{i0} \sim i.i.d.$ $\mathrm{N}(0, \mathbb{I}_2)$ and the setups of the $g_i(\cdot) = (g_{i1}(\cdot), g_{i2}(\cdot))'$ are given by:

DGP.P9: $g_{ik}(u) = \mathbf{1}(|u| \leq \frac{1}{2})$ for $k = 1, 2$, $e_{it} = \sigma_i v_{it}$, where $\sigma_i \sim i.i.d.$ $\mathrm{U}(0.5, 1.5)$, $v_{it} \sim i.i.d.$ $\mathrm{N}(0,1)$;

DGP.P10: $g_{i1}(u) = 0$, $g_{i2}(u) = G(u; 20, 0.5)$, $e_{it} \sim i.i.d.$ $\mathrm{N}(0,1)$.

As mentioned above, our test does not require the trimming parameter used to control the minimum length of each subsample under the alternative. To check the performance of our test near the end of the sample, we follow the advice of a referee and consider the following DGP:

DGP.P11: $\lambda_{it,k} = \begin{cases} \lambda_{i0,k}, & \text{for } t = 1, 2, \ldots, cT \\ \lambda_{i0,k} + b, & \text{for } t = cT + 1, \ldots, T \end{cases}$, $\lambda_{i0,k} \sim i.i.d.$ $\mathrm{N}(1,1)$ for $k = 1, 2, e_{it} = \sigma_i v_{it}$, where $\sigma_i \sim i.i.d.$ $\mathrm{U}(0.5, 1.5)$, $v_{it} \sim i.i.d.$ $\mathrm{N}(0,1)$, and $\lambda_{i0,1}$, $\lambda_{i0,2}$, σ_i and v_{it} are mutually independent of each other. We consider the cases where $b = 0.2$, 0.5 and $c = 0.5$, 0.6, 0.7, 0.8, 0.9. Apparently, the structural break point moves from the middle to the end of the sample as c increases from 0.5 to 0.9.

4.2 Tests Statistics and Implementation

For each DGP, we simulate 500 data sets with $N = 100$, 200 and $T = 100$, 200, respectively. Since the factor loadings are assumed to be nonrandom, we generate them once for all and fix them across the Monte Carlo replications.

To implement our $\widehat{SM}_{NT}$ test, we apply the Epanechnikov kernel and the Silverman's rule-of thumb bandwidth $h = (2.35/\sqrt{12})T^{-1/5}$. [Note that $\{\frac{1}{T}, \frac{2}{T}, ..., \frac{T-1}{T}, 1\}$ behave like a $\mathrm{U}(0,1)$ random variable with variance 1/12.] We have also tried the Uniform kernel and the Quartic kernel, and the rule-of thumb bandwidth with different tuning parameters. Our simulation studies show that the choice of kernel function has little impact on the performance of our test. However, the empirical sizes and powers are a bit sensitive to the bandwidth selection. To alleviate this problem, we follow the nonparametric literature and apply the bootstrap procedure proposed in Section 3.4. We consider 500 replications with $B = 200$ bootstrap number for the bootstrap-based test. Moreover, we also examine the performance of our nonparametric for different choices of bandwidth sequences by setting $h = c(2.35/\sqrt{12})T^{-1/5}$ for $c = 0.5$, 1 and 1.5. Nevertheless, due to the space constraint, we only report the results with $c = 1$ in the paper and relegate the results for other choices of c to the online supplement.

In addition to our test, we also consider Breitung and Eickmeier's (2011) sup-LM N-variable-specific test, Chen et al.'s (2014) sup-LM test, and Han and Inoue's (2014) sup-LM test. Following these papers, we choose the trimming parameter $\tau = 0.15$ that restricts the one-time break, if it exists, to occur within the time interval $[0.15T,\ 0.85T]$. We also examine the performance of these tests with $\tau = 0.1$ and 0.25 and find the results are quite similar. The tests of Chen et al. (2014) and Han and Inoue (2015) involve the long-run variance estimation. We follow the HAC literature by choosing the Bartlett kernel and setting the truncation parameter to be $\lfloor T^{1/3} \rfloor$ to estimate the long-run variance. The critical values presented in Andrews (1993) are applied for the tests of Breitung and Eickmeier (2011), Chen et al. (2014) and Han and Inoue (2015). Note that Breitung and Eickmeier's (2011) sup-LM tests are implemented for each one of the N cross-sectional units and we can only report the average rejection frequency for this test where the averages are taken over these N cross-sectional units and the designated number of simulations. The number of replications is set to be 500.

Moreover, we also implement the bootstrap-version of Su and Wang's (2017) nonparametric test which contrasts the local-PCA estimates of the common components under the alternative with the conventional PCA estimates under the null. To implement their test, we follow their recommendation to choose the bandwidth parameter. Here, we also consider 500 replications with 200 bootstrap resamples for each replication.

4.3 Simulation Results

In this section we first report the comparison of the size behavior of various tests and then report the comparison of the power behavior of these tests under the global alternatives when the number of factors is set to be the true value and determined from the data, respectively. Then we study the local power performance of our test. Finally, we compare the performance of different tests when the one-time break date is near the end of the sample.

4.3.1 Size comparison with correctly specified R

Table 1 reports the size performance of our test as well as that of the parametric tests of Breitung and Eickmeier (2011), Chen et al. (2014) and Han and Inoue (2015) and the nonparametric test of Su and Wang (2017) at the 5% and 10% significance levels when the number of common factors are fixed as the true value $R = 2$. For our test, we report the results using the bootstrap critical values. As shown in this table, our test has reasonable sizes using bootstrap critical values. Su and Wang's (2017) test tends to overreject slightly but is still acceptable. The sup-LM tests of Han and Inoue (2015) and Chen et al. (2014) tend to under-reject slightly. In addition, Breitung and Eickmeier's (2011) N-variable-specific sup-LM test also suffers from slight under-rejection for DGP.S1-S2 and S4 but severe over-rejection for DGP.S3. It may not be difficult to understand the bad size performance of Breitung and Eickmeier's (2011) test, as their tests require the independence between common factors and the error term, which is violated in DGP.S3.

Table 1: Size of tests under DGP.S1-S4 when the number of factors is fixed to the true value

DGP	N	T	SM_B		$SW17$		HI_{LM}		CDG_{LM}		BE_{LM}	
			5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
S1	100	100	5.0	10.8	6.6	13.4	0.6	3.8	3.4	8.2	2.8	6.5
	100	200	5.8	12.4	7.4	13.0	2.4	6.8	4.8	7.4	3.4	7.5
	200	100	4.8	8.8	5.2	10.2	1.6	4.4	1.4	7.0	2.7	6.3
	200	200	5.4	10.8	5.8	12.0	1.6	6.8	3.6	8.8	3.4	7.5
S2	100	100	5.2	9.6	7.4	12.0	0.4	2.4	2.0	8.2	2.7	6.4
	100	200	4.6	9.8	5.0	11.4	1.0	5.8	2.0	6.6	3.7	7.8
	200	100	5.4	10.6	6.4	14.0	0.4	1.8	1.0	4.6	2.8	6.4
	200	200	6.6	11.2	7.0	14.0	0.6	5.4	2.6	6.8	3.6	7.7
S3	100	100	5.6	10.8	7.2	11.2	0.4	2.2	2.2	8.8	11.9	20.3
	100	200	4.8	9.8	6.0	11.4	1.6	5.2	2.0	6.0	15.2	24.5
	200	100	6.8	12.2	7.8	11.6	0.4	1.8	1.2	5.0	11.9	20.2
	200	200	7.4	13.4	8.2	13.0	0.8	5.2	2.4	7.0	15.2	24.7
S4	100	100	5.2	12.0	6.2	12.2	0.4	4.6	2.8	8.0	2.8	6.4
	100	200	5.2	10.4	4.2	10.4	2.0	6.8	4.6	8.6	3.4	7.5
	200	100	6.0	12.0	6.8	12.0	1.6	3.2	2.6	6.6	2.8	6.3
	200	200	5.0	9.6	5.6	10.2	2.2	7.0	4.0	8.4	3.4	7.4

Note: (i) SM_B denotes the results of our $\widehat{SM}_{NT}$ test using bootstrap critical values; (ii) $SW17$ denotes the results of Su and Wang's (2017) bootstrap-based test; (iii) HI_{LM} denotes Han and Inoue's (2014) sup-LM test; (iv) CDG_{LM} denotes Chen et al.'s (2014) sup-LM test; (v) BE_{LM} denotes Breitung and Eickmeier's (2011) N variable-specific sup-LM test. The main entries report the average percentage of rejection.

4.3.2 Global power comparison with correctly specified R

Table 2 reports the power performance of the tests under DGP.P1-P8 at the 5% and 10% significance levels when the number of common factors is fixed as the true value $R = 2$. Our bootstrap-based test is powerful in detecting all forms of time-varying factor loadings given by DGP.P1-P8 and its power increases as either T or N increases. Recall that DGP.P1-P2, P5-P6 and P8 are factor models with sudden structural breaks, while DGP.P3-P4 and P7 are factor models with smooth structural changes. The simulation results are consistent with our theoretical claim that the test is able to detect both a finite number of sudden structural breaks and smooth structural changes. In addition, Su and Wang's (2017) test is also powerful in detecting the deviation from the null in these DGPs. Moreover, the power of the new test is usually higher than that of Su and Wang's (2017) test in all cases except DGP.P3 and DGP.P7, which consider the monotonic smooth structural changes in factor loadings. Hence the power ranking of these two tests are ambiguous. In contrast, Han and Inoue's (2015) sup-LM test has relatively low power against DGP.P1-P2 and P4-P6. However, it is most powerful in detecting DGP.P3 and P7. This is because the factor loadings under DGP.P3 and P7 are monotonic functions of the scaled time t/T for each i. If we apply the method of PCA to estimate the factor model, the estimated factor series would behave like an explosive process with increasing volatility over time. Since Han and Inoue's (2015) test checks the time-invariance property of the second order moment of common factors, it is able to capture such smooth structural changes as in DGP.P3. In addition, both Chen et al.'s (2014) sup-LM test and Breitung and Eickmeier's (2011) N-variable-specific sup-LM test have quite low power against DGP.P1-P8, which exhibit either sudden structural breaks or smooth structural changes in factor loadings.

Table 2: Power of tests under DGP.P1-P8 when the number of factors is fixed to the true value

DGP	N	T	SM_B		$SW17$		HI_{LM}		CDG_{LM}		BE_{LM}	
			5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
P1	100	100	72.2	81.4	67.4	79.4	0.8	4.4	2.4	7.2	5.9	11.1
	100	200	98.4	99.6	98.4	99.4	4.2	10.6	2.0	6.8	11.2	17.8
	200	100	94.0	97.2	92.2	96.4	0.8	4.0	2.4	6.6	5.7	10.7
	200	200	100	100	100	100	5.0	12.2	2.2	6.6	11.1	17.5
P2	100	100	29.4	41.4	26.0	40.4	0.6	2.2	2.2	8.6	3.9	8.3
	100	200	82.2	86.8	76.8	84.0	1.6	6.4	2.2	6.4	6.7	12.7
	200	100	41.0	51.8	27.6	40.6	0.8	2.8	1.8	8.6	3.7	8.1
	200	200	93.0	95.8	85.2	91.6	1.6	5.8	1.8	7.6	6.5	12.4
P3	100	100	37.2	47.8	41.2	55.2	35.8	67.0	6.8	16.8	4.9	10.3
	100	200	64.8	73.8	77.0	86.6	97.4	99.8	10.2	18.4	9.8	17.2
	200	100	42.4	53.8	45.2	60.2	37.4	71.4	6.6	15.4	5.2	10.7
	200	200	76.0	82.2	84.2	92.0	99.2	100	10.2	20.0	9.8	17.7
P4	100	100	25.0	38.0	25.6	36.8	0.4	1.6	1.0	4.0	3.5	7.9
	100	200	74.2	83.6	72.2	81.6	0.6	4.0	3.0	5.6	5.4	10.6
	200	100	40.6	52.8	34.4	45.2	0.4	1.4	1.0	5.8	3.5	7.8
	200	200	92.0	94.4	86.8	92.8	0.2	3.8	3.2	6.4	5.5	10.7
P5	100	100	67.8	79.8	63.0	75.6	1.4	5.8	3.2	8.8	4.9	10.1
	100	200	97.4	99.2	96.8	99.0	6.0	12.8	4.4	8.4	9.8	16.6
	200	100	90.0	94.2	88.0	92.0	2.0	6.6	1.2	6.6	4.9	9.9
	200	200	100	100	99.6	99.8	3.8	11.4	4.8	10.6	9.4	15.8
P6	100	100	29.6	38.6	27.2	36.2	0.8	5.0	3.6	9.2	3.7	8.1
	100	200	81.2	86.0	75.8	82.8	3.2	10.4	5.6	10.8	6.2	12.1
	200	100	38.4	52.6	27.4	38.2	1.4	4.6	1.6	7.6	3.6	7.9
	200	200	92.4	95.8	85.2	90.8	3.0	9.8	4.6	11.0	6.2	11.9
P7	100	100	34.0	45.8	37.0	54.6	32.4	65.0	7.4	14.6	5.0	10.5
	100	200	62.4	72.2	74.4	86.2	98.2	99.6	12.0	18.0	9.5	16.9
	200	100	44.0	53.0	44.0	60.2	36.6	68.8	7.0	15.2	5.0	10.5
	200	200	78.8	85.0	86.4	92.6	99.0	99.8	10.8	19.6	9.7	17.5
P8	100	100	38.2	50.6	35.2	47.6	0.6	2.4	2.0	8.6	5.2	10.4
	100	200	91.2	94.8	88.6	92.2	1.6	6.2	2.2	6.6	10.7	18.1
	200	100	49.4	60.8	32.8	43.8	0.8	2.8	1.8	8.8	4.7	9.5
	200	200	97.8	99.0	93.2	95.4	1.6	5.8	2.0	7.6	9.3	15.9

Note: See the note in Table 1.

4.3.3 Size and global power comparison with R determined from the data

As the exact number R of common factors is typically unknown in practice, one should determine the number of common factors before estimating and testing. In the literature on testing for structural breaks in factor loadings, the number of common factors is either determined by Bai and Ng's (2002, BN hereafter) information criteria (e.g., Han and Inoue, 2015) or specified by some fixed numbers, which may be equal to, less than, or greater than the correct number of factors (e.g., Chen et al., 2014). Of course, one can also consider applying Onatski's (2009, 2010) or Ahn and Horenstein's (2013) testing procedures to determine the number of factors, which work well in the presence of moderate or strong cross-sectional dependence. Alternatively, one can apply Su and Wang's (2017) nonparametric method to determine the number of factors that is robust to the presence of structural changes in the factor loadings. In general, all the aforementioned methods can select the correct number of factors consistently under the null hypothesis, but only Su and Wang's (2017) method has been proven valid under the alternative too. Indeed, if we apply Su and Wang's (2017) method to determine the number of factors, the size and power performance of all tests will be similar to that in Tables 1 and 2 (see Tables A.3 and A.4 in the online supplement). To allow the possible misspecification of the number of factors under the alternative, here we follow Han and Inoue (2015) and select the number of factors based on BN's information criteria IC_{p1} and IC_{p2}. To implement IC_{p1} and IC_{p2}, we need to prescribe the maximum number of factors, $R_{\max}$. Given the true value of R is 2, we set $R_{\max} = 6$ in our simulations. We find that the results based on IC_{p1} and IC_{p2} are quite similar and thus we only report the results using IC_{p1} below to save space.

Tables 3 and 4 report the size and power performance of various tests at the 5% and 10% significance levels when the number of factors is determined by BN's IC_{p1}. The results are similar to those reported in Tables 1 and 2. In fact, for all DGPs, our simulation results show that BN's IC_{p1} only tends to overparameterize slightly, and the problem alleviates as the sample size increases. Moreover, we also examine the performance of various tests by setting the number of common factors as 3. The power of our bootstrap-based test is a little bit lower than that in the case of correctly specified factors as reported in Table 2. However, our test still has reasonable power that increases as either N or T increases; and more importantly, it is still the most powerful test among all tests for most DGPs under consideration. To save space, we do not report the results for this case here.

4.3.4 Local power performance of our test

We now study the local power property of our test by using DGP.P9-P10. Table 5 reports the empirical rejection frequency of our test at the 5% and 10% significance levels when the number of factors is fixed to be the true value or determined by BN's IC_{p1}, respectively. As shown in the table, the power of our test increases fast in both cases as c increases.

4.3.5 Performance of various tests when the break date is near the end of the sample

Finally, we compare the performance of various tests when the break date is near the end of the sample by using DGP.P11. Tables 6 and 7 report the empirical rejection rate under this DGP when the number of common factors is fixed as the true number and determined from the data respectively.

We summarize the findings from Tables 6-7. First, all of the parametric tests considered by Breitung

Table 3: Size of tests under DGP.S1-S4 when the number of factors is determined from the data

DGP	N	T	SM_B		$SW17$		HI_{LM}		CDG_{LM}		BE_{LM}	
			5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
S1	100	100	5.0	10.8	6.6	13.4	0.6	3.8	3.4	8.2	2.8	6.5
	100	200	5.8	12.4	7.4	13.0	2.4	6.8	4.8	7.4	3.4	7.5
	200	100	4.8	8.8	5.2	10.2	1.6	4.4	1.4	7.0	2.7	6.2
	200	200	5.4	10.8	5.8	12.0	1.6	6.8	3.6	8.8	3.4	7.5
S2	100	100	5.2	9.6	7.4	12.0	0.4	2.4	2.0	8.2	2.8	6.5
	100	200	4.6	9.8	5.0	11.4	1.0	5.8	2.0	6.6	3.7	7.8
	200	100	5.4	10.6	6.4	14.0	0.4	1.8	1.0	4.6	2.8	6.4
	200	200	6.6	11.2	7.0	14.0	0.6	5.4	2.6	6.8	3.6	7.7
S3	100	100	5.6	10.8	7.2	11.2	0.4	2.2	2.2	8.8	11.9	20.1
	100	200	4.8	9.8	6.0	11.4	1.6	5.2	2.0	6.0	15.3	24.5
	200	100	6.8	12.2	7.8	11.6	0.4	1.8	1.2	5.0	11.8	20.2
	200	200	7.4	13.4	8.2	13.0	0.8	5.2	2.4	7.0	15.2	24.7
S4	100	100	5.2	12.0	6.2	12.2	0.4	4.6	2.8	8.0	2.8	6.4
	100	200	5.2	10.4	4.2	10.4	2.0	6.8	4.6	8.6	3.4	7.5
	200	100	6.0	12.0	6.8	12.0	1.6	3.2	2.6	6.6	2.7	6.3
	200	200	5.0	9.6	5.6	10.2	2.2	7.0	4.0	8.4	3.3	7.5

Note: See the note in Table 1.

and Eickmeier(2011), Han and Inoue (2015) and Chen et al. (2014) have extremely low power against this DGP for both choices of c when $b = 0.2$, and the rejection frequency is close to the nominal level in most cases. As b increases to 0.5, the powers of these parametric tests increase but are still significantly lower than the powers of the two nonparametric tests. Second, when the structural break point moves from the middle to the end of the sample, the empirical rejection rates of the parametric tests decrease significantly and almost lose power for $b = 0.5$ when $c = 0.9$. Third, both our test and Su and Wang's (2017) test have reasonably high power to detect the structural changes near the end of the sample and the power increases as either T or N increases. However, the empirical rejection rates of these nonparametric tests also decrease when the structural break point moves from the middle to the end. This simulation result is as expected in the structure change literature and is consistent with our theoretical claim. We note that when the break point moves from the middle to the end, the post-break period gets shorter and shorter, and "the cumulative effect of structural changes" is smaller. That is, π_0 given by Theorem 3.2 gets smaller despite the fact it is still significantly different from zero as along as c does not tend to 1 as $(N, T) \to \infty$.

5 An Empirical Application

In this section, we apply our test to check whether the factor loadings for asset returns suffer from structural changes. Factor models for asset returns have received extensive attention in the finance literature. Since the factor loadings depend on the nature of the information available to investors at any given time, they may vary over time. Li and Yang (2011) and Ang and Kristensen (2012) consider the conditional factor models when the number of assets/portfolios is fixed and small. Li and Yang (2011) model the factor loadings as smooth functions of time; Ang and Kristensen model them as smooth functions of some macroeconomic and financial variables that are thought to capture systematic risks as observable factors. Both devise Wald-

Table 4: Power of tests under DGP.P1-P8 when the number of factors is determined from the data

DGP	N	T	SM_B		$SW17$		HI_{LM}		CDG_{LM}		BE_{LM}	
			5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
P1	100	100	72.8	80.8	70.0	78.8	0.8	4.4	2.4	7.2	5.9	11.1
	100	200	98.8	99.6	98.4	99.4	4.2	10.6	2.0	6.8	11.2	17.8
	200	100	94.2	97.2	92.4	97.2	0.8	4.0	2.4	6.6	5.7	10.7
	200	200	100	100	100	100	5.0	12.2	2.2	6.6	11.1	17.5
P2	100	100	29.8	41.6	26.4	40.8	0.4	1.6	1.8	7.8	3.7	8.0
	100	200	85.2	88.6	80.0	85.8	1.0	5.2	2.4	6.4	5.3	11.7
	200	100	37.8	50.0	27.6	41.2	0.6	2.2	1.6	7.6	3.4	7.6
	200	200	92.4	94.6	83.2	89.0	1.0	3.8	2.0	8.4	4.7	9.7
P3	100	100	36.0	41.0	52.0	63.0	35.4	66.2	6.8	16.6	5.0	10.4
	100	200	65.2	74.8	76.2	85.6	97.4	99.8	10.2	18.4	9.8	17.2
	200	100	42.4	52.0	45.6	60.4	37.4	71.4	6.6	15.4	5.2	10.7
	200	200	76.0	81.6	85.0	92.0	99.2	100	10.2	20.0	9.8	17.7
P4	100	100	25.4	35.8	26.4	35.2	0.4	1.6	1.0	4.0	3.5	7.9
	100	200	73.4	83.6	70.2	81.4	0.6	4.0	3.0	5.6	5.4	10.6
	200	100	40.0	52.6	32.2	45.2	0.4	1.4	1.0	5.8	3.5	7.8
	200	200	91.0	94.6	85.6	92.2	0.2	3.8	3.2	6.4	5.5	10.7
P5	100	100	69.0	80.2	63.2	76.2	1.4	5.8	3.2	8.8	4.9	10.1
	100	200	97.6	99.4	96.8	98.8	6.0	12.8	4.4	8.4	9.8	16.6
	200	100	90.4	94.8	87.0	91.8	2.0	6.6	1.2	6.6	4.9	9.9
	200	200	100	100	99.6	99.6	3.8	11.4	4.8	10.6	9.5	15.8
P6	100	100	27.8	39.8	27.0	35.6	0.8	4.8	3.2	8.4	3.6	7.9
	100	200	80.8	85.2	76.2	82.2	3.2	8.6	4.8	10.0	5.0	10.2
	200	100	40.4	52.6	27.6	38.6	1.0	4.0	1.4	6.8	3.4	7.5
	200	200	92.4	95.4	86.2	90.8	1.2	6.0	2.8	7.6	4.5	9.2
P7	100	100	33.2	46.0	35.6	54.0	32.2	64.2	7.4	14.4	5.0	10.4
	100	200	63.6	72.2	74.2	86.4	98.2	99.6	12.0	18.0	9.5	16.9
	200	100	44.8	52.2	40.6	59.4	36.6	68.8	7.0	15.2	5.0	10.5
	200	200	78.2	84.4	86.4	93.4	99.0	99.8	10.8	19.6	9.7	17.5
P8	100	100	38.6	51.2	36.4	47.0	0.2	1.8	1.6	6.4	3.6	8.0
	100	200	90.6	95.4	88.0	92.6	0.2	1.4	2.2	9.8	4.3	8.8
	200	100	49.4	60.8	33.0	44.4	0.6	1.6	1.4	5.2	3.4	7.5
	200	200	97.8	99.2	92.8	95.8	0.2	1.8	2.6	8.6	4.0	8.6

Note: See the note in Table 1.

Table 5: Local power performance of our test

DGP	N	T	R : fixed to the true value						R : determined from BN's IC_{p1}					
			$c=1$		$c=2$		$c=4$		$c=1$		$c=2$		$c=4$	
			5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
P9	100	100	8.0	14.0	14.8	23.4	55.0	66.4	6.6	10.8	11.0	19.8	54.4	68.0
	100	200	6.6	12.8	19.2	29.2	65.2	77.4	6.8	14.4	15.8	22.4	63.8	76.2
	200	100	8.0	14.8	13.0	22.0	47.8	60.2	6.2	13.4	13.2	22.2	60.2	72.4
	200	200	8.0	13.6	15.6	27.4	56.0	69.8	9.2	14.8	16.6	27.0	69.4	78.8
P10	100	100	7.6	15.2	15.0	24.2	52.4	62.8	7.2	10.2	10.2	20.2	49.4	64.4
	100	200	5.4	12.8	11.8	21.6	56.0	67.0	6.6	13.2	12.6	20.8	52.4	66.6
	200	100	7.6	15.6	13.0	22.2	52.2	66.0	8.2	14.4	13.8	23.2	50.8	65.0
	200	200	8.0	13.8	14.2	21.2	58.2	70.6	8.4	14.2	15.8	23.2	60.0	72.8

Note: The main entries report the average percentage of rejection; c signifies the magnitude of local deviation from the null hypothesis.

type tests for the significance of long-run conditional alphas and find substantial variation in the conditional factor loadings. More recently, Ma et al. (2019) propose a high-dimensional alpha test to assess whether there exist abnormal excess returns on high-dimensional assets by allow the factor loadings to evolve over time. In all these studies, the factors are assumed to be observed. When the factors are not observed, we can also check whether the factor loadings are time-varying by using the method developed in this paper.

Monthly data between 2000.1-2015.9 are available for 9145 stocks traded on the New York Exchange, AMEX, and NASDAQ, which are obtained from the WIND data base. The data include live stocks whose suspensions are no more than two years between this period. Finally, we get a balanced panel with $T = 189$, $N = 2684$.

We use BN's four information criteria (namely, $PC_{p1}, PC_{p2}, IC_{p1}, IC_{p2}$), Ahn and Horenstein's (2013) two criterion functions (ER for eigenvalue ratio and GR for growth ratio) and Onatski's (2009) sequential testing procedure to determine the number of common factors. The maximum number of factors is set to be 8 in this empirical study. The estimated number of factors by PC_{p1} and PC_{p2} is 3, the other two BN's information criteria (IC_{p1}, IC_{p2}) and Onatski's (2009) procedure all choose 2 common factors, while Ahn and Horenstein's (2013) testing procedures choose 1 common factor. Therefore, in the following context, we report the test results for the cases of one, two and three common factors, respectively.

We apply our nonparametric test $\widehat{SM}_{NT}$, Han and Inoue's (2014) sup-LM and sup-Wald tests, as well as Chen et al.'s (2014) sup-LM and sup-Wald tests to investigate the possible structural changes in factor loadings. The smooth parameter, kernel functions and other presettings for these tests are all the same to those used in the simulation studies. For our test, we focus on the bootstrap results based on $B = 500$ bootstrap replications. Since the one-sided $N(0, 1)$ critical values at the 5% and 10% levels are 1.64 and 1.28, respectively, it is obvious that we can reject the null hypothesis of time-invariant factor loadings at the 5% significance level by using asymptotic critical value.

Table 8 reports the results of the tests and the corresponding critical values at the 5% and 10% significance levels. Our test rejects the null hypothesis for all the cases of one, two, and three common factors. In contrast, both the results of Han and Inoue (2015) and Chen et al. (2014) are mixed. Han and Inoue's

Table 6: Empirical rejection rates under DGP.P11 when the number of factors is fixed to the true value

	N	T	SM_B		$SW17$		HI_{LM}		CDG_{LM}		BE_{LM}	
			5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
$c = 0.5$	100	100	72.2	81.4	67.6	79.4	0.8	4.4	2.4	7.2	5.9	11.1
$b = 0.2$	100	200	98.4	99.6	98.4	99.4	4.2	10.6	2.0	6.8	11.2	17.8
	200	100	94.0	97.2	92.2	96.4	0.8	4.0	2.4	6.6	5.7	10.7
	200	200	100	100	100	100	5.0	12.2	2.2	6.6	11.1	17.5
$c = 0.6$	100	100	68.8	77.8	66.0	77.0	0.8	5.2	2.4	7.6	5.8	10.9
$b = 0.2$	100	200	98.2	98.8	98.4	99.0	5.6	11.6	2.0	7.0	11.1	17.3
	200	100	92.8	95.4	91.6	95.0	1.0	5.0	2.6	6.8	5.6	10.6
	200	200	100	100	99.8	99.8	6.0	12.2	2.4	6.8	10.9	17.3
$c = 0.7$	100	100	54.2	68.2	56.0	68.4	1.2	5.8	2.8	7.4	5.3	10.3
$b = 0.2$	100	200	95.2	98.2	96.0	97.8	5.6	12.4	2.0	7.0	10.0	16.1
	200	100	80.4	89.8	81.2	87.8	1.4	6.0	3.0	7.2	5.2	10.1
	200	200	99.2	99.4	99.4	99.4	6.6	13.4	2.4	6.4	9.9	16.1
$c = 0.8$	100	100	30.2	42.2	35.4	47.6	1.4	5.8	3.0	7.8	4.5	9.1
$b = 0.2$	100	200	73.2	83.2	77.0	84.2	5.2	11.6	2.2	6.6	8.1	13.8
	200	100	49.8	61.8	55.2	68.4	1.6	6.4	2.4	7.6	4.5	9.0
	200	200	90.8	93.8	91.0	95.0	6.0	12.4	2.2	6.6	8.0	13.9
$c = 0.9$	100	100	9.8	18.2	12.8	21.6	0.8	3.2	1.8	7.4	3.4	7.4
$b = 0.2$	100	200	23.0	33.8	28.2	38.8	1.8	8.4	2.0	6.2	5.1	9.9
	200	100	13.4	25.2	18.8	27.8	0.8	3.2	1.8	7.6	3.3	7.2
	200	200	31.0	44.4	38.2	49.8	2.6	8.4	2.2	6.0	5.0	9.8
$c = 0.5$	100	100	100	100	100	100	3.8	11.0	2.8	9.2	21.1	28.5
$b = 0.5$	100	200	100	100	100	100	22.8	39.8	3.4	7.6	36.6	44.1
	200	100	100	100	100	100	4.0	13.6	4.2	8.8	18.9	26.2
	200	200	100	100	100	100	26.0	44.6	4.2	9.2	33.3	40.8
$c = 0.6$	100	100	100	100	100	100	5.0	14.6	3.4	9.8	20.8	28.5
$b = 0.5$	100	200	100	100	100	100	29.2	44.4	4.0	9.4	36.5	43.9
	200	100	100	100	100	100	4.8	17.4	3.8	9.6	19.0	26.1
	200	200	100	100	100	100	33.2	49.8	5.0	10.0	33.4	40.8
$c = 0.7$	100	100	100	100	99.8	100	7.4	17.6	4.0	9.6	19.2	26.7
$b = 0.5$	100	200	100	100	100	100	30.4	44.2	3.8	9.8	34.0	41.6
	200	100	100	100	100	100	8.0	19.8	4.4	10.0	17.9	24.8
	200	200	100	100	100	100	34.8	48.8	5.0	10.6	31.5	38.7
$c = 0.8$	100	100	98.6	99.0	97.2	98.6	7.4	17.0	3.6	8.4	15.4	22.4
$b = 0.5$	100	200	100	100	100	100	27.6	40.0	3.6	9.6	28.6	36.0
	200	100	100	100	99.6	99.8	8.6	19.2	4.6	9.4	14.9	21.5
	200	200	100	100	100	100	30.4	44.8	4.4	10.0	27.0	34.1
$c = 0.9$	100	100	48.2	61.4	56.4	68.6	2.4	8.2	3.2	7.8	7.7	13.4
$b = 0.5$	100	200	91.4	94.2	93.8	96.6	8.0	16.6	2.2	7.4	15.3	22.1
	200	100	67.2	78.2	62.8	74.8	3.0	8.0	2.4	8.2	7.9	13.2
	200	200	97.2	98.6	97.2	98.6	9.4	18.2	2.4	7.4	15.5	21.9

Note: See the note in Table 1.

Table 7: Empirical rejection rates under DGP.P11 when the number of factors is determined from the data

	N	T	SM_B		$SW17$		HI_{LM}		CDG_{LM}		BE_{LM}	
			5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
$c = 0.5$	100	100	72.2	81.4	67.6	79.4	0.8	4.4	2.4	7.2	5.9	11.1
$b = 0.2$	100	200	98.4	99.6	98.4	99.4	4.2	10.6	2.0	6.8	11.2	17.8
	200	100	94.0	97.2	92.2	96.4	0.8	4.0	2.4	6.6	5.7	10.7
	200	200	100	100	100	100	5.0	12.2	2.2	6.6	11.1	17.5
$c = 0.6$	100	100	68.8	77.8	66.0	77.0	0.8	5.2	2.4	7.6	5.8	10.9
$b = 0.2$	100	200	98.2	98.8	98.4	99.0	5.6	11.6	2.0	7.0	11.1	17.3
	200	100	92.8	95.4	91.6	95.0	1.0	5.0	2.6	6.8	5.6	10.6
	200	200	100	100	99.8	99.8	6.0	12.2	2.4	6.8	10.9	17.3
$c = 0.7$	100	100	54.2	68.2	56.0	68.4	1.2	5.8	2.8	7.4	5.3	10.3
$b = 0.2$	100	200	95.4	98.2	96.0	97.8	5.6	12.4	2.0	7.0	10.0	16.1
	200	100	80.4	89.8	81.2	87.8	1.4	6.0	3.0	7.2	5.2	10.1
	200	200	99.2	99.4	99.4	99.4	6.6	13.4	2.4	6.4	9.9	16.1
$c = 0.8$	100	100	30.2	42.2	35.4	47.6	1.4	5.8	3.0	7.8	4.5	91
$b = 0.2$	100	200	73.2	83.2	77.0	84.2	5.2	11.6	2.2	6.6	8.1	13.8
	200	100	49.8	61.8	55.2	68.4	1.6	6.4	2.4	7.6	4.5	9.0
	200	200	90.8	93.8	91.0	95.0	6.0	12.4	2.2	6.6	8.0	13.9
$c = 0.9$	100	100	9.8	18.2	12.8	21.6	0.8	3.2	1.8	7.4	3.4	7.4
$b = 0.2$	100	200	23.0	33.8	28.2	38.8	1.8	8.4	2.0	6.2	5.1	9.9
	200	100	13.4	25.2	18.8	27.8	0.8	3.2	1.8	7.6	3.3	7.2
	200	200	31.0	44.4	38.2	49.8	2.6	8.4	2.2	6.0	5.0	9.8
$c = 0.5$	100	100	91.6	93.8	93.4	96.0	2.4	10.2	25.4	31.6	15.9	22.8
$b = 0.5$	100	200	100	100	100	100	89.0	91.8	87.2	87.8	11.9	18.2
	200	100	100	100	100	100	2.0	10.0	74.0	77.8	6.3	11.3
	200	200	100	100	100	100	99.6	99.6	99.4	99.4	5.3	10.7
$c = 0.6$	100	100	91.4	93.0	93.0	95.8	3.6	12.6	27.6	33.4	15.3	22.2
$b = 0.5$	100	200	100	100	100	100	89.6	92.6	88.0	88.8	11.5	17.8
	200	100	100	100	100	100	2.2	14.6	74.8	78.0	6.4	11.2
	200	200	100	100	100	100	99.8	99.8	99.8	99.8	5.1	10.4
$c = 0.7$	100	100	94.4	95.8	93.6	96.2	5.8	18.2	21.4	26.2	15.3	22.3
$b = 0.5$	100	200	100	100	100	100	75.2	80.2	71.2	72.6	14.9	21.5
	200	100	100	100	100	100	5.4	24.6	60.0	63.2	8.0	13.1
	200	200	100	100	100	100	97.2	97.6	97.2	97.4	5.6	10.8
$c = 0.8$	100	100	96.6	97.2	94.0	96.2	7.6	18.6	9.4	14.2	14.0	20.9
$b = 0.5$	100	200	100	100	100	100	41.8	49.2	31.2	33.8	21.4	28.3
	200	100	100	100	100	100	9.8	26.4	32.8	36.4	10.2	15.9
	200	200	100	100	100	100	72.4	76.6	70.4	72.2	10.4	15.9
$c = 0.9$	100	100	50.8	65.2	60.0	71.0	2.2	7.6	4.0	8.0	7.5	13.1
$b = 0.5$	100	200	91.6	94.6	93.2	96.2	8.0	17.2	5.0	10.2	14.7	21.6
	200	100	64.8	76.2	60.8	73.0	2.4	6.4	6.2	12.8	7.1	12.2
	200	200	96.0	97.6	93.0	95.0	13.8	23.8	13.4	27.8	13.4	19.6

Note: See the note in Table 1.

Table 8: Tests for Structural Changes for the Stock Returns

	Our test: bootstrap			Han and Inoue (2015)				Chen et al. (2014)			
	SM	5%	10%	sup-LM	sup-Wald	5%	10%	sup-LM	sup-Wald	5%	10%
$r = 1$	21.45	2.99	2.57	3.61	9.75	8.85	7.17	–	–	–	–
$r = 2$	21.08	8.12	6.25	24.78	16.06	14.15	12.27	14.65	9.35	8.85	7.17
$r = 3$	26.60	8.62	6.62	25.28	17.61	20.26	18.12	6.27	21.50	11.79	10.01

Notes (i) SM denotes the results of our $\widehat{SM}_{NT}$ test using the bootstrap critical value based on $B = 500$ iterations; (ii) entries below 5% and 10% denote the corresponding critical values.

(2014) sup-Wald test cannot reject the null for the case of three common factors, and their sup-LM test cannot reject the null for the case of one common factor, while Chen et al.'s (2014) sup-LM test cannot reject the null for the case of three common factors. This is consistent with the results of our simulation studies that the tests of Han and Inoue (2015) and Chen et al. (2014) have relatively low power.

As suggested by one anonymous referee, it is interesting to study the structural change features of the factor loadings. However, we want to mention that the factor loadings estimated in the second step is inconsistent under the global alternative, due to the inconsistent PCA estimation of common factors in the case of non-local structural changes. Figure A.2 in the online supplement plots Su and Wang's (2017) local PCA estimates of the time-varying factor loadings for some representative stocks. From the figure we can see that the estimated factor loadings show significant structural changes that appear very likely to be smooth changes.

6 Conclusion

Conventional factor models assume the factor loadings, which capture the relationship between random variables and the latent common factors, to be time-invariant. In fact, shocks induced by policy switch, preference change, and technology progress may cause structural changes in the relationship. Therefore, the assumption of time-invariant factor loadings may not hold in practice. In this paper, we propose a nonparametric test for structural changes in large dimensional factor models. Our test follows a convenient asymptotic $N(0, 1)$ distribution under the null hypothesis. By construction, it is powerful in detecting both smooth structural changes and sudden structural breaks with possibly unknown break dates or unknown number of breaks. Unlike existing tests such as Breitung and Eickmeier (2011), Chen et al. (2014), Han and Inoue (2015), Yamamoto and Tanaka (2015), and Cheng et al. (2016), our test does not require any trimming of the boundary regions and hence could detect any structural changes that occur close to the starting and ending points of the sample period. We also study the local power property and propose a bootstrap procedure to improve the finite sample performance of our test. Monte Carlo studies show that our test has reasonable size and excellent power in detecting various time-varying factor loadings. In an application to the U.S. asset returns, we find significant evidence against time-invariant factor loadings.

References

Ahn, S. & A. Horenstein (2013) Eigenvalue ratio test for the number of factors. *Econometrica* 81, 1203-1227.

Andrews, W.K. (1993) Tests for parameter instability and structural change with unknown change point. *Econometrica* 61, 821-856.

Ang, A. & D. Kristensen (2012) Testing conditional factor models. *Journal of Financial Economics* 106, 132-156.

Bai, J. (2003) Inferential theory for factor models of large dimensions. *Econometrica* 71, 135-172.

Bai, J. & S. Ng (2002) Determining the number of factors in approximate factor models. *Econometrica* 90, 191-221.

Baltagi, B., C. Kao, & F. Wang (2016) Estimating and testing high dimensional factor models with multiple structural changes. *Working paper* , Department of Economics, Syracuse University.

Banerjee, A., M. Marcellino, & I. Masten (2008) Forecasting macroeconomic variables using diffusion indexes in short samples with structural change. *Frontiers of Economics and Globalization* 3, 149-194.

Bates, B.J., M. Plagborg-Møller, J.H. Stock, & M.W. Watson (2013) Consistent factor estimation in dynamic factor models with structural instability. *Journal of Econometrics* 177, 289-304.

Breitung, J. & S. Eickmeier (2011) Testing for structural breaks in dynamic factor models. *Journal of Econometrics* 163, 71-84.

Cai, Z. (2007) Trending time-varying coefficient time series models with serially correlated errors. *Journal of Econometrics* 136, 163-188.

Chen, J., J. Gao, & D. Li (2012) Semiparametric testing for smooth structural changes in panel data models. *Journal of Econometrics* 171, 71-85.

Chen, L., J. Dolado, & J. Gonzalo (2014) Detecting big structural breaks in large factor models. *Journal of Econometrics* 180, 30-48.

Cheng, X., Z. Liao, & F. Schorfheide (2016) Shrinkage estimation of high-dimensional factor models with structural instabilities. *Review of Economic Studies* 83, 1511-1543.

Corradi, V. & N.R. Swanson (2014) Testing for structural stability of factor augmented forecasting models. *Journal of Econometrics* 182, 100-118.

Del Negro, M., & C. Otrok (2009) Dynamic factor models with time-varying parameters: measuring changes in international business cycles. *Staff Report No. 326, Federal Reserve Bank of New York.*

Eickmeier S., W. Lemke, & M. Marcellino (2015) Classical time varying factor-augmented vector auto-regressive models-estimation, forecasting and structural analysis. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 178, 493-533.

Fan J., Y. Liao, & M. Mincheva (2013) Large covariance estimation by thresholding principal orthogonal complements. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 75, 603-680.

Gao, J. & I. Gijbels (2008) Bandwidth selection in nonparametric kernel testing. *Journal of American Statistical Association* 103, 1584-1594.

Han, X. & A. Inoue (2015) Tests for parameter instability in dynamic factor models. *Econometric Theory* 31, 1117-1152.

Hong, Y. & H. Li (2005) Nonparametric specification testing for continuous-time models with applications to term structure of interest rate. *Review of Financial Studies* 18, 37-84.

Li, Q. & J.S. Racine (2007) *Nonparametric Econometrics: Theory and Practice.* Princeton University Press, New Jersey.

Li, Y. & L. Yang (2011) Testing conditional factor models: a nonparametric approach. *Journal of Empirical Finance* 18, 972-992.

Ma, S., W. Lan, L. Su, & C-L. Tsai (2019) Testing alphas in conditional time-varying factor models with high dimensional assets. *Journal of Business & Economic Statistics,* forthcoming.

Ma, S. & L. Su (2018) Estimation of large dimensional factor models with an unknown number of breaks. *Journal of Econometrics* 207, 1-29.

Mikkelsen, J.G., E. Hillebrand & G. Urga (2015) Maximum likelihood estimation of time-varying factor loadings in high-dimensional factor models. *CREATES research paper 2015-61.*

Moon, H. & M. Weidner (2015) Linear regression for panel with unknown number of factors as interactive fixed effects. *Econometrica* 83, 1543-1579.

Motta, C., C.M. Hafner & R. von Sachs (2011) Local stationary factor models: identification and nonparametric estimation. *Econometric Theory* 27, 1279-1319.

Onatski, A. (2009) Testing hypotheses about the number of factors in large factor models. *Econometrica* 77, 1447-1479.

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

Robinson, P.M. (2012) Nonparametric trending regression with cross-sectional dependence. *Journal of Econometrics* 169, 4-14.

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

Stock, J.H. & M.W. Watson (2009) Forecasting in dynamic factor models subject to structural instability, In J. Castle & N. Shephard (eds.), *The Methodology and Practice of Econometrics, A Festschrift in Honour of Professor David F. Hendry*, pp. 1-57, Oxford University Press, Oxford.

Su, L. & Q. Chen (2013) Testing homogeneity in panel data models with interactive fixed effects. *Econometric Theory* 29, 1079-1135.

Su, L., S. Jin, & Y. Zhang (2015) Specification test for panel data models with interactive fixed effects. *Journal of Econometrics* 185, 230-258.

Su, L. & X. Wang (2017) On time-varying factor models: estimation and testing. *Journal of Econometrics* 198, 84-101.

Sun, Y., P. Phillips, & S. Jin (2008) Optimal bandwidth selection in heteroskedasticity-autocorrelation robust testing. *Econometrica* 76, 175-194.

Yamamoto, Y. & S. Tanaka (2015) Testing for factor loading structural change under common breaks. *Journal of Econometrics* 189, 187-206.

Online Supplement for
" Testing for Structural Changes in Factor Models via a Nonparametric Regression"

Liangjun Su[a], Xia Wang[b]
[a] School of Economics, Singapore Management University, Singapore
[b] Lingnan (University) College, Sun Yat-Sen University, China

This Online Supplement contains two appendices. Appendix A is a mathematical appendix that contains some technical lemmas and the proofs of the theorems and lemmas in the paper. Appendix B contains some additional simulation and application results.

A Mathematical Appendix

This Mathematical Appendix is composed of three parts. Section A.1 provides some technical lemmas that are used in the proof of the theorems in Section 3. Section A.2 provides the proofs of the theorems in Section 3. Section A.3 gives the proofs of the technical lemmas in Section A.1.

A.1 Technical Lemmas

Let V_{NT} denote the $R \times R$ diagonal matrices of the first R largest eigenvalues of $(NT)^{-1} XX'$ arranged in decreasing order along its diagonal line. Let $H = (\Lambda_0'\Lambda_0/N)(F'\tilde{F}/T)V_{NT}^{-1}$, $C_{NT} = \min\{\sqrt{T}, \sqrt{N}\}$, $\tau_{ij,s} = E[e_{is}e_{js}F_s'F_s]$, $S_{Tr} = \frac{1}{T}\sum_{t=1}^{T} k_{h,tr}\tilde{F}_t\tilde{F}_t'$ and $S_{Tr}^{(0)} = \frac{1}{T}\sum_{t=1}^{T} k_{h,tr}F_tF_t'$. We state some technical lemmas whose proofs are relegated to Section A.3.

Lemma A.1 *Suppose Assumptions A.1–A.3 and A.5 hold. Then under* $\mathbb{H}_1(a_{NT})$ *with* $a_{NT} = T^{-1/2}N^{-1/4}h^{-1/4}$,

(i) $T^{-1}\tilde{F}'(NT)^{-1} XX'\tilde{F} = V_{NT} \xrightarrow{P} V_0$,

(ii) $(T^{-1}\tilde{F}'F)\left(N^{-1}\Lambda_0'\Lambda_0\right)(T^{-1}F'\tilde{F}) \xrightarrow{P} V_0$,

where V_{NT} *is an* $R \times R$ *diagonal matrix consisting of the* R *largest eigenvalues of* $(NT)^{-1} XX'$, *and* V_0 *is an* $R \times R$ *matrix consisting of the* R *eigenvalues of* $\Sigma_{\Lambda_0}\Sigma_F$, *both arranged in descending order.*

Lemma A.2 *Suppose that Assumptions A.1–A.3 and A.5 hold. Then under* $\mathbb{H}_1(a_{NT})$ *with* $a_{NT} = T^{-1/2}N^{-1/4}h^{-1/4}$,

(i) $\frac{1}{T}\left\|\tilde{F} - FH\right\|^2 = O_P\left(C_{NT}^{-2}\right)$,

(ii) $\frac{1}{T}(\tilde{F} - FH)'FH = O_P\left(C_{NT}^{-2}\right) + o_P(a_{NT})$,

(iii) $\frac{1}{T}(\tilde{F} - FH)'\tilde{F} = O_P\left(C_{NT}^{-2}\right) + o_P(a_{NT})$,

(iv) $\frac{1}{T}(\tilde{F}'\tilde{F} - H'F'FH) = O_P\left(C_{NT}^{-2}\right) + o_P(a_{NT})$,

(v) $V_{NT} = V_0 + O_P\left(C_{NT}^{-1}\right)$,

(vi) $H = Q_0^{-1} + O_P\left(C_{NT}^{-1}\right)$,

where $Q_0 = V_0^{1/2}\Upsilon_0^{-1}\Sigma_{\Lambda_0}^{-1/2}$ *and* Υ_0 *denotes the probability limit of* Υ_{NT} *defined in the proof of (v).*

Lemma A.3 *Suppose that Assumptions A.1–A.5 hold. Then under* $\mathbb{H}_1(a_{NT})$ *with* $a_{NT} = T^{-1/2}N^{-1/4}h^{-1/4}$,

(i) $\max_r\left\|\frac{1}{T}\sum_{t=1}^{T} k_{h,tr}(\tilde{F}_t - H'F_t)(\tilde{F}_t - H'F_t)'\right\| = O_P(T^{-1}\ln T + N^{-1})$,

(ii) $\max_r\left\|\frac{1}{T}\sum_{r=1}^{T} k_{h,tr}(\tilde{F}_t - H'F_t)F_tH'\right\| = O_P(T^{-1}\ln T + N^{-1}) + o_P(a_{NT})$,

(iii) $\max_r\left\|S_{Tr}^{(0)} - \Sigma_F\right\| = O_P\left(T^{-1/2}(\ln T)^{1/2}\right)$,

(iv) $\max_r \left\| S_{Tr} - \left(Q_0^{-1}\right)' \Sigma_F Q_0^{-1} \right\| = O_P((Th)^{-1/2} (\ln T)^{1/2} + N^{-1/2})$,
(v) $\max_r \left\| H S_{Tr}^{-1} S_{Tr}^{-1} H' - \Sigma_F^{-1} Q_0' Q_0 \Sigma_F^{-1} \right\| = O_P((Th)^{-1/2} (\ln T)^{1/2} + N^{-1/2})$.

Lemma A.4 *Suppose that Assumptions A.1–A.5 hold. Then under $\mathbb{H}_1(a_{NT})$ with $a_{NT} = T^{-1/2}N^{-1/4}h^{-1/4}$,*
(i) $\max_{i,r} \left\| \frac{1}{T} \sum_{t=1}^T k_{h,tr} \tilde{F}_t F_t' g_{it} \right\| = O_P(1)$,
(ii) $\max_{i,r} \left\| \frac{1}{T} \sum_{t=1}^T k_{h,tr} F_t' e_{it} \right\| = O_P(T^{-1/2} h^{-1/2} \ln(NT))$,
(iii) $\frac{1}{N} \sum_{i=1}^N \left\| \frac{1}{T} \sum_{s=1}^T (\tilde{F}_s - HF_s) e_{is}^\dagger \right\|^2 = O_P\left(C_{NT}^{-4}\right)$,
(iv) $\frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T \left\| \frac{1}{T} \sum_{s=1}^T (\tilde{F}_s - H'F_s) e_{is} k_{h,st} \right\|^2 = O_P\left(N^{-3/2} + T^{-2}\right) + o_P\left(a_{NT}^2\right)$.

Lemma A.5 *Suppose that Assumptions A.1 and A.3–A.5 hold. Suppose that $\mathbb{H}_1(a_{NT})$ holds with $a_{NT} = T^{-1/2}N^{-1/4}h^{-1/4}$. Then uniformly in (i, r),*
(i) $\frac{1}{T} \sum_{t=1}^T k_{h,tr} F_t F_t' g_{it} = \Sigma_F g_i\left(\frac{r}{T}\right) + o_P(1)$,
(ii) $\frac{1}{T} \sum_{t=1}^T F_t F_t' g_{it} = \Sigma_F \frac{1}{T} \sum_{t=1}^T g_i\left(\frac{t}{T}\right) + o_P(1) = o_P(1)$.

Lemma A.6 *Suppose that Assumptions A.1–A.5 hold. Then under $\mathbb{H}_1(a_{NT})$ with $a_{NT} = T^{-1/2}N^{-1/4}h^{-1/4}$,*
(i) $\frac{1}{N} \sum_{i=1}^N \left\| \tilde{\lambda}_{i0} - H^{-1} \lambda_{i0} \right\|^2 = O_P\left(C_{NT}^{-2}\right)$,
(ii) $\frac{1}{T} \sum_{t=1}^T \left\| \tilde{F}_t \right\|^l = O_P(1)$ *for* $l = 4, 6$,
(iii) $\max_s \frac{1}{T} \sum_{t=1}^T k_{h,st}^2 = O\left(h^{-1}\right)$,
(iv) $\frac{1}{T^2} \sum_{t=1}^T \sum_{s=1}^T \tilde{F}_s' L_{st} \tilde{F}_s \tilde{F}_s' \tilde{F}_s = O\left(h^{-1}\right)$,
(v) $\frac{1}{T} \sum_{t=1}^T \left\| (\tilde{F}_t - H'F_t) \tilde{F}_t \right\|^2 = O_P\left(C_{NT}^{-2} + TN^{-2}\right)$,
(vi) $\frac{1}{NT} \sum_{t=1}^T \left\| \sum_{i=1}^N \lambda_{i0}' H'^{-1} (\tilde{F}_t - H'F_t) e_{it} \right\|^2 = O_P\left(C_{NT}^{-2}\right)$,
(vii) $\frac{1}{NT} \sum_{i=1}^N \sum_{s=1}^T (\tilde{F}_s - H'F_s) F_s' H e_{is}^2 = O_P(a_{NT})$.

In addition, we need the following lemma from Sun and Chiang (1997).

Lemma A.7 *Let $\{V_t, t \geq 1\}$ be a strong mixing process with mixing coefficient $\alpha(\cdot)$. Let $G_{t_1,...,t_m}$ denote the distribution function of $(V_{t_1}, ..., V_{t_m})$. For any integer $m > 1$ and integers $(t_1, ..., t_m)$ such that $1 \leq t_1 < t_2 < \ldots < t_m$, let ϑ be a Borel measurable function such that* $\max\{\int |\vartheta(v_1, \ldots, v_m)|^{1+\tilde{\eta}} dG_{t_1,\ldots,t_j}(v_1, \ldots, v_j) dG_{t_j+1,\ldots,t_m}(v_{j+1}, \ldots, v_m), \int |\vartheta(v_1, \ldots, v_m)|^{1+\tilde{\eta}} dG_{t_1,\ldots,t_m}\} \leq M$ *for some $\tilde{\eta} > 0$. Then* $|\int \vartheta(v_1, \ldots, v_m) dG_{t_1,\ldots,t_m}(v_1, \ldots, v_m) - \int \vartheta(v_1, \ldots, v_m) dG_{t_1,\ldots,t_j}(v_1, \ldots, v_j) dG_{t_j+1,\ldots,t_m}(v_{j+1}, \ldots, v_m)| \leq 4M^{1/(1+\tilde{\eta})} \alpha(t_{j+1} - t_j)^{\tilde{\eta}/(1+\tilde{\eta})}$.

A.2 Proof of the Theorems in Section 3

Proof of Theorem 3.1. The result in Theorem 3.1 follows as a special case of Theorem 3.2 with $g_i(t/T) = 0$ for each i and t. ■

Proof of Theorem 3.2. Under $\mathbb{H}_1(a_{NT}) : \lambda_{it} = \lambda_{i0} + a_{NT} g_i(t/T)$, we can decompose $TN^{1/2}h^{1/2}\hat{M}$ as follows:

$$\begin{aligned} TN^{1/2}h^{1/2}\hat{M} &= N^{-1/2}h^{1/2}\sum_{i=1}^{N}\sum_{t=1}^{T}\left\|\left(\hat{\lambda}_{it}-H^{-1}\lambda_{i0}\right)-\left(\tilde{\lambda}_{i}-H^{-1}\lambda_{i0}\right)\right\|^{2} \\ &= N^{-1/2}h^{1/2}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\hat{\lambda}_{it}-H^{-1}\lambda_{i0}\right)'\left(\hat{\lambda}_{it}-H^{-1}\lambda_{i0}\right) \\ &+ N^{-1/2}h^{1/2}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\tilde{\lambda}_{i}-H^{-1}\lambda_{i0}\right)'\left(\tilde{\lambda}_{i}-H^{-1}\lambda_{i0}\right) \\ &- 2N^{-1/2}h^{1/2}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\hat{\lambda}_{it}-H^{-1}\lambda_{i0}\right)'\left(\tilde{\lambda}_{i}-H^{-1}\lambda_{i0}\right) \\ &\equiv M_1+M_2-2M_3, \text{ say,} \end{aligned}$$

where for notational simplicity we suppress the dependence of M_l on (N,T) for $l=1,2,3$. We complete the proof by showing that under $\mathbb{H}_1(a_{NT})$, (i) $M_1-\mathbb{B}_{1NT}-\Pi_{1NT}\xrightarrow{d}N(0,\mathbb{V}_0)$, (ii) $M_2-\mathbb{B}_{2NT}-\Pi_{2NT}=o_P(1)$, and (iii) $M_3-\mathbb{B}_{3NT}-\Pi_{3NT}=o_P(1)$, (iv) $\hat{\mathbb{B}}_{NT}=\mathbb{B}_{NT}+o_P(1)$, and (v) $\hat{\mathbb{V}}_{NT}=\mathbb{V}_{NT}+o_P(1)$, where $\mathbb{B}_{NT}=\mathbb{B}_{1NT}+\mathbb{B}_{2NT}-2\mathbb{B}_{3NT}$, and

$$\begin{aligned} \mathbb{B}_{1NT} &= \frac{h^{1/2}}{T^2N^{1/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}k_{h,st}^2F_s'HS_{Tt}^{-1}S_{Tt}^{-1}H'F_se_{is}^2, \\ \mathbb{B}_{2NT} &= \frac{h^{1/2}}{TN^{1/2}}\sum_{i=1}^{N}\sum_{s=1}^{T}F_s'HH'F_se_{is}^2, \\ \mathbb{B}_{3NT} &= \frac{h^{1/2}}{T^2N^{1/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}k_{h,st}F_s'HS_{Tt}^{-1}H'F_se_{is}^2, \\ \Pi_{1NT} &= \frac{1}{TN}\sum_{i=1}^{N}\sum_{t=1}^{T}\operatorname{tr}\left(Q_0g_i\left(\frac{t}{T}\right)g_i\left(\frac{t}{T}\right)'Q_0'\right), \\ \Pi_{2NT} &= \frac{1}{N}\sum_{i=1}^{N}\operatorname{tr}\left[(Q_0^{-1})'\Sigma_F\frac{1}{T}\sum_{r=1}^{T}g_i\left(\frac{r}{T}\right)\frac{1}{T}\sum_{s=1}^{T}g_i\left(\frac{s}{T}\right)'\Sigma_FQ_0^{-1}\right], \\ \Pi_{3NT} &= \frac{1}{N}\sum_{i=1}^{N}\operatorname{tr}\left[\frac{1}{T}\sum_{r=1}^{T}g_i\left(\frac{r}{T}\right)\frac{1}{T}\sum_{t=1}^{T}g_i\left(\frac{t}{T}\right)'\Sigma_F\right], \end{aligned}$$

$\mathbb{V}_{NT}$ are defined in Theorem 3.2, and $\mathbb{V}_0=\lim_{(N,T)\to\infty}\mathbb{V}_{NT}$. We prove these claims in Propositions A.8-A.12 below. Noting that $\frac{1}{T}\sum_{r=1}^{T}g_i\left(\frac{r}{T}\right)=\int_0^1 g_i(u)\,du+O(1/T)=O(1/T)$ under the normalization rule $\int_0^1 g_i(u)\,du=0$, we have $\Pi_{lNT}=O(1/T)$ for $l=2,3$. Combining these results yields $\widehat{SM}_{NT}=\hat{\mathbb{V}}_{NT}^{-1/2}(TN^{1/2}h^{1/2}\hat{M}-\hat{\mathbb{B}}_{NT})\xrightarrow{d}N(\pi_0,1)$, where $\pi_0=\lim_{(N,T)\to\infty}\Pi_{1NT}/\mathbb{V}_{NT}^{1/2}$. ■

Proposition A.8 *Suppose that the conditions in Theorem 3.2 hold. Then* $M_1-\mathbb{B}_{1NT}-\Pi_{1NT}\xrightarrow{d}N(0,\mathbb{V}_0)$ *under* $\mathbb{H}_1(a_{NT})$.

Proof. Using $X_{it}=F_t'\lambda_{it}+e_{it}=F_t'HH^{-1}\lambda_{i0}+e_{it}+a_{NT}F_t'g_{it}=\tilde{F}_t'H^{-1}\lambda_{i0}+e_{it}+a_{NT}F_t'g_{it}-(\tilde{F}_t-$

$H'F_t)'H^{-1}\lambda_{i0}$, we have

$$\hat{\lambda}_{it} - H^{-1}\lambda_{i0} = \left(\frac{1}{T}\sum_{s=1}^{T} k_{h,st}\tilde{F}_s\tilde{F}_s'\right)^{-1}\frac{1}{T}\sum_{s=1}^{T} k_{h,st}\tilde{F}_s X_{is} - H^{-1}\lambda_{i0}$$
$$= S_{Tt}^{-1}H'\frac{1}{T}\sum_{s=1}^{T} k_{h,st}F_s e_{is} + a_{NT}S_{Tt}^{-1}\frac{1}{T}\sum_{s=1}^{T} k_{h,st}\tilde{F}_s F_s' g_{is}$$
$$- S_{Tt}^{-1}\frac{1}{T}\sum_{s=1}^{T} k_{h,st}\tilde{F}_s\left(\tilde{F}_s - H'F_s\right)' H^{-1}\lambda_{i0} + S_{Tt}^{-1}\frac{1}{T}\sum_{s=1}^{T} k_{h,st}\left(\tilde{F}_s - H'F_s\right) e_{is}$$
$$\equiv D_1(i,t) + D_2(i,t) - D_3(i,t) + D_4(i,t), \quad \text{say}, \tag{A.1}$$

where $S_{Tt} = \frac{1}{T}\sum_{s=1}^{T} k_{h,st}\tilde{F}_s\tilde{F}_s'$. By (A.1) we decompose M_1 as follows:

$$M_1 = h^{1/2}N^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T}\|D_1(i,t) + D_2(i,t) - D_3(i,t) + D_4(i,t)\|^2$$
$$= h^{1/2}N^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T}[\|D_1(i,t)\|^2 + \|D_2(i,t)\|^2 + \|D_3(i,t)\|^2 + \|D_4(i,t)\|^2$$
$$+ 2D_1(i,t)'D_2(i,t) - 2D_1(i,t)'D_3(i,t)) + 2D_1(i,t)'D_4(i,t)$$
$$-2D_2(i,t)'D_3(i,t) + 2D_2(i,t)'D_4(i,t)) - 2D_3(i,t)'D_4(i,t)]$$
$$\equiv M_{1,1} + M_{1,2} + M_{1,3} + M_{1,4} + 2M_{1,5} - 2M_{1,6} + 2M_{1,7} - 2M_{1,8} + 2M_{1,9} - 2M_{1,10}, \text{ say.}$$

We prove the proposition by showing that (i) $M_{1,1} - \mathbb{B}_{1NT} \xrightarrow{d} N(0, \mathbb{V}_0)$, (ii) $M_{1,2} = \Pi_{1NT} + o_P(1)$, and (iii) $M_{1,j} = o_P(1)$ for $j = 3, ..., 10$.

We first prove (i). We decompose the $M_{1,1}$ term as follows:

$$M_{1,1} = \frac{h^{1/2}}{N^{1/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\left\|S_{Tt}^{-1}H'\frac{1}{T}\sum_{s=1}^{T}F_s e_{is}k_{h,st}\right\|^2$$
$$= \frac{h^{1/2}}{T^2N^{1/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}k_{h,st}F_s'e_{is}HS_{Tt}^{-1}S_{Tt}^{-1}H'\sum_{r=1}^{T}k_{h,rt}F_r e_{ir}$$
$$= \frac{h^{1/2}}{T^2N^{1/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}k_{h,st}^2F_s'HS_{Tt}^{-1}S_{Tt}^{-1}H'F_s e_{is}^2 + \frac{2h^{1/2}}{T^2N^{1/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{1\le r<s\le T}k_{h,st}k_{h,rt}F_s'\mathbb{S}F_r e_{is}e_{ir}$$
$$+ \frac{2h^{1/2}}{T^2N^{1/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{1\le r<s\le T}k_{h,st}k_{h,rt}F_s'\left(HS_{Tt}^{-1}S_{Tt}^{-1}H' - \mathbb{S}\right)F_r e_{is}e_{ir} \equiv M_{1,1}^{(1)} + M_{1,1}^{(2)} + M_{1,1}^{(3)},$$

where $\mathbb{S} \equiv \Sigma_F^{-1}Q_0'Q_0\Sigma_F^{-1}$. Apparently, $M_{1,1}^{(1)} = \mathbb{B}_{1NT}$. For $M_{1,1}^{(2)}$, we make the following decomposition

$$M_{1,1}^{(2)} = \frac{2h^{1/2}}{T^2N^{1/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{1\le r<s\le T}k_{h,st}k_{h,rt}F_s'\mathbb{S}F_r e_{is}e_{ir}$$
$$= \frac{2}{TN^{1/2}h^{1/2}}\sum_{i=1}^{N}\sum_{1\le r<s\le T}\bar{K}\left(\frac{s-r}{Th}\right)F_s'\mathbb{S}F_r e_{is}e_{ir}$$
$$+ \frac{2}{TN^{1/2}h^{1/2}}\sum_{i=1}^{N}\sum_{1\le r<s\le T}\left[\frac{h}{T}\sum_{t=1}^{T}k_{h,st}k_{h,rt} - \bar{K}\left(\frac{s-r}{Th}\right)\right]F_s'\mathbb{S}F_r e_{is}e_{ir} \equiv M_{1,1}^{(2,1)} + M_{1,1}^{(2,2)},$$

where $\bar{K}(v)=\int_{-1}^{1}K(u)K(u-v)\,du$. Let $Z_{NT,s}=T^{-1}N^{-1/2}h^{-1/2}\sum_{r=1}^{s-1}\bar{K}\left(\frac{s-r}{Th}\right)F_s'\mathbb{S}F_r e_s' e_r$, then $M_{1,1}^{(2,1)}=2\sum_{s=2}^{T}Z_{NT,s}$ and $E\left(Z_{NT,s}|\mathcal{F}_{NT,s-1}\right)=T^{-1}N^{-1/2}h^{-1/2}\sum_{r=1}^{s-1}\bar{K}\left(\frac{s-r}{Th}\right)F_s'\mathbb{S}F_r E\left(e_s'|\mathcal{F}_{NT,s-1}\right)e_r=0$. By the martingale central limit theorem (e.g., Pollard, 1984, p.171), it suffices to prove $\mathbb{V}_{NT}^{-1/2}M_{1,1}^{(2,1)}\xrightarrow{d}N(0,1)$ by showing that

$$\mathcal{Z}\equiv\sum_{s=2}^{T}E\left(Z_{NT,s}^4|\mathcal{F}_{NT,s-1}\right)=o_P(1)\ \text{ and }\ \sum_{s=2}^{T}Z_{NT,s}^2-\mathbb{V}_{NT}=o_P(1)\,. \tag{A.2}$$

First, we verify the first part of (A.2). Observing that $\mathcal{Z}\geq 0$, it suffices to show $\mathcal{Z}=o_P(1)$ by showing that $E(\mathcal{Z})=o(1)$ by Markov inequality. Let $\bar{k}_{sr}=\bar{K}\left(\frac{s-r}{Th}\right)$ and $\phi_{sr}=F_s'\mathbb{S}F_r e_s' e_r$. We have

$$\begin{aligned}E(\mathcal{Z})&=\sum_{s=2}^{T}E\left\{\left[\frac{2}{TN^{1/2}h^{1/2}}\sum_{r=1}^{s-1}\bar{k}_{sr}\phi_{sr}\right]^4\right\}\\&=\frac{16}{T^4N^2h^2}\sum_{s=2}^{T}E\left[\sum_{r=1}^{s-1}\bar{k}_{sr}^4\phi_{sr}^4+2\sum_{1\leq r_1<r_2\leq s-1}\bar{k}_{sr_1}^2\bar{k}_{sr_2}^2\phi_{sr_1}^2\phi_{sr_2}^2\right.\\&\quad\left.+4\sum_{t=1}^{s-1}\sum_{1\leq r_1<r_2\leq s-1}\bar{k}_{st}^2\bar{k}_{sr_1}\bar{k}_{sr_2}\phi_{st}^2\phi_{sr_1}\phi_{sr_2}+4\sum_{1\leq r_1<r_2\leq s-1,1\leq t_1<t_2\leq s-1}\bar{k}_{r_1s}\bar{k}_{r_2s}\bar{k}_{t_1s}\bar{k}_{t_2s}\phi_{sr_1}\phi_{sr_2}\phi_{st_1}\phi_{st_2}\right]\\&\equiv\mathcal{Z}_1+\mathcal{Z}_2+\mathcal{Z}_3+\mathcal{Z}_4,\ \text{say.}\end{aligned}$$

Noting that $\max_{r<s}\left\|N^{-1/2}\phi_{sr}\right\|_4^4\leq C$ under Assumption A.3(v), we can readily show that under Assumption A.4

$$\begin{aligned}\mathcal{Z}_1&\leq\max_{r<s}\left\|N^{-1/2}\phi_{sr}\right\|_4^4\frac{16}{T^4h^2}\sum_{s=2}^{T}\sum_{r=1}^{s-1}\bar{k}_{sr}^4=O\left(T^{-2}h^{-1}\right)\\\mathcal{Z}_2&\leq\max_{r<s}\left\|N^{-1/2}\phi_{sr}\right\|_4^4\frac{32}{T^4h^2}\sum_{s=2}^{T}\sum_{1\leq r_1<r_2\leq s-1}\bar{k}_{sr_1}^2\bar{k}_{sr_2}^2=O\left(T^{-1}\right),\\\mathcal{Z}_3&\leq\max_{r<s}\left\|N^{-1/2}\phi_{sr}\right\|_4^4\frac{64}{T^4h^2}\sum_{s=2}^{T}\sum_{t=1}^{s-1}\sum_{1\leq r_1<r_2\leq s-1}\bar{k}_{st}^2\bar{k}_{sr_1}\bar{k}_{sr_2}=O(h)\,.\end{aligned}$$

For $\mathcal{Z}_4$, we can apply Assumptions A.3(iii) and (v) and A.5 along with the Davydov inequality to show that

$$\mathcal{Z}_4=\frac{64}{T^4N^2h^2}\sum_{s=2}^{T}\sum_{1\leq r_1<r_2\leq s-1,1\leq t_1<t_2\leq s-1}\bar{k}_{r_1s}\bar{k}_{r_2s}\bar{k}_{t_1s}\bar{k}_{t_2s}E\left(\phi_{sr_1}\phi_{sr_2}\phi_{st_1}\phi_{st_2}\right)=O(h)\,.$$

Thus $E(\mathcal{Z})=o(1)$ and $\mathcal{Z}=o_P(1)$.

To verify the second part of (A.2), it suffices to show (I) $\sum_{s=2}^{T}E\left(Z_{NT,s}^2\right)=\mathbb{V}_{NT}+o(1)$, and (II) $\text{Var}(\sum_{s=2}^{T}Z_{NT,s}^2)=o_P(1)$ by Chebyshev inequality. These two claims can be easily proved if we also assume independence of $\{e_{i\cdot}=(e_{i1},...,e_{iT})'\}$ across i conditional on the factor. Here we prove them without imposing such a cross-sectional independence condition. We first prove (I). Observe that

$$\begin{aligned}\operatorname{Var}(M_{1,1}^{(2,1)}) &= \sum_{s=2}^{T} E\left(Z_{NT,s}^2\right) = 4T^{-2}N^{-1}h^{-1}\sum_{s=2}^{T}\sum_{r=1}^{s-1}\bar{k}_{sr}^2 E\left(F_s'\mathbb{S}F_r e_s' e_r\right)^2 \\ &\quad + 4T^{-2}N^{-1}h^{-1}\sum_{s=2}^{T}\sum_{1\le r_1\ne r_2\le s-1}\bar{k}_{sr_1}\bar{k}_{sr_2}E\left(F_s'\mathbb{S}F_{r_1}e_s'e_{r_1}F_{r_2}'\mathbb{S}F_s e_{r_2}'e_s\right) \\ &\equiv \mathbb{V}_{NT} + b_{NT}.\end{aligned}$$

To study b_{NT}, let $\mathbb{S} = \Sigma_F^{-1}Q_0'Q_0\Sigma_F^{-1} = \{s_{mn}\}$. Then $\phi_{sr} = F_s'\mathbb{S}F_r e_s'e_r = \sum_{m=1}^{R}\sum_{n=1}^{R} s_{mn}F_{sm}F_{rn}e_s'e_r$, and

$$\begin{aligned}b_{NT} &= 4T^{-2}N^{-1}h^{-1}\sum_{s=2}^{T}\sum_{1\le r_1\ne r_2\le s-1}\bar{k}_{sr_1}\bar{k}_{sr_2}E\left(F_s'\mathbb{S}F_{r_1}e_s'e_{r_1}F_{r_2}'\mathbb{S}F_s e_{r_2}'e_s\right) \\ &= 4T^{-2}N^{-1}h^{-1}\sum_{1\le m_1,m_2\le R}\sum_{1\le n_1,n_2\le R}\sum_{s=2}^{T}\sum_{1\le r_1\ne r_2\le s-1}\bar{k}_{sr_1}\bar{k}_{sr_2}s_{m_1n_1}s_{m_2n_2}E\left(F_{sm_1}F_{r_1n_1}e_s'e_{r_1}F_{sm_2}F_{r_2n_2}e_s'e_{r_2}\right) \\ &= 4\sum_{1\le m_1,m_2\le R}\sum_{1\le n_1,n_2\le R} s_{m_1n_1}s_{m_2n_2}b_{NT}(m_1,m_2,n_1,n_2)\end{aligned}$$

where $b_{NT}(m_1,m_2,n_1,n_2) = T^{-2}N^{-1}h^{-1}\sum_{s=2}^{T}\sum_{1\le r_1\ne r_2\le s-1}\sum_{i=1}^{N}\sum_{j=1}^{N}\bar{k}_{sr_1}\bar{k}_{sr_2}E(F_{sm_1}F_{r_1n_1}e_{is}e_{jr_1}F_{sm_2}$ $\times F_{r_2n_2}e_{is}e_{jr_2})$. Since R is fixed and s_{mn}'s are finite, $b_{NT} = o(1)$ provided $b_{NT}(m_1,m_2,n_1,n_2) = o(1)$ for each quadruple (m_1,m_2,n_1,n_2). We consider three cases (1) $|s-r_2| > T_0$, (2) $|s-r_2|\le T_0$ and $|r_2-r_1| > T_0$, and (3) $|s-r_2|\le T_0$ and $|r_2-r_1|\le T_0$. We use $b_{NT}^{(l)}(m_1,m_2,n_1,n_2)$ to denote $b_{NT}(m_1,m_2,n_1,n_2)$ when the time indices are restricted to case (l) for $l=1,2,3$. In case (1), we apply Lemma A.7 and the fact that $E(F_{r_1}F_{r_2}'e_{ir_1}e_{ir_2}) = 0$ for $r_1\ne r_2$ under Assumption A.3(iii) to obtain

$$|b_{NT}^{(l)}(m_1,m_2,n_1,n_2)| \le CT^{-2}N^{-1}h^{-1}\sum_{r_1<r_2<s}\sum_{i=1}^{N}\sum_{j=1}^{N}\bar{k}_{sr_1}\bar{k}_{sr_2}\alpha(T_0)^{\delta/(1+\delta)} = O\left(NTh\alpha(T_0)^{\delta/(1+\delta)}\right) = o(1)$$

In case (2), we apply Lemma A.7 and the fact that $E(F_{r_1}e_{ir_1}) = 0$ to obtain

$$|b_{NT}^{(l)}(m_1,m_2,n_1,n_2)| \le CT^{-2}N^{-1}h^{-1}\sum_{r_1<r_2<s}\sum_{i=1}^{N}\sum_{j=1}^{N}\bar{k}_{sr_1}\bar{k}_{sr_2}\alpha(T_0)^{\delta/(1+\delta)} = O\left(NTh\alpha(T_0)^{\delta/(1+\delta)}\right) = o(1)$$

In case (3), we have

$$\begin{aligned}\left|b_{NT}^{(3)}(m_1,m_2,n_1,n_2)\right| &= T^{-2}N^{-1}h^{-1}\sum_{r_1<r_2<s,\ \text{case (3)}}\bar{k}_{sr_1}\bar{k}_{sr_2}\left|E\left(F_sF_se_{r_1}'e_se_{r_2}'e_sF_{r_1}F_{r_2}\right)\right| \\ &\le \max_{m,n}\max_{r<s}\left\|N^{-1/2}F_rF_se_r'e_s\right\|_2^2 T^{-2}h^{-1}\sum_{r_1<r_2<s,\ \text{case (3)}}\bar{k}_{sr_1}\bar{k}_{sr_2} = O\left(T^{-1}T_0^2h\right) = o(1),\end{aligned}$$

where we use the fact that the total number of terms in the summation over the three time indices for $b_{NT}^{(3)}$ are of order $O\left(TT_0^2\right)$. In sum, we have shown that $b_{NT} = o(1)$ and $\sum_{s=2}^{T}E\left(Z_{NT,s}^2\right) = \mathbb{V}_{NT} + o(1)$.

Now, we want to prove (II) by showing that $E(\sum_{s=2}^{T} Z_{NT,s}^2)^2 = \mathbb{V}_{NT}^2 + o(1)$. Noting that

$$
\begin{aligned}
E\left(\sum_{s=2}^{T} Z_{NT,s}^2\right)^2 &= \frac{1}{T^4N^2h^2}E\left(\sum_{s=2}^{T}\left[\sum_{r=1}^{s-1}\bar{k}_{sr}\phi_{sr}\right]^2\right)^2 \\
&= \frac{1}{T^4N^2h^2}E\left(\sum_{s=2}^{T}\sum_{r=1}^{s-1}\bar{k}_{sr}^2\phi_{sr}^2\right)^2 + \frac{1}{T^4N^2h^2}E\left(\sum_{s=2}^{T}\sum_{1\le r_1\ne r_2\le s-1}\bar{k}_{sr_1}\bar{k}_{sr_2}\phi_{sr_1}\phi_{sr_2}\right)^2 \\
&\quad + \frac{2}{T^4N^2h^2}E\left[\left(\sum_{s=2}^{T}\sum_{r=1}^{s-1}\bar{k}_{sr}^2\phi_{sr}^2\right)\sum_{s=2}^{T}\sum_{1\le r_1\ne r_2\le s-1}\bar{k}_{sr_1}\bar{k}_{sr_2}\phi_{sr_1}\phi_{sr_2}\right] \\
&\equiv b_{1NT} + b_{2NT} + b_{3NT}, \text{ say,}
\end{aligned}
$$

it suffices to show that (a) $b_{1NT} = \mathbb{V}_{NT}^2 + o_P(1)$ and (b) $b_{2NT} = o_P(1)$, because then $b_{3NT} \le 2\{b_{1NT}b_{2NT}\}^{1/2} = o_P(1)$ by Cauchy-Schwarz (CS) inequality. Note that $b_{1NT} = \frac{1}{T^4N^2h^2}\sum_{1\le r_1<s_1\le T, 1\le r_1<s_1\le T}\bar{k}_{s_1r_1}^2\bar{k}_{s_2r_2}^2 \times E(\phi_{s_1r_1}^2\phi_{s_2r_2}^2)$ and $\mathbb{V}_{NT}^2 = \frac{1}{T^4N^2h^2}\sum_{1\le r_1<s_1\le T, 1\le r_2<s_2\le T}\bar{k}_{s_1r_1}^2\bar{k}_{s_2r_2}^2 E\left(\phi_{s_1r_1}^2\right)E\left(\phi_{s_2r_2}^2\right)$. Let $\mathcal{S}_3 = \{r_1, s_1, r_2, s_2\}$. We consider two cases: (1) for each $t \in \mathcal{S}_3$, $|t-l| > T_0$ for all $l \in \mathcal{S}_3$ with $l \ne t$, and (2) all the other remaining cases. Let $\mathcal{S}_{3,1}$ and $\mathcal{S}_{3,2}$ denote the subsets of $\mathcal{S}_3$ corresponding to these two cases, respectively. For $l = 1, 2$, let $b_{1NT}(l)$ and $\mathbb{V}_{NT}^2(l)$ denote b_{1NT} and $\mathbb{V}_{NT}^2$ when the time indices are restricted to lie in $\mathcal{S}_{3,l}$, respectively. Note that $b_{1NT} = b_{1NT}(1) + b_{1NT}(2)$ and $\mathbb{V}_{NT}^2 = \mathbb{V}_{NT}^2(1) + \mathbb{V}_{NT}^2(2)$. In case (2), we have by Assumptions A.3(iii), (v) and A.4

$$
\begin{aligned}
b_{1NT}(2) &\le \max_{s<r}\left\|N^{-1}\phi_{sr}^2\right\|_2^2 \frac{1}{T^4h^2}\sum_{\substack{1\le r_1<s_1\le T, 1\le r_2<s_2\le T,\\ \text{case (2)}}}\bar{k}_{s_1r_1}^2\bar{k}_{s_2r_2}^2 = O\left(T_0T^{-1}\right) = o(1), \\
\mathbb{V}_{NT}^2(2) &\le \max_{s<r}\left[E\left(N^{-1}\phi_{sr}^2\right)\right]^2 \frac{1}{T^4h^2}\sum_{\substack{1\le r_1<s_1\le T, 1\le r_2<s_2\le T\\ \text{case (2)}}}\bar{k}_{s_1r_1}^2\bar{k}_{s_2r_2}^2 = O\left(T_0T^{-1}\right) = o(1),
\end{aligned}
$$

where we use the fact that there are at most T^3T_0 terms in the above displayed summations. In case (1), we consider six subcases: (1a) $r_1 < s_1 < r_2 < s_2$, (1b) $r_2 < s_2 < r_1 < s_1$, (1c) $r_1 < r_2 < s_1 < s_2$, (1d) $r_2 < r_1 < s_1 < s_2$, (1e) $r_1 < r_2 < s_2 < s_1$, and (1f) $r_2 < r_1 < s_2 < s_1$. We use $b_{1NT}(1, v)$ and $\mathbb{V}_{NT}^2(1, v)$ to denote $b_{1NT}(1)$ and $\mathbb{V}_{NT}^2(1)$, respectively, when the summation over the time indices are restricted to satisfy the conditions in subcase (1v) for $v = a, b, c, d, e, f$. First, we study subcase (1a). By Lemma A.7, Assumptions A.3(iii), (v) and A.4

$$
\begin{aligned}
b_{1NT}(1,a) &= \frac{1}{T^4N^2h^2}\sum_{r_1<s_1<r_2<s_2,\ \mathcal{S}_{3,1}}\bar{k}_{s_1r_1}^2\bar{k}_{s_2r_2}^2 E\left(\phi_{s_1r_1}^2\phi_{s_2r_2}^2\right) \\
&= \frac{1}{T^4N^2h^2}\sum_{r_1<s_1<r_2<s_2,\ \mathcal{S}_{3,1}}\sum_{i_1,j_1,i_2,j_2}\bar{k}_{s_1r_1}^2\bar{k}_{s_2r_2}^2 E(f_{s_1r_1}^2 e_{i_1s_1}e_{i_1r_1}e_{j_1s_1}e_{j_1r_1}f_{s_2r_2}^2 e_{i_2s_2}e_{i_2r_2}e_{j_2s_2}e_{j_2r_2}) \\
&\le \frac{1}{T^4N^2h^2}\sum_{r_1<s_1<r_2<s_2,\ \mathcal{S}_{3,1}}\sum_{i_1,j_1,i_2,j_2}\bar{k}_{s_1r_1}^2\bar{k}_{s_2r_2}^2\{E\left(f_{s_1r_1}^2 e_{i_1s_1}e_{i_1r_1}e_{j_1s_1}e_{j_1r_1}\right) \\
&\quad \times E\left(f_{s_2r_2}^2 e_{i_2s_2}e_{i_2r_2}e_{j_2s_2}e_{j_2r_2}\right) + C\alpha(T_0)^{\delta/(1+\delta)}\} \\
&= \frac{1}{T^4N^2h^2}\sum_{r_1<s_1<r_2<s_2,\ \mathcal{S}_{3,1}}\sum_{i_1,j_1,i_2,j_2}\bar{k}_{s_1r_1}^2\bar{k}_{s_2r_2}^2 E\left(\phi_{s_1r_1}^2\right)E\left(\phi_{s_2r_2}^2\right) + O(N^2\alpha(T_0)^{\delta/(1+\delta)}) \\
&= \mathbb{V}_{NT}^2(1,a) + o(1),
\end{aligned}
$$

where $f_{sr} = F_s'\mathbb{S}F_r$, $\sum_{i_1,j_1,i_2,j_2}$ denotes $\sum_{i_1=1}^{N}\sum_{j_1=1}^{N}\sum_{i_2=1}^{N}\sum_{j_2=1}^{N}$, and $\sum_{r_1<s_1<r_2<s_2,\ \mathcal{S}_{3,1}}$ indicates the summation is done over the four time indices satisfying the condition in case (1) (corresponding to $\mathcal{S}_{3,1}$).

By the same token, $b_{1NT}(1,b) = \mathbb{V}_{NT}^2(1,b) + o(1)$. Now, consider subcase (1c). For notational simplicity, we assume that $R = 1$ so that each term in $F_s' \mathbb{S} F_t$ is a scalar. [Otherwise, we need to utilize $F_s' \mathbb{S} F_t = \sum_{m=1}^{R} \sum_{n=1}^{R} s_{mn} F_{s,m} F_{s,n}$ as in the analysis of Part (I)]. By applying Lemma A.7 three times, we have

$$\begin{aligned}
b_{1NT}(1,c) &= \frac{1}{T^4 N^2 h^2} \sum_{r_1<r_2<s_1<s_2, \mathcal{S}_{3,1}} \bar{k}_{s_1 r_1}^2 \bar{k}_{s_2 r_2}^2 E\left(\phi_{s_1 r_1}^2 \phi_{s_2 r_2}^2\right) \\
&= \frac{\mathbb{S}^4}{T^4 N^2 h^2} \sum_{r_1<r_2<s_1<s_2, \mathcal{S}_{3,1}} \sum_{i_1, j_1, i_2, j_2} \bar{k}_{s_1 r_1}^2 \bar{k}_{s_2 r_2}^2 E\left(F_{s_1}^2 F_{s_2}^2 F_{r_1}^2 F_{r_2}^2 e_{i_1 s_1} e_{i_1 r_1} e_{j_1 s_1} e_{j_1 r_1} e_{i_2 s_2} e_{i_2 r_2} e_{j_2 s_2} e_{j_2 r_2}\right) \\
&\leq \frac{\mathbb{S}^4}{T^4 N^2 h^2} \sum_{r_1<r_2<s_1<s_2, \mathcal{S}_{3,1}} \sum_{i_1, j_1, i_2, j_2} \bar{k}_{s_1 r_1}^2 \bar{k}_{s_2 r_2}^2 \{E\left(F_{r_1}^2 F_{r_2}^2 e_{i_1 r_1} e_{j_1 r_1} e_{i_2 r_2} e_{j_2 r_2}\right) \\
&\quad \times E\left(F_{s_1}^2 F_{s_2}^2 e_{i_1 s_1} e_{j_1 s_1} e_{i_2 s_2} e_{j_2 s_2}\right) + C\alpha(T_0)^{\delta/(1+\delta)}\} \\
&\leq \frac{\mathbb{S}^4}{T^4 N^2 h^2} \sum_{r_1<r_2<s_1<s_2, \mathcal{S}_{3,1}} \sum_{i_1, j_1, i_2, j_2} \bar{k}_{s_1 r_1}^2 \bar{k}_{s_2 r_2}^2 \{E\left(F_{r_1}^2 e_{i_1 r_1} e_{j_1 r_1}\right) E\left(F_{r_2}^2 e_{i_2 r_2} e_{j_2 r_2}\right) \\
&\quad \times E\left(F_{s_1}^2 e_{i_1 s_1} e_{j_1 s_1}\right) E\left(F_{s_2}^2 e_{i_2 s_2} e_{j_2 s_2}\right) + 2C\alpha(T_0)^{\delta/(1+\delta)}\} \\
&= \frac{\mathbb{S}^4}{T^4 N^2 h^2} \sum_{r_1<r_2<s_1<s_2, \mathcal{S}_{3,1}} \sum_{i_1, j_1, i_2, j_2} \bar{k}_{s_1 r_1}^2 \bar{k}_{s_2 r_2}^2 E\left(F_{r_1}^2 e_{i_1 r_1} e_{j_1 r_1}\right) E\left(F_{r_2}^2 e_{i_2 r_2} e_{j_2 r_2}\right) \\
&\quad \times E\left(F_{s_1}^2 e_{i_1 s_1} e_{j_1 s_1}\right) E\left(F_{s_2}^2 e_{i_2 s_2} e_{j_2 s_2}\right) + o(1).
\end{aligned}$$

Similarly,

$$\begin{aligned}
\mathbb{V}_{NT}^2(1,c) &= \frac{1}{T^4 N^2 h^2} \sum_{r_1<r_2<s_1<s_2, \mathcal{S}_{3,1}} \bar{k}_{s_1 r_1}^2 \bar{k}_{s_2 r_2}^2 E\left(\phi_{s_1 r_1}^2\right) E\left(\phi_{s_2 r_2}^2\right) \\
&= \frac{\mathbb{S}^4}{T^4 N^2 h^2} \sum_{r_1<r_2<s_1<s_2, \mathcal{S}_{3,1}} \sum_{i_1, j_1, i_2, j_2} \bar{k}_{s_1 r_1}^2 \bar{k}_{s_2 r_2}^2 E\left(F_{r_1}^2 F_{r_2}^2 e_{i_1 r_1} e_{j_1 r_1} e_{i_2 r_2} e_{j_2 r_2}\right) \\
&\quad \times E\left(F_{s_1}^2 F_{s_2}^2 e_{i_1 s_1} e_{j_1 s_1} e_{i_2 s_2} e_{j_2 s_2}\right) \\
&\leq \frac{\mathbb{S}^4}{T^4 N^2 h^2} \sum_{r_1<r_2<s_1<s_2, \mathcal{S}_{3,1}} \sum_{i_1, j_1, i_2, j_2} \bar{k}_{s_1 r_1}^2 \bar{k}_{s_2 r_2}^2 \{E\left(F_{r_1}^2 e_{i_1 r_1} e_{j_1 r_1}\right) E\left(F_{r_2}^2 e_{i_2 r_2} e_{j_2 r_2}\right) \\
&\quad \times E\left(F_{s_1}^2 e_{i_1 s_1} e_{j_1 s_1}\right) E\left(F_{s_2}^2 e_{i_2 s_2} e_{j_2 s_2}\right) + C\alpha(T_0)^{\delta/(1+\delta)}\} \\
&= \frac{\mathbb{S}^4}{T^4 N^2 h^2} \sum_{r_1<r_2<s_1<s_2, \mathcal{S}_{3,1}} \sum_{i_1, j_1, i_2, j_2} \bar{k}_{s_1 r_1}^2 \bar{k}_{s_2 r_2}^2 E\left(F_{r_1}^2 e_{i_1 r_1} e_{j_1 r_1}\right) E\left(F_{r_2}^2 e_{i_2 r_2} e_{j_2 r_2}\right) \\
&\quad \times E\left(F_{s_1}^2 e_{i_1 s_1} e_{j_1 s_1}\right) E\left(F_{s_2}^2 e_{i_2 s_2} e_{j_2 s_2}\right) + o(1).
\end{aligned}$$

It follows that $b_{1NT}(1,c) = \mathbb{V}_{NT}^2(1,c) + o(1)$. Analogously, we can show that $b_{1NT}(1,v) = \mathbb{V}_{NT}^2(1,v) + o(1)$ for $v = d, e, f$. Consequently, we have $b_{1NT}(1) = \mathbb{V}_{NT}^2(1) + o(1)$ and $b_{1NT} = \mathbb{V}_{NT}^2 + o(1)$. Using arguments as used in the analysis of b_{1NT} and Lemma A.7, we can also show that

$$\begin{aligned}
b_{2NT} &= \frac{1}{T^4 N^2 h^2} \sum_{s_1=2}^{T} \sum_{s_2=2}^{T} \sum_{1 \leq r_1 \neq r_2 \leq s_1 - 1} \sum_{1 \leq r_3 \neq r_4 \leq s_2 - 1} \bar{k}_{s_1 r_1} \bar{k}_{s_1 r_2} \bar{k}_{s_2 r_3} \bar{k}_{s_2 r_4} E\left(\phi_{s_1 r_1} \phi_{s_1 r_2} \phi_{s_2 r_3} \phi_{s_2 r_4}\right) \\
&= O\left(T^{-1} h^{-2} + N^2 T h^2 \alpha(T_0)^{\delta/(1+\delta)} + T^{-2} T_0^4 + T^{-2} T_0^3 h^{-1} + T^{-2} T_0^2 h^{-2}\right) = o(1).
\end{aligned}$$

It follows that $E(\sum_{s=2}^{T} Z_{NT,s}^2)^2 = \mathbb{V}_{NT}^2 + o(1)$ and $\mathrm{Var}(\sum_{s=2}^{T} Z_{NT,s}^2) = o(1)$. Then the second part of (A.2) follows by Chebyshev inequality. In addition, by straightforward moment calculations, we can show that

$M_{1,1}^{(2,2)} = o_P(1)$. It follows that $M_{1,1}^{(2)} \xrightarrow{d} N(0, \mathbb{V}_0)$. For $M_{1,1}^{(3)}$, by the matrix version of Cauchy-Schwarz inequality, Jensen inequality, and Lemma A.3(v), we have

$$\begin{aligned}
\left|M_{1,1}^{(3)}\right| &= \frac{h^{1/2}}{T^2 N^{1/2}} \left|\sum_{t=1}^{T} \operatorname{tr}\left[\left(HS_{Tt}^{-1}S_{Tt}^{-1}H' - \mathbb{S}\right)\sum_{i=1}^{N}\sum_{1\le s<r\le T} k_{h,st}k_{h,rt}F_rF_s'e_{is}e_{ir}\right]\right| \\
&\le \max_t \left\|HS_{Tt}^{-1}S_{Tt}^{-1}H' - \mathbb{S}\right\| \frac{h^{1/2}}{T^2 N^{1/2}} \sum_{t=1}^{T}\left\|\sum_{i=1}^{N}\sum_{1\le s<r\le T} k_{h,st}k_{h,rt}F_rF_s'e_{is}e_{ir}\right\| \\
&\le \max_t \left\|HS_{Tt}^{-1}S_{Tt}^{-1}H' - \mathbb{S}\right\| \left\{\frac{h}{T^3 N}\sum_{t=1}^{T}\left\|\sum_{i=1}^{N}\sum_{1\le s<r\le T} k_{h,st}k_{h,rt}F_rF_s'e_{is}e_{ir}\right\|^2\right\}^{1/2} \\
&= O_P\left((Th)^{-1/2}(\ln T)^{1/2} + N^{-1/2}\right) O_P(1) = o_P(1),
\end{aligned}$$

where we also use the fact that $E\left(\frac{h}{T^3N}\sum_{t=1}^{T}\left\|\sum_{i=1}^{N}\sum_{1\le s<r\le T} k_{h,st}k_{h,rt}F_rF_s'e_{is}e_{ir}\right\|^2\right) = O(1)$ by using Lemma A.7 and arguments as used in the above study of b_{1NT}. Consequently, we have shown that $M_{1,1} - \mathbb{B}_{1NT} \xrightarrow{d} N(0, \mathbb{V}_0)$.

Next, we prove (ii). Using $a_{NT} = T^{-1/2}N^{-1/4}h^{-1/4}$, $\tilde{F}_s = H'F_s + (\tilde{F}_s - H'F_s)$, and Lemmas A.3(i), (v) and A.5(i), we have

$$\begin{aligned}
M_{1,2} &= \frac{h^{1/2}}{N^{1/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\|D_2(i,t)\|^2 = \frac{a_{NT}^2 h^{1/2}}{N^{1/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\left\|S_{Tt}^{-1}\frac{1}{T}\sum_{s=1}^{T}k_{h,st}\tilde{F}_sF_s'g_{is}\right\|^2 \\
&= \frac{1}{TN}\sum_{i=1}^{N}\sum_{t=1}^{T}\operatorname{tr}\left(S_{Tt}^{-1}\frac{1}{T}\sum_{s=1}^{T}k_{h,st}\tilde{F}_sF_s'g_{is}\frac{1}{T}\sum_{r=1}^{T}k_{h,rt}g_{ir}'F_r\tilde{F}_r'S_{Tt}^{-1}\right) \\
&= \frac{1}{TN}\sum_{i=1}^{N}\sum_{t=1}^{T}\operatorname{tr}\left(HS_{Tt}^{-1}S_{Tt}^{-1}H'\frac{1}{T}\sum_{s=1}^{T}k_{h,st}F_sF_s'g_{is}\frac{1}{T}\sum_{r=1}^{T}k_{h,rt}g_{ir}'F_rF_r'\right) + O_P\left(C_{NT}^{-2}\right) \\
&= \frac{1}{TN}\sum_{i=1}^{N}\sum_{t=1}^{T}\operatorname{tr}\left(\Sigma_F^{-1}Q_0'Q_0\Sigma_F^{-1}\Sigma_F g_i\left(\frac{t}{T}\right)g_i\left(\frac{t}{T}\right)'\Sigma_F\right) + o_P(1) \\
&= \frac{1}{TN}\sum_{i=1}^{N}\sum_{t=1}^{T}\operatorname{tr}\left(Q_0 g_i\left(\frac{t}{T}\right)g_i\left(\frac{t}{T}\right)'Q_0'\right) + o_P(1) = \Pi_{1NT} + o_P(1).
\end{aligned}$$

Now, we prove (iii). For $M_{1,3}$, we apply Lemmas A.3(i)-(ii) and (iv) and triangle inequality to obtain

$$\begin{aligned}
M_{1,3} &= \frac{h^{1/2}}{N^{1/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\|D_3(i,t)\|^2 = \frac{h^{1/2}}{N^{1/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\left\|S_{Tt}^{-1}\frac{1}{T}\sum_{s=1}^{T}k_{h,st}\tilde{F}_s\left(\tilde{F}_s - H'F_s\right)'H^{-1}\lambda_{i0}\right\|^2 \\
&\le TN^{1/2}h^{1/2}\left\|H^{-1}\right\|\max_t\left\|S_{Tt}^{-1}\right\|^2\left\{\frac{1}{N}\sum_{i=1}^{N}\|\lambda_{i0}\|^2\right\}\max_t\left\|\frac{1}{T}\sum_{s=1}^{T}k_{h,st}\tilde{F}_s\left(\tilde{F}_s - H'F_s\right)'\right\|^2 \\
&= TN^{1/2}h^{1/2}O_P(1)\left(O_P\left(T^{-1}\ln T + N^{-1}\right)^2 + o_P\left(a_{NT}^2\right)\right) = o_P(1).
\end{aligned}$$

By Lemma A.4(iv) and triangle inequality, we obtain

$$\begin{aligned} M_{1,4} &= \frac{h^{1/2}}{N^{1/2}} \sum_{i=1}^{N} \sum_{t=1}^{T} \|D_4(i,t)\|^2 = \frac{h^{1/2}}{N^{1/2}} \sum_{i=1}^{N} \sum_{t=1}^{T} \left\| S_{Tt}^{-1} \frac{1}{T} \sum_{s=1}^{T} \left(\tilde{F}_s - H' F_s \right) e_{is} k_{h,st} \right\|^2 \\ &\leq TN^{1/2} h^{1/2} \max_t \left\| S_{Tt}^{-1} \right\| \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \left\| \frac{1}{T} \sum_{s=1}^{T} \left(\tilde{F}_s - H' F_s \right) e_{is} k_{h,st} \right\|^2 \\ &= TN^{1/2} h^{1/2} O_P(1) \left(O_P\left(N^{-3/2} + T^{-2} \right) + o_P\left(a_{NT}^2 \right) \right) = o_P(1). \end{aligned}$$

For $M_{1,5}$, we have

$$\begin{aligned} |M_{1,5}| &= \frac{a_{NT} h^{1/2}}{T^2 N^{1/2}} \left| \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{s=1}^{T} k_{h,st} F_s' e_{is} H S_{Tt}^{-1} S_{Tt}^{-1} \sum_{r=1}^{T} k_{h,rt} \tilde{F}_r F_r' g_{ir} \right| \\ &\leq \max_{i,t} \left\| \frac{1}{T} \sum_{r=1}^{T} k_{h,rt} \tilde{F}_r F_r' g_{ir} \right\| \frac{h^{1/4}}{T^{3/2} N^{3/4}} \sum_{t=1}^{T} \left\| \sum_{i=1}^{N} \sum_{s=1}^{T} F_s' e_{is} k_{h,st} H S_{Tt}^{-1} S_{Tt}^{-1} \right\|. \end{aligned}$$

By Lemma A.4(i), $\max_{i,t} \left\| \frac{1}{T} \sum_{r=1}^{T} k_{h,rt} \tilde{F}_r F_r' g_{ir} \right\| = O_P(1)$. In addition,

$$\begin{aligned} & \frac{h^{1/4}}{T^{3/2} N^{3/4}} \sum_{t=1}^{T} \left\| \sum_{i=1}^{N} \sum_{s=1}^{T} k_{h,st} F_s' e_{is} H S_{Tt}^{-1} S_{Tt}^{-1} H' H'^{-1} \right\| \\ \leq\ & \frac{h^{1/4}}{T^{3/2} N^{3/4}} \sum_{t=1}^{T} \left\| \sum_{i=1}^{N} \sum_{s=1}^{T} k_{h,st} F_s' e_{is} \Sigma_F^{-1} Q_0' Q_0 \Sigma_F^{-1} \right\| \left\| H'^{-1} \right\| \\ & + \frac{h^{1/4}}{T^{3/2} N^{3/4}} \sum_{t=1}^{T} \left\| \sum_{i=1}^{N} \sum_{s=1}^{T} k_{h,st} F_s' e_{is} \left(H S_{Tt}^{-1} S_{Tt}^{-1} H' - \Sigma_F^{-1} Q_0' Q_0 \Sigma_F^{-1} \right) \right\| \left\| H'^{-1} \right\| \\ \equiv\ & \{ II_1 + II_2 \} \left\| H'^{-1} \right\|, \text{ say.} \end{aligned}$$

Noting that under Assumptions A.3(ii), (v) and A.4

$$\begin{aligned} \frac{h}{T^2 N} \sum_{t=1}^{T} E \left\| \sum_{i=1}^{N} \sum_{s=1}^{T} F_s' e_{is} k_{h,st} \right\|^2 &= \frac{h}{T^2 N} \sum_{t=1}^{T} \sum_{i=1}^{N} \sum_{s=1}^{T} \sum_{j=1}^{N} k_{h,st}^2 E\left(F_s' F_s e_{is} e_{js} \right) \\ &\leq \max_s \left\| \frac{h}{T} \sum_{t=1}^{T} k_{h,st}^2 \right\| \|\Sigma_F\|^2 \frac{1}{NT} \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{s=1}^{T} |\tau_{ij,s}| = O(1), \end{aligned}$$

we have $\frac{h^{1/2}}{TN^{1/2}} \sum_{i=1}^{N} \left\| \sum_{t=1}^{T} \sum_{s=1}^{T} F_s' e_{is} k_{h,st} \right\| \leq \left\{ \frac{h}{T^2 N} \sum_{i=1}^{N} \left\| \sum_{t=1}^{T} \sum_{s=1}^{T} k_{h,st} F_s' e_{is} \right\|^2 \right\}^{1/2} = O_P(1)$ and $II_1 = O_P(a_{NT})$. For II_2, we have by Lemmas A.4(ii) and A.3(v)

$$\begin{aligned} II_2 &\leq \frac{h^{1/4}}{T^{3/2} N^{3/4}} \sum_{t=1}^{T} \left\| \sum_{i=1}^{N} \sum_{s=1}^{T} k_{h,st} F_s' e_{is} \left(H S_{Tt}^{-1} S_{Tt}^{-1} H' - \Sigma_F^{-1} Q_0' Q_0 \Sigma_F^{-1} \right) \right\| \\ &\leq h^{1/4} T^{1/2} N^{1/4} \max_{i,t} \left\| \frac{1}{T} \sum_{s=1}^{T} k_{h,st} F_s' e_{is} \right\| \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \left\| H S_{Tt}^{-1} S_{Tt}^{-1} - \Sigma_F^{-1} Q_0' Q_0 \Sigma_F^{-1} \right\| \\ &= h^{1/4} T^{1/2} N^{1/4} O_P\left(T^{-1/2} h^{-1/2} \ln(NT) \right) O_P\left(T^{-1/2} h^{-1/2} (\ln T)^{1/2} + N^{-1/2} \right) \\ &= O_P\left(T^{-1/2} h^{-3/4} N^{1/4} \ln(NT) (\ln T)^{1/2} + N^{-1/4} h^{-1/4} \ln(NT) \right) = o_P(1). \end{aligned}$$

It follows that $M_{1,5} = o_P(1)$.

For $M_{1,6}$, we have by Lemmas A.3(i)-(ii) and (iv),

$$\begin{aligned}
|M_{1,6}| &= \frac{h^{1/2}}{T^2 N^{1/2}} \left| \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{s=1}^{T} F_s' e_{is} k_{h,st} H S_{Tt}^{-1} S_{Tt}^{-1} \sum_{r=1}^{T} k_{h,rt} \tilde{F}_r \left(\tilde{F}_r - H' F_r \right)' H^{-1} \lambda_{i0} \right| \\
&= \frac{h^{1/2}}{T^2 N^{1/2}} \left| \sum_{t=1}^{T} \operatorname{tr} \left[\left(\sum_{i=1}^{N} \sum_{s=1}^{T} \lambda_{i0} F_s' e_{is} k_{h,st} \right) H S_{Tt}^{-1} S_{Tt}^{-1} \sum_{r=1}^{T} k_{h,rt} \tilde{F}_r \left(\tilde{F}_r - H' F_r \right)' H^{-1} \right] \right| \\
&\leq \max_t \left\| S_{Tt}^{-1} \right\| \left\| H^{-1} \right\| \left\| H \right\| \max_t \left\| \frac{1}{T} \sum_{r=1}^{T} k_{h,rt} \tilde{F}_r \left(\tilde{F}_r - H' F_r \right)' \right\| \\
&\quad \times \left\{ \frac{h}{TN} \sum_{t=1}^{T} \left\| \sum_{i=1}^{N} \sum_{s=1}^{T} \lambda_{i0} F_s' e_{is} k_{h,st} \right\|^2 \right\}^{1/2} \\
&= \left[O_P \left(T^{-1} \ln T + N^{-1} \right) + o_P \left(a_{NT} \right) \right] O(T^{1/2}) = o_P(1),
\end{aligned}$$

where we also use the fact that under Assumptions A.3(ii), (v) and A.4

$$\begin{aligned}
\frac{h}{TN} \sum_{t=1}^{T} E \left\| \sum_{i=1}^{N} \sum_{s=1}^{T} \lambda_{i0} F_s' e_{is} k_{h,st} \right\|^2 &= \frac{h}{TN} \sum_{t=1}^{T} \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{s=1}^{T} k_{h,st}^2 E \left(F_s' F_s e_{is} e_{js} \right) \lambda_{j0}' \lambda_{i0} \\
&\leq CT \left(\max_s \frac{h}{T} \sum_{t=1}^{T} k_{h,st}^2 \right) \frac{1}{NT} \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{s=1}^{T} |\tau_{ij,s}| \\
&= TO(1) O(1) = O(T).
\end{aligned}$$

For $M_{1,7}$, we apply (A.8) in the supplementary appendix to make the following decomposition

$$\begin{aligned}
M_{1,7} &= \frac{h^{1/2}}{T^2 N^{1/2}} \sum_{i=1}^{N} \sum_{t=1}^{T} F_s' e_{is} k_{h,st} H S_{Tt}^{-1} \sum_{r=1}^{T} k_{h,rt} e_{ir} (\tilde{F}_r - H' F_r) \\
&= \frac{h^{1/2}}{T^2 N^{1/2}} \sum_{i=1}^{N} \sum_{t=1}^{T} F_s' e_{is} k_{h,st} H S_{Tt}^{-1} \sum_{r=1}^{T} k_{h,rt} e_{ir} \left[A_1(r) + A_2(r) + A_3(r) + A_4(r) \right] \\
&\equiv M_{1,7}^{(1)} + M_{1,7}^{(2)} + M_{1,7}^{(3)} + M_{1,7}^{(4)}, \text{ say.}
\end{aligned}$$

By straightforward calculations, we can show that $M_{1,7}^{(l)} = o_P(1)$ for $l = 1, 2, 3, 4$. It follows that $M_{1,7} = o_P(1)$.

Finally, $M_{1,8} \leq \{M_{1,2} M_{1,3}\}^{1/2} = o_P(1)$, $M_{1,9} \leq \{M_{1,2} M_{1,4}\}^{1/2} = o_P(1)$, and $M_{1,10} \leq \{M_{1,3} M_{1,4}\}^{1/2}$ $= o_P(1)$ by CS inequality and the fact that $M_{1,2} = O_P(1)$ and $M_{1,j} = o_P(1)$ for $j = 3, 4$. Consequently, $M_1 - \mathbb{B}_{1NT} - \Pi_{1NT} \xrightarrow{d} N(0, \mathbb{V}_0)$. ■

Proposition A.9 *Suppose that the conditions in Theorem 3.2 hold. Then* $M_2 - \mathbb{B}_{2NT} - \Pi_{2NT} = o_P(1)$ *under* $\mathbb{H}_1(a_{NT})$.

Proof. Using $X_{it} = F_t' \lambda_{it} + e_{it} = F_t' \lambda_{i0} + (e_{it} + a_{NT} F_t' g_{it}) = F_t' \lambda_{i0} + e_{it}^{\dagger}$ with $e_{it}^{\dagger} = e_{it} + a_{NT} F_t' g_{it}$ and by Bai (2003, p.165), we have

$$\begin{aligned}
\tilde{\lambda}_{i0} - H^{-1} \lambda_{i0} &= H' \frac{1}{T} \sum_{s=1}^{T} F_s e_{is}^{\dagger} + \frac{1}{T} \sum_{s=1}^{T} (\tilde{F}_s - H F_s) e_{is}^{\dagger} - \frac{1}{T} \tilde{F}' (\tilde{F} H^{-1} - F) \lambda_{i0} \\
&\equiv D_5(i) + D_6(i) - D_7(i), \text{ say.} \qquad \text{(A.3)}
\end{aligned}$$

By (A.3), we make the following decomposition for M_2 :

$$M_2 = N^{-1/2}h^{1/2}\sum_{i=1}^{N}\sum_{t=1}^{T}\left\|\tilde{\lambda}_{i0} - H^{-1}\lambda_{i0}\right\|^2 = TN^{-1/2}h^{1/2}\sum_{i=1}^{N}\|D_5(i) + D_6(i) - D_7(i)\|^2$$
$$= TN^{-1/2}h^{1/2}\sum_{i=1}^{N}\left[\|D_5(i)\|^2 + \|D_6(i)\|^2 + \|D_7(i)\|^2 + 2D_5(i)'D_6(i) - 2D_5(i)'D_7(i) - 2D_6(i)'D_7(i)\right]$$
$$\equiv M_{2,1} + M_{2,2} + M_{2,3} + 2M_{2,4} - 2M_{2,5} - 2M_{2,6}, \text{ say.}$$

We prove the proposition by showing that (i) $M_{2,1} - \mathbb{B}_{2NT} - \Pi_{2NT} = o_P(1)$ and (ii) $M_{2,j} = o_P(1)$ for $j = 2, 3, ..., 6$.

To prove (i), we use $e^{\dagger}_{is} = e_{is} + a_{NT}F'_s g_{is}$ and further make the following decomposition:

$$M_{2,1} = \frac{h^{1/2}}{TN^{1/2}}\sum_{i=1}^{N}\sum_{s=1}^{T}e^{\dagger}_{is}F'_sHH'\sum_{r=1}^{T}F_re^{\dagger}_{ir}$$
$$= \frac{h^{1/2}}{TN^{1/2}}\sum_{i=1}^{N}\sum_{s=1}^{T}\sum_{r=1}^{T}F'_sHH'F_re_{is}e_{ir} + \frac{a^2_{NT}h^{1/2}}{TN^{1/2}}\sum_{i=1}^{N}\sum_{s=1}^{T}\sum_{r=1}^{T}F'_sF_sg'_{is}HH'F_rF'_rg_{ir}$$
$$+ \frac{2a_{NT}h^{1/2}}{TN^{1/2}}\sum_{i=1}^{N}\sum_{s=1}^{T}\sum_{r=1}^{T}F'_sHH'F_rF'_rg_{ir}e_{is} \equiv M^{(1)}_{2,1} + M^{(2)}_{2,1} + 2M^{(3)}_{2,1}, \quad \text{say.}$$

For $M^{(1)}_{2,1}$ we make the following decomposition:

$$M^{(1)}_{2,1} = \frac{h^{1/2}}{TN^{1/2}}\sum_{i=1}^{N}\sum_{s=1}^{T}F'_sHH'F_se^2_{is} + 2T^{-1}N^{-1/2}h^{1/2}\sum_{i=1}^{N}\sum_{1\le s<r\le T}F'_sQ_0^{-1}Q_0^{-1'}F_re_{is}e_{ir}$$
$$+ 2T^{-1}N^{-1/2}h^{1/2}\sum_{i=1}^{N}\sum_{1\le s<r\le T}F'_s(HH' - Q_0^{-1}Q_0^{-1'})F_re_{is}e_{ir}$$
$$\equiv M^{(1,1)}_{2,1} + 2M^{(1,2)}_{2,1} + 2M^{(1,3)}_{2,1}, \quad \text{say.}$$

Apparently, $M^{(1,1)}_{21} = \mathbb{B}_{2NT}$. Using the fact that $H - Q_0^{-1} = O_P(C^{-1}_{NT})$ under $\mathbb{H}_1(a_{NT})$, we can show that $M^{(1,2)}_{2,1} = o_P(1)$ and $M^{(1,3)}_{2,1} = o_P(1)$ by arguments as used in the analyses of $M^{(2)}_{1,1}$ and $M^{(3)}_{1,1}$, respectively. By Lemmas A.2(vi) and A.5(ii), we have

$$\begin{aligned} M^{(2)}_{2,1} &= \frac{a^2_{NT}h^{1/2}}{TN^{1/2}}\sum_{i=1}^{N}\sum_{s=1}^{T}F'_sF_sg'_{is}HH'\sum_{r=1}^{T}F_rF'_rg_{ir} \\ &= \operatorname{tr}\left[HH'\frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{T}\sum_{r=1}^{T}F_rF'_rg_{ir}\right)\left(\frac{1}{T}\sum_{s=1}^{T}F'_sF_sg'_{is}\right)\right] \\ &= \frac{1}{N}\sum_{i=1}^{N}\operatorname{tr}\left[Q_0^{-1}(Q_0^{-1})'\Sigma_F\frac{1}{T}\sum_{r=1}^{T}g_{ir}\frac{1}{T}\sum_{s=1}^{T}g'_{is}\Sigma_F\right] + o_P(1) = \Pi_{2NT} + o_P(1). \end{aligned}$$

For $M_{2,1}^{(3)}$, we have

$$\begin{aligned}
\left|M_{2,1}^{(3)}\right| &= \frac{a_{NT}h^{1/2}}{TN^{1/2}}\left|\operatorname{tr}\left(HH'\sum_{r=1}^{T}F_rF_r'\sum_{i=1}^{N}\sum_{s=1}^{T}g_{ir}e_{is}F_s'\right)\right| \\
&\leq \|H\|^2\frac{h^{1/4}}{T^{3/2}N^{3/4}}\left\|\sum_{r=1}^{T}E\left(F_rF_r'\right)\sum_{i=1}^{N}\sum_{s=1}^{T}g_{ir}e_{is}F_s'\right\| \\
&\quad +\|H\|^2\frac{h^{1/4}}{T^{3/2}N^{3/4}}\left\|\sum_{r=1}^{T}\left[F_rF_r'-E\left(F_rF_r'\right)\right]\sum_{i=1}^{N}\sum_{s=1}^{T}g_{ir}e_{is}F_s'\right\| \equiv M_{2,1}^{(3,1)}+M_{2,1}^{(3,2)}.
\end{aligned}$$

Noting that under Assumptions A.1(i), A.3(ii), (v) and A.5(i),

$$\begin{aligned}
E\left(\frac{1}{T^{3/2}N^{3/4}}\left\|\sum_{r=1}^{T}\Sigma_F\sum_{i=1}^{N}\sum_{s=1}^{T}g_{ir}e_{is}F_s'\right\|\right)^2 &= \frac{1}{T^3N^{3/2}}\sum_{i=1}^{N}\sum_{r=1}^{T}\sum_{s=1}^{T}\sum_{j=1}^{N}\sum_{r_1=1}^{T}\operatorname{tr}(\Sigma_Fg_{ir}\tau_{ij,s}g_{jr_1}'\Sigma_F) \\
&\leq C\frac{1}{TN^{3/2}}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{s=1}^{T}|\tau_{ij,s}| = O(N^{-1/2}),
\end{aligned}$$

we have $M_{2,1}^{(3,1)} = \|H\|^2\frac{h^{1/4}}{T^{3/2}N^{3/4}}\left\|\sum_{r=1}^{T}\Sigma_F\sum_{i=1}^{N}\sum_{s=1}^{T}g_{ir}e_{is}F_s'\right\| = h^{1/4}O_P(N^{-1/4}) = o_P(1)$. Similarly, noting that $E\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\sum_{s=1}^{T}e_{is}F_s'\right\|\right)^2 = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{s=1}^{T}\sum_{r=1}^{T}E\left(F_s'F_re_{jr}e_{is}\right) = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{s=1}^{T}\tau_{ij,s}$ $= O(T/N)$, we have

$$\begin{aligned}
M_{2,1}^{(3,2)} &= \|H\|^2\frac{h^{1/4}}{T^{3/2}N^{3/4}}\left\|\sum_{i=1}^{N}\sum_{r=1}^{T}\left[F_rF_r'-E\left(F_rF_r'\right)\right]g_{ir}\sum_{s=1}^{T}e_{is}F_s'\right\| \\
&\leq \left\{\max_i\left\|\frac{1}{T}\sum_{r=1}^{T}\left[F_rF_r'-E\left(F_rF_r'\right)\right]g_{ir}\right\|\right\}\|H\|^2\frac{h^{1/4}}{T^{1/2}N^{3/4}}\sum_{i=1}^{N}\left\|\sum_{s=1}^{T}e_{is}F_s'\right\| \\
&= O_P(T^{-1/2}\ln N)O_P(h^{1/4}N^{-1/4}) = o_P(1).
\end{aligned}$$

Thus $M_{2,1} = \mathbb{B}_{2NT}+\Pi_{2NT}+o_P(1)$.

Now we prove (ii). By Lemma A.4(iii) and Lemma A.2 (iii)-(vi),

$$\begin{aligned}
M_{2,2} &= TN^{-1/2}h^{1/2}\sum_{i=1}^{N}\|D_6(i)\|^2 = TN^{1/2}h^{1/2}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{1}{T}\sum_{s=1}^{T}(\tilde{F}_s-HF_s)e_{is}^{\dagger}\right\|^2 \\
&= TN^{1/2}h^{1/2}O_P(C_{NT}^{-4}) = o_P(1) \text{ and} \\
M_{2,3} &= TN^{-1/2}h^{1/2}\sum_{i=1}^{N}\|D_7(i)\|^2 \leq TN^{1/2}h^{1/2}\left\|\frac{1}{T}\tilde{F}'(\tilde{F}H^{-1}-F)\right\|^2\frac{1}{N}\sum_{i=1}^{N}\|\lambda_{i0}\|^2 \\
&= TN^{1/2}h^{1/2}O_P(C_{NT}^{-4}) = o_P(1).
\end{aligned}$$

By CS inequality $M_{2,6}\leq\{M_{2,2}M_{2,3}\}^{1/2} = o_P(1)$. For $M_{2,4}$, we apply Lemma A.4(iii) to obtain

$$\begin{aligned}
|M_{2,4}| &= TN^{-1/2}h^{1/2}\left|\sum_{i=1}^{N}D_5(i)'D_6(i)\right| = TN^{-1/2}h^{1/2}\left|\sum_{i=1}^{N}\frac{1}{T}\sum_{s=1}^{T}e_{is}^{\dagger}F_s'H\frac{1}{T}\sum_{t=1}^{T}(\tilde{F}_t-HF_t)e_{it}^{\dagger}\right| \\
&\leq TN^{1/2}h^{1/2}\left\{\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{1}{T}\sum_{s=1}^{T}e_{is}^{\dagger}F_s'\right\|^2\right\}^{1/2}\left\{\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{1}{T}\sum_{t=1}^{T}(\tilde{F}_t-HF_t)e_{it}^{\dagger}\right\|^2\right\}^{1/2} \\
&= TN^{1/2}h^{1/2}O_P(T^{-1/2})O_P(C_{NT}^{-2}) = o_P(1),
\end{aligned}$$

where we use the fact that $\frac{1}{N}\sum_{i=1}^N E||\frac{1}{T}\sum_{s=1}^T e_{is}^{\dagger}F_s'||^2 \leq \frac{2}{N}\sum_{i=1}^N E||\frac{1}{T}\sum_{s=1}^T e_{is}F_s'||^2 + \frac{a_{NT}^2}{N}\sum_{i=1}^N E||\frac{1}{T}\sum_{s=1}^T g_{is}'F_sF_s'||^2 = O(T^{-1}+a_{NT}^2) = O(T^{-1})$. Now,

$$\begin{aligned}
M_{2,5} &= TN^{-1/2}h^{1/2}\left|\sum_{i=1}^N D_5(i)'D_7(i)\right| = TN^{-1/2}h^{1/2}\left|\sum_{i=1}^N \frac{1}{T}\sum_{s=1}^T e_{is}F_s'H\frac{1}{T}\tilde{F}'(\tilde{F}H^{-1}-F)\lambda_{i0}\right| \\
&= T^{-1}N^{-1/2}h^{1/2}\left|\operatorname{tr}\left(H\tilde{F}'(\tilde{F}H^{-1}-F)\sum_{i=1}^N\sum_{s=1}^T \lambda_{i0}F_s'e_{is}\right)\right| \\
&\leq T^{1/2}h^{1/2}\|H\|\left\{T^{-1}\left\|\tilde{F}'(\tilde{F}H^{-1}-F)\right\|\right\}\left\|N^{-1/2}T^{-1/2}\sum_{i=1}^N\sum_{s=1}^T\lambda_{i0}F_s'e_{is}\right\| \\
&= T^{1/2}h^{1/2}O_P(C_{NT}^{-2}) = o_P(1).
\end{aligned}$$

Thus we have shown that $M_{2,j} = o_P(1)$ for $j = 2, 3, ..., 6$ and the second part of the lemma follows. ■

Proposition A.10 *Suppose that the conditions in Theorem 3.2 hold. Then* $M_3 - \mathbb{B}_{3NT} - \Pi_{3NT} = o_P(1)$ *under* $\mathbb{H}_1(a_{NT})$.

Proof. By (A.1) and (A.3), we can write M_3 as follows:

$$\begin{aligned}
M_3 &= \frac{h^{1/2}}{N^{1/2}}\sum_{i=1}^N\sum_{t=1}^T\left(\hat{\lambda}_{it}-H^{-1}\lambda_{i0}\right)'\left(\tilde{\lambda}_i - H^{-1}\lambda_{i0}\right) \\
&= \frac{h^{1/2}}{N^{1/2}}\sum_{i=1}^N\sum_{t=1}^T [D_1(i,t)+D_2(i,t)-D_3(i,t)+D_4(i,t)]'[D_5(i)+D_6(i)-D_7(i)] \\
&= \frac{h^{1/2}}{N^{1/2}}\sum_{i=1}^N\sum_{t=1}^T [D_1(i,t)'D_5(i)+D_1(i,t)'D_6(i)-D_1(i,t)'D_7(i)+D_2(i,t)'D_5(i) \\
&+ D_2(i,t)'D_6(i)-D_2(i,t)'D_7(i)-D_3(i,t)'D_5(i)-D_3(i,t)'D_6(i)+D_3(i,t)'D_7(i) \\
&+ D_4(i,t)'D_5(i)+D_4(i,t)'D_6(i)-D_4(i,t)'D_7(i)] \\
&\equiv \sum_{i=1}^{12} M_{3,i}, \quad \text{say.}
\end{aligned}$$

We prove the proposition by showing that (i) $M_{3,1} = \mathbb{B}_{3NT} + o_P(1)$, (ii) $M_{3,4} = \Pi_{3NT} + o_P(1)$ and (iii) $M_{3,j} = o_P(1)$ for $j = 2, 3, 5, 6, ..., 12$.

First, we show (i). We decompose $M_{3,1}$ as follows:

$$\begin{aligned}
M_{3,1} &= \frac{h^{1/2}}{N^{1/2}}\sum_{i=1}^N\sum_{t=1}^T D_1(i,t)'D_5(i) = \frac{h^{1/2}}{T^2N^{1/2}}\sum_{i=1}^N\sum_{t=1}^T\sum_{s=1}^T\sum_{r=1}^T k_{h,st}F_s'HS_{Tt}^{-1}H'F_re_{ir}^{\dagger}e_{is} \\
&= \frac{h^{1/2}}{T^2N^{1/2}}\sum_{i=1}^N\sum_{t=1}^T\sum_{s=1}^T\sum_{r=1}^T k_{h,st}F_s'HS_{Tt}^{-1}H'F_re_{ir}e_{is} + \frac{a_{NT}h^{1/2}}{T^2N^{1/2}}\sum_{i=1}^N\sum_{t=1}^T\sum_{s=1}^T\sum_{r=1}^T k_{h,st}F_s'HS_{Tt}^{-1}H'F_rF_r'g_{ir}e_{is} \\
&= \frac{h^{1/2}}{T^2N^{1/2}}\sum_{i=1}^N\sum_{t=1}^T\sum_{s=1}^T k_{h,st}F_s'HS_{Tt}^{-1}H'F_se_{is}^2 + \frac{h^{1/2}}{T^2N^{1/2}}\sum_{i=1}^N\sum_{t=1}^T\sum_{1\leq s\neq r\leq T} k_{h,st}F_s'HS_{Tt}^{-1}H'F_re_{ir}e_{is} \\
&+ \frac{a_{NT}h^{1/2}}{T^2N^{1/2}}\sum_{i=1}^N\sum_{t=1}^T\sum_{s=1}^T\sum_{r=1}^T k_{h,st}F_s'HS_{Tt}^{-1}H'F_rF_r'g_{ir}e_{is} \\
&\equiv M_{3,1}^{(1)} + M_{3,1}^{(2)} + M_{3,1}^{(3)}, \text{ say.}
\end{aligned}$$

Apparently $M_{3,1}^{(1)} = \mathbb{B}_{3NT}$. Following the analysis of $M_{1,1}$, we can readily show that $M_{3,1}^{(2)} = O_P(h^{1/2})$ and $M_{3,1}^{(2)} = O_P(T^{-1/2}N^{1/4}h^{-1/4}) = o_P(1)$. It follows that $M_{3,1} = \mathbb{B}_{3NT} + o_P(1)$.

Next, we show (ii). Using $e_{ir}^{\dagger} = e_{ir} + a_{NT}F_r' g_{ir}$ we decompose $M_{3,4}$ as follows:

$$
\begin{aligned}
M_{3,4} &= \frac{h^{1/2}}{N^{1/2}} \sum_{i=1}^{N} \sum_{t=1}^{T} D_2(i,t)' D_5(i) = \frac{a_{NT}h^{1/2}}{T^2 N^{1/2}} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{s=1}^{T} k_{h,st} g_{is}' F_s \tilde{F}_s' S_{Tt}^{-1} H' \sum_{r=1}^{T} F_r e_{ir}^{\dagger} \\
&= \frac{a_{NT}h^{1/2}}{T^2 N^{1/2}} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{s=1}^{T} k_{h,st} g_{is}' F_s \tilde{F}_s' S_{Tt}^{-1} H' \sum_{r=1}^{T} F_r e_{ir} \\
&\quad + \frac{a_{NT}^2 h^{1/2}}{T^2 N^{1/2}} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{s=1}^{T} k_{h,st} g_{is}' F_s \tilde{F}_s' S_{Tt}^{-1} H' \sum_{r=1}^{T} F_r F_r' g_{ir} \equiv M_{3,4}^{(1)} + M_{3,4}^{(2)}, \text{ say.}
\end{aligned}
$$

For $M_{3,4}^{(1)}$, by Lemmas A.2(iv) and A.3(iv) we have

$$
\begin{aligned}
\left| M_{3,4}^{(1)} \right| &\leq \frac{a_{NT}h^{1/2}}{T^2 N^{1/2}} \left| \sum_{t=1}^{T} \sum_{i=1}^{N} \sum_{s=1}^{T} k_{h,st} g_{is}' F_s F_s' H S_{Tt}^{-1} H' \sum_{r=1}^{T} F_r e_{ir} \right| + o_P(1) \\
&= \frac{a_{NT}h^{1/2}}{T^2 N^{1/2}} \left| \sum_{t=1}^{T} \operatorname{tr} \left(H S_{Tt}^{-1} H' \sum_{i=1}^{N} \sum_{s=1}^{T} \sum_{r=1}^{T} F_r e_{ir} k_{h,st} g_{is}' F_s F_s' \right) \right| + o_P(1) \\
&= \frac{h^{1/4}}{T^{5/2} N^{3/4}} \left\| H \right\|^2 \max_t \left\| S_{Tt}^{-1} \right\| \sum_{t=1}^{T} \left\| \sum_{i=1}^{N} \sum_{s=1}^{T} k_{h,st} g_{is}' F_s F_s' \sum_{r=1}^{T} F_r e_{ir} \right\| + o_P(1) \\
&= \frac{h^{1/4}}{T^{5/2} N^{3/4}} O_P(N^{1/2} T^{5/2} h^{-1/2}) + o_P(1) = o_P(1).
\end{aligned}
$$

Noting that $H S_{Tt}^{-1} H' = Q_0^{-1} \left((Q_0^{-1})' \Sigma_F Q_0^{-1} \right)^{-1} (Q_0^{-1})' + o_P(1) = \Sigma_F^{-1} + o_P(1)$ uniformly in t by Lemmas A.2(vi) and A.3(iv), we have by Lemmas A.2(vi) and A.5(i)-(ii)

$$
\begin{aligned}
M_{3,4}^{(2)} &= \frac{1}{T^3 N} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{s=1}^{T} k_{h,st} g_{is}' F_s \tilde{F}_s' S_{Tt}^{-1} H' \sum_{r=1}^{T} F_r F_r' g_{ir} \\
&= \frac{1}{N} \sum_{i=1}^{N} \frac{1}{T^2} \sum_{t=1}^{T} \sum_{s=1}^{T} k_{h,st} g_{is}' F_s F_s' H S_{Tt}^{-1} H' \frac{1}{T} \sum_{r=1}^{T} F_r F_r' g_{ir} + o_P(1) \\
&= \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \operatorname{tr} \left[H S_{Tt}^{-1} H' \left(\frac{1}{T} \sum_{r=1}^{T} F_r F_r' g_{ir} \right) \left(\frac{1}{T} \sum_{s=1}^{T} k_{h,st} g_{is}' F_s F_s' \right) \right] \\
&= \frac{1}{N} \sum_{i=1}^{N} \operatorname{tr} \left[\Sigma_F^{-1} \Sigma_F \frac{1}{T} \sum_{r=1}^{T} g_i \left(\frac{r}{T} \right) \frac{1}{T} \sum_{t=1}^{T} g_i \left(\frac{t}{T} \right)' \Sigma_F \right] + o_P(1) = \Pi_{3NT} + o_P(1).
\end{aligned}
$$

It follows that $M_{3,4} = \Pi_{3NT} + o_P(1)$.

Now, we prove (iii). We first consider $M_{3,2}$ and $M_{3,3}$. Note that by Lemmas A.2(vi) and A.3(iv),

$$\begin{aligned}
&\frac{h^{1/2}}{T^2}\sum_{i=1}^{N}\left\|\sum_{t=1}^{T}\sum_{s=1}^{T}k_{h,st}F_s'HS_{Tt}^{-1}e_{is}\right\|^2\\
\leq\ & \frac{2h^{1/2}}{T^2}\sum_{i=1}^{N}\left\|\sum_{t=1}^{T}\sum_{s=1}^{T}k_{h,st}F_s'\Sigma_F^{-1}Q_0'e_{is}\right\|^2+\frac{2h^{1/2}}{T^2}\sum_{i=1}^{N}\left\|\sum_{t=1}^{T}\left(\sum_{s=1}^{T}k_{h,st}F_s'e_{is}\left(HS_{Tt}^{-1}-\Sigma_F^{-1}Q_0'\right)\right)\right\|^2\\
\leq\ & \frac{2h^{1/2}}{T^2}\sum_{i=1}^{N}\left\|\sum_{t=1}^{T}\sum_{s=1}^{T}k_{h,st}F_s'\Sigma_F^{-1}Q_0'e_{is}\right\|^2\\
&+2h^{1/2}N\max_{i,t}\left\|\frac{1}{T}\sum_{s=1}^{T}k_{h,st}F_s'e_{is}\right\|^2\left\|\sum_{t=1}^{T}\left\|HS_{Tt}^{-1}-\Sigma_F^{-1}Q_0'\right\|\right\|^2\\
=\ & O_P(NTh^{1/2})+h^{1/2}NO_P(T^{-1}h^{-1}\ln(NT))O_P(T^2((Th)^{-1}\ln T+N^{-1}))=O_P(NTh^{1/2}).
\end{aligned}$$

By analogous analysis as used in the study of $M_{1,1}^{(2)}$ and $M_{1,1}^{(3)}$ and Lemma A.4(ii), we have

$$\begin{aligned}
M_{3,2} &= \frac{h^{1/2}}{N^{1/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}D_1(i,t)'D_6(i)=\frac{h^{1/2}}{TN^{1/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}k_{h,st}F_s'HS_{Tt}^{-1}e_{is}\frac{1}{T}\sum_{r=1}^{T}(\tilde{F}_r-HF_r)e_{ir}^{\dagger}\\
&\leq \left\{\frac{h}{T^2}\sum_{i=1}^{N}\left\|\sum_{t=1}^{T}\sum_{s=1}^{T}k_{h,st}F_s'HS_{Tt}^{-1}e_{is}\right\|^2\right\}^{1/2}\left\{\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{1}{T}\sum_{r=1}^{T}(\tilde{F}_r-HF_r)e_{ir}^{\dagger}\right\|^2\right\}^{1/2}\\
&= O_P(N^{1/2}T^{1/2}h^{1/2})O_P(C_{NT}^{-2})=o_P(1).
\end{aligned}$$

Similarly, noting that

$$\begin{aligned}
&\frac{h^{1/2}}{TN^{1/2}}\left\|\sum_{t=1}^{T}\sum_{i=1}^{N}\sum_{s=1}^{T}k_{h,st}\lambda_{i0}F_s'HS_{Tt}^{-1}e_{is}\right\|\\
\leq\ & \frac{h^{1/2}}{TN^{1/2}}\left\|\sum_{t=1}^{T}\sum_{i=1}^{N}\sum_{s=1}^{T}k_{h,st}\lambda_{i0}F_s'\Sigma_F^{-1}Q_0'e_{is}\right\|+\frac{h^{1/2}}{TN^{1/2}}\left\|\sum_{t=1}^{T}\sum_{i=1}^{N}\sum_{s=1}^{T}k_{h,st}\lambda_{i0}F_s'\left(HS_{Tt}^{-1}-\Sigma_F^{-1}Q_0'\right)e_{is}\right\|\\
=\ & O_P(T^{1/2}h^{1/2}),
\end{aligned}$$

we have by Lemma A.2(iii)

$$\begin{aligned}
M_{3,3} &= \frac{h^{1/2}}{N^{1/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}D_1(i,t)'D_7(i)=\frac{h^{1/2}}{TN^{1/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}k_{h,st}F_s'HS_{Tt}^{-1}e_{is}\frac{1}{T}\tilde{F}'\left[\tilde{F}H^{-1}-F\right]\lambda_{i0}\\
&= \frac{h^{1/2}}{TN^{1/2}}\mathrm{tr}\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}k_{h,st}\lambda_{i0}F_s'HS_{Tt}^{-1}e_{is}\frac{1}{T}\tilde{F}'\left[\tilde{F}H^{-1}-F\right]\right)\\
&\leq \left\|\frac{1}{T}\tilde{F}'\left[\tilde{F}H^{-1}-F\right]\right\|\frac{h^{1/2}}{TN^{1/2}}\left\|\sum_{t=1}^{T}\sum_{i=1}^{N}\sum_{s=1}^{T}k_{h,st}\lambda_{i0}F_s'HS_{Tt}^{-1}e_{is}\right\|\\
&= O_P\left(C_{NT}^{-2}+a_{NT}\right)O_P(T^{1/2}h^{1/2})=o_P(1).
\end{aligned}$$

For $M_{3,5}$, $M_{3,6}$, $M_{3,8}$, $M_{3,9}$, $M_{3,11}$, and $M_{3,12}$, we apply CS inequality and the fact that $M_{1,2}=O_P(1)$, $M_{1,l}=o_P(1)$ for $l=3,4$, and $M_{2,j}=o_P(1)$ for $j=2,3$ to obtain

$$\begin{aligned}
|M_{3,5}| &\leq \{M_{1,2}M_{2,2}\}^{1/2}=o_P(1),\ |M_{3,6}|\leq\{M_{1,2}M_{2,3}\}^{1/2}=o_P(1),\ |M_{3,8}|\leq\{M_{1,3}M_{2,2}\}^{1/2}=o_P(1),\\
|M_{3,9}| &\leq \{M_{1,3}M_{2,3}\}^{1/2}=o_P(1),\ |M_{3,11}|\leq\{M_{1,4}M_{2,2}\}^{1/2}=o_P(1),\ |M_{3,12}|\leq\{M_{1,4}M_{2,3}\}^{1/2}=o_P(1).
\end{aligned}$$

For $M_{3,7}$, we have

$$
\begin{aligned}
|M_{3,7}| &= \frac{h^{1/2}}{N^{1/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}D_3(i,t)'D_5(i) = \frac{a_{NT}h^{1/2}}{T^2N^{1/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}k_{h,st}\lambda_{i0}'H'^{-1}\left(\tilde{F}_s - H'F_s\right)\tilde{F}_s'S_{Tt}^{-1}H'\sum_{r=1}^{T}F_re_{ir}^{\dagger} \\
&= \frac{h^{1/4}}{T^{5/2}N^{3/4}}\left|\sum_{t=1}^{T}\sum_{s=1}^{T}\mathrm{tr}\left(k_{h,st}H'^{-1}\left(\tilde{F}_s - H'F_s\right)\tilde{F}_s'S_{Tt}^{-1}H'\sum_{i=1}^{N}\sum_{r=1}^{T}F_re_{ir}^{\dagger}\lambda_{i0}'\right)\right| \\
&= \frac{h^{1/4}}{T^{5/2}N^{3/4}}\left|\sum_{t=1}^{T}\sum_{s=1}^{T}\mathrm{tr}\left(k_{h,st}H'^{-1}\left(\tilde{F}_s - H'F_s\right)\tilde{F}_s'S_{Tt}^{-1}H'\sum_{i=1}^{N}\sum_{r=1}^{T}F_re_{ir}^{\dagger}\lambda_{i0}'\right)\right| \\
&\leq N^{-1/4}h^{1/4}\left\|H^{-1}\right\|\left\{\frac{1}{T}\sum_{t=1}^{T}\left\|\frac{1}{T}\sum_{s=1}^{T}k_{h,st}\left(\tilde{F}_s - H'F_s\right)\tilde{F}_s'\right\|^2\right\}^{1/2}\left\{\frac{1}{T^2N}\sum_{t=1}^{T}\left\|\sum_{i=1}^{N}\sum_{r=1}^{T}F_re_{ir}^{\dagger}\lambda_{i0}'\right\|^2\right\}^{1/2} \\
&= N^{-1/4}h^{1/4}O_P\left(C_{NT}^{-2}\right)O_P\left(1+N^{1/2}T^{1/2}a_{NT}\right) = o_P\left(1\right),
\end{aligned}
$$

Similarly, we can show that $M_{3,10} = \frac{h^{1/2}}{N^{1/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}D_4(i,t)'D_5(i) = \frac{h^{1/2}}{T^2N^{1/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}k_{h,st}e_{is}(\tilde{F}_s - H'F_s)'\ S_{Tt}^{-1}H'\frac{1}{T}\sum_{r=1}^{T}F_re_{ir}^{\dagger} = o_P\left(1\right).$ Consequently, we have $M_3 = \mathbb{B}_{3NT} + \Pi_{3NT} + o_P\left(1\right).$ ■

Proposition A.11 *Suppose that the conditions in Theorem 3.2 hold. Then* $\hat{\mathbb{B}}_{NT} - \mathbb{B}_{NT} = o_P(1)$ *under* $\mathbb{H}_1\left(a_{NT}\right).$

Proof. Let $L_{st} = \left(k_{h,st}S_{Tt}^{-1} - \mathbb{I}_R\right)\left(k_{h,st}S_{Tt}^{-1} - \mathbb{I}_R\right).$ Using $\tilde{e}_{is}^2 - e_{is}^2 = \left(\tilde{e}_{is} - e_{is}\right)^2 + 2\left(\tilde{e}_{is} - e_{is}\right)e_{is}$, we have

$$
\begin{aligned}
\hat{\mathbb{B}}_{NT} - \mathbb{B}_{NT} &= \frac{h^{1/2}}{T^2N^{1/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\mathrm{tr}\left[L_{st}(\tilde{F}_s\tilde{F}_s'\tilde{e}_{is}^2 - H'F_sF_s'He_{is}^2)\right] \\
&= \frac{h^{1/2}}{T^2N^{1/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\{\mathrm{tr}(L_{st}\tilde{F}_s\tilde{F}_s')\left(\tilde{e}_{is} - e_{is}\right)^2 + 2\mathrm{tr}(L_{st}\tilde{F}_s\tilde{F}_s')\left(\tilde{e}_{is} - e_{is}\right)e_{is} \\
&\quad + \mathrm{tr}[L_{st}(\tilde{F}_s\tilde{F}_s' - H'F_sF_s'H)]e_{is}^2\} \\
&\equiv B_1 + 2B_2 + B_3,\ \text{say.}
\end{aligned}
$$

It suffices to show that (i1) $B_1 = o_P\left(1\right)$, (i2) $B_2 = o_P\left(1\right)$, and (i3) $B_3 = o_P\left(1\right).$

We first show (i1). We make the following decomposition:

$$
\begin{aligned}
e_{is} - \tilde{e}_{is} &= \tilde{\lambda}_{i0}'\tilde{F}_s - \lambda_{is}'F_s = \tilde{\lambda}_{i0}'\tilde{F}_s - \lambda_{i0}'{H'}^{-1}H'F_s - a_{NT}F_t'g_{is} \\
&= (\tilde{\lambda}_{i0} - H^{-1}\lambda_{i0})'\tilde{F}_s + \lambda_{i0}'{H'}^{-1}(\tilde{F}_s - H'F_s) - a_{NT}F_s'g_{is} \equiv d_{1is} + d_{2is} - d_{3is},\ \text{say.}
\end{aligned}
\tag{A.4}
$$

By CS inequality, $B_1 \leq \frac{3h^{1/2}}{T^2N^{1/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\mathrm{tr}\left(L_{st}\tilde{F}_s\tilde{F}_s'\right)\left(d_{1is}^2 + d_{2is}^2 + d_{3is}^2\right) \equiv 3B_{1,1} + 3B_{1,2} + 3B_{1,3}$, say. By Lemmas A.6(i) and (iv),

$$
\begin{aligned}
B_{1,1} &= \frac{h^{1/2}}{T^2N^{1/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\tilde{F}_s'L_{st}\tilde{F}_s\tilde{F}_s'(\tilde{\lambda}_i - H^{-1}\lambda_{i0})(\tilde{\lambda}_i - H^{-1}\lambda_{i0})'\tilde{F}_s \\
&\leq N^{1/2}h^{1/2}\left\{\frac{1}{N}\sum_{i=1}^{N}\left\|\tilde{\lambda}_i - H^{-1}\lambda_{i0}\right\|^2\right\}\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\tilde{F}_s'L_{st}\tilde{F}_s\tilde{F}_s'\tilde{F}_s \\
&= N^{1/2}h^{1/2}O_P(C_{NT}^{-2})O_P(h^{-1}) = o_P\left(1\right).
\end{aligned}
$$

Noting that $L_{st} \leq k_{h,st}^2 S_{Tt}^{-1} S_{Tt}^{-1} + \mathbb{I}_R$, by Lemmas A.3(iv) and A.6(iii) and (v)

$$\begin{aligned} B_{1,2} &= \frac{h^{1/2}}{T^2 N^{1/2}} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{s=1}^{T} \operatorname{tr}\left[L_{st} \tilde{F}_s (\tilde{F}_s - H' F_s)' H^{-1} \lambda_{i0} \lambda_{i0}' H'^{-1} (\tilde{F}_s - H' F_s) \tilde{F}_s' \right] \\ &\leq \frac{h^{1/2}}{T^2 N^{1/2}} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{s=1}^{T} \operatorname{tr}\left[\left(k_{h,st}^2 S_{Tt}^{-1} S_{Tt}^{-1} + \mathbb{I}_R \right) \tilde{F}_s (\tilde{F}_s - H' F_s)' H^{-1} \lambda_{i0} \lambda_{i0}' H'^{-1} (\tilde{F}_s - H' F_s) \tilde{F}_s' \right] \\ &\leq N^{1/2} h^{1/2} \left\| H^{-1} \right\|^2 c_{1NT} \left\{ \frac{1}{N} \sum_{i=1}^{N} \left\| \lambda_{i0} \right\|^2 \right\} \frac{1}{T} \sum_{s=1}^{T} \left\| \tilde{F}_s (\tilde{F}_s - H' F_s)' \right\|^2 \\ &= N^{1/2} h^{1/2} O_P\left(h^{-1} \right) O\left(1 \right) O_P(C_{NT}^{-2} + T^{-1} N^{-2}) = o_P\left(1 \right), \end{aligned}$$

and

$$\begin{aligned} B_{1,3} &= \frac{h^{1/2}}{T^2 N^{1/2}} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{s=1}^{T} \tilde{F}_s' L_{st} \tilde{F}_s d_{3is}^2 = \frac{a_{NT}^2 h^{1/2}}{T^2 N^{1/2}} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{s=1}^{T} \tilde{F}_s' L_{st} \tilde{F}_s \left(F_s' g_{is} \right)^2 \\ &\leq \frac{1}{T^3 N} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{s=1}^{T} \operatorname{tr}\left[\left(k_{h,st}^2 S_{Tt}^{-1} S_{Tt}^{-1} + \mathbb{I}_R \right) \tilde{F}_s \tilde{F}_s' \left(F_s' g_{is} \right)^2 \right] \\ &\leq c_{1NT} \frac{1}{T^2 N} \sum_{i=1}^{N} \sum_{s=1}^{T} \left\| \tilde{F}_s \right\|^2 \left\| F_s \right\|^2 \left\| g_{is} \right\|^2 \leq \frac{\bar{c}_g^2 c_{1NT}}{T} \left\{ \frac{1}{T} \sum_{s=1}^{T} \left\| \tilde{F}_s \right\|^4 \frac{1}{T} \sum_{s=1}^{T} \left\| F_s \right\|^4 \right\}^{1/2} \\ &= O_P\left(T^{-1} h^{-1} \right) O_P\left(1 \right) = o_P\left(1 \right), \end{aligned}$$

where $c_{1NT} \equiv \max_t \left\| S_{Tt}^{-1} \right\|^2 \max_s \frac{1}{T} \sum_{t=1}^{T} k_{h,st}^2 + 1 = O_P\left(h^{-1} \right)$.

Next, we show (i2). Using (A.4), we decompose B_2 as follows

$$\begin{aligned} B_2 &= \frac{h^{1/2}}{T^2 N^{1/2}} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{s=1}^{T} \operatorname{tr}(L_{st} \tilde{F}_s \tilde{F}_s') \left(\tilde{e}_{is} - e_{is} \right) e_{is} \\ &= \frac{h^{1/2}}{T^2 N^{1/2}} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{s=1}^{T} \operatorname{tr}(L_{st} \tilde{F}_s \tilde{F}_s') \left(-d_{1is} - d_{2is} + d_{3is} \right) e_{is} \equiv -B_{2,1} - B_{2,2} + B_{2,3}, \text{ say.} \end{aligned}$$

By (A.3), we further decompose $B_{2,1}$:

$$\begin{aligned} B_{2,1} &= \frac{h^{1/2}}{T^2 N^{1/2}} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{s=1}^{T} \operatorname{tr}(L_{st} \tilde{F}_s \tilde{F}_s') \tilde{F}_s' \left(\tilde{\lambda}_i - H^{-1} \lambda_{i0} \right) e_{is} \\ &= \frac{h^{1/2}}{T^2 N^{1/2}} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{s=1}^{T} \operatorname{tr}(L_{st} \tilde{F}_s \tilde{F}_s') \tilde{F}_s' \left[D_5(i) + D_6(i) - D_7(i) \right] e_{is} \equiv B_{2,1}^{(1)} + B_{2,1}^{(2)} - B_{2,1}^{(3)}, \text{ say.} \end{aligned}$$

For $B_{2,1}^{(1)}$, we have

$$\begin{aligned} B_{2,1}^{(1)} &= \frac{h^{1/2}}{T^2 N^{1/2}} \sum_{t=1}^{T} \sum_{s=1}^{T} \operatorname{tr}(L_{st} \tilde{F}_s \tilde{F}_s') \tilde{F}_s' H' \left(\frac{1}{T} \sum_{i=1}^{N} \sum_{r=1}^{T} F_r e_{ir}^{\dagger} e_{is} \right) \\ &\leq \left\| H \right\| \left\{ \frac{h}{T^3} \sum_{s=1}^{T} \left\| \sum_{t=1}^{T} \operatorname{tr}(L_{st} \tilde{F}_s \tilde{F}_s') \tilde{F}_s' \right\|^2 \right\}^{1/2} \left\{ \frac{1}{NT} \sum_{s=1}^{T} \left\| \frac{1}{T} \sum_{i=1}^{N} \sum_{r=1}^{T} F_r e_{ir}^{\dagger} e_{is} \right\|^2 \right\}^{1/2}. \end{aligned}$$

Using $L_{st} \leq k_{h,st}^2 S_{Tt}^{-1} S_{Tt}^{-1} + \mathbb{I}_R$ and Lemma A.6(ii),

$$\begin{aligned}
\frac{h}{T^3}\sum_{s=1}^T \left\|\sum_{t=1}^T \mathrm{tr}(L_{st}\tilde{F}_s\tilde{F}_s')\tilde{F}_s'\right\|^2 &\leq \frac{h}{T^3}\sum_{s=1}^T \left\|\sum_{t=1}^T \mathrm{tr}[\left(k_{h,st}^2 S_{Tt}^{-1} S_{Tt}^{-1} + \mathbb{I}_R\right)\tilde{F}_s\tilde{F}_s']\tilde{F}_s'\right\|^2 \leq c_{1NT}^2 \frac{h}{T}\sum_{s=1}^T \left\|\tilde{F}_s\right\|^6 \\
&= O_P((Th)^{-1}).
\end{aligned}$$

In addition, $\frac{1}{NT}\sum_{s=1}^T E\left\|\frac{1}{T}\sum_{i=1}^N\sum_{r=1}^T F_r e_{ir}^{\dagger} e_{is}\right\|^2 = O\left(T^{-1} + NT^{-2} + a_{NT}^2(1+N/T)\right) = o(1)$. It follows that $B_{2,1}^{(1)} = o_P(1)$. For $B_{2,1}^{(2)}$ and $B_{2,1}^{(3)}$, we have by Lemmas A.2(iii) and A.6(ii), and the proof of Lemma A.6(i),

$$\begin{aligned}
B_{2,1}^{(2)} &= \frac{h^{1/2}}{T^2N^{1/2}}\sum_{i=1}^N\sum_{t=1}^T\sum_{s=1}^T \mathrm{tr}(L_{st}\tilde{F}_s\tilde{F}_s')\tilde{F}_s' D_6(i)e_{is} \\
&= \frac{h^{1/2}}{T^2N^{1/2}}\sum_{t=1}^T\sum_{s=1}^T \mathrm{tr}(L_{st}\tilde{F}_s\tilde{F}_s')\tilde{F}_s'\sum_{i=1}^N\frac{1}{T}\sum_{r=1}^T(\tilde{F}_r - HF_r)e_{ir}^{\dagger}e_{is} \\
&\leq \frac{h^{1/2}}{T^2N^{1/2}}\sum_{t=1}^T\sum_{s=1}^T \mathrm{tr}[\left(k_{h,st}^2 S_{Tt}^{-1} S_{Tt}^{-1} + \mathbb{I}_R\right)\tilde{F}_s\tilde{F}_s']\tilde{F}_s'\sum_{i=1}^N\frac{1}{T}\sum_{r=1}^T(\tilde{F}_r - HF_r)e_{ir}^{\dagger}e_{is} \\
&= \frac{h^{1/2}}{T^2N^{1/2}}\sum_{s=1}^T\sum_{t=1}^T \mathrm{tr}[\left(k_{h,st}^2 S_{Tt}^{-1} S_{Tt}^{-1} + \mathbb{I}_R\right)\tilde{F}_s\tilde{F}_s']\tilde{F}_s'\left\|\sum_{i=1}^N\frac{1}{T}\sum_{r=1}^T(\tilde{F}_r - HF_r)e_{ir}^{\dagger}e_{is}\right\| \\
&\leq c_{1NT}N^{1/2}h^{1/2}\left\{\frac{1}{T}\sum_{s=1}^T\left\|\tilde{F}_s\right\|^3\right\}\left\{\frac{1}{N}\sum_{i=1}^N\left\|\frac{1}{T}\sum_{r=1}^T(\tilde{F}_r - HF_r)e_{ir}^{\dagger}\right\|^2\right\}^{1/2}\max_s\left\{\frac{1}{N}\sum_{i=1}^N e_{is}^2\right\}^{1/2} \\
&= O_P(N^{1/2}h^{-1/2})O_P(1)\,O_P\left(C_{NT}^{-2}\right)O_P(1) = o_P(1),
\end{aligned}$$

and

$$\begin{aligned}
B_{2,1}^{(3)} &= \frac{h^{1/2}}{T^2N^{1/2}}\sum_{i=1}^N\sum_{t=1}^T\sum_{s=1}^T \mathrm{tr}(L_{st}\tilde{F}_s\tilde{F}_s')\tilde{F}_s' D_7(i)e_{is} \\
&= \frac{h^{1/2}}{T^2N^{1/2}}\sum_{t=1}^T\sum_{s=1}^T \mathrm{tr}(L_{st}\tilde{F}_s\tilde{F}_s')\tilde{F}_s'\frac{1}{T}\tilde{F}'(\tilde{F}H^{-1} - F)\sum_{i=1}^N\lambda_{i0}e_{is} \\
&\leq N^{1/2}h^{-1/2}\frac{1}{T}\left\|\tilde{F}'(\tilde{F}H^{-1} - F)\right\|\max_s\left\|\frac{1}{N}\sum_{i=1}^N\lambda_{i0}e_{is}\right\|\frac{1}{T^2}\sum_{t=1}^T\sum_{s=1}^T\left\|\mathrm{tr}(L_{st}\tilde{F}_s\tilde{F}_s')\tilde{F}_s'\right\| \\
&\leq c_{1NT}N^{1/2}h^{1/2}\frac{1}{T}\left\|\tilde{F}'(\tilde{F}H^{-1} - F)\right\|\max_s\left\|\frac{1}{N}\sum_{i=1}^N\lambda_{i0}e_{is}\right\|\frac{1}{T}\sum_{s=1}^T\left\|\tilde{F}_s\right\|^3 \\
&= O_P(N^{1/2}h^{-1/2})\left(O_P(C_{NT}^{-2}) + o_P(a_{NT})\right)O_P(N^{-1/2}\ln T)O_P(1) = o_P(1).
\end{aligned}$$

Thus $B_{2,1} = o_P(1)$. By Lemma A.6(vi),

$$\begin{aligned}
|B_{2,2}| &= \frac{h^{1/2}}{T^2N^{1/2}}\left|\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\operatorname{tr}(L_{st}\tilde{F}_s\tilde{F}_s')\lambda_{i0}'H'^{-1}(\tilde{F}_s - H'F_s)e_{is}\right| \\
&\leq \frac{h^{1/2}}{TN^{1/2}}\left|\sum_{s=1}^{T}\left[\frac{1}{T}\sum_{t=1}^{T}\operatorname{tr}(L_{st}\tilde{F}_s\tilde{F}_s')\right]\left[\sum_{i=1}^{N}\lambda_{i0}'H'^{-1}(\tilde{F}_s - H'F_s)e_{is}\right]\right| \\
&\leq h^{1/2}\left\{\frac{1}{T}\sum_{s=1}^{T}\left\|\frac{1}{T}\sum_{t=1}^{T}\operatorname{tr}(L_{st}\tilde{F}_s\tilde{F}_s')\right\|^2\right\}^{1/2}\left\{\frac{1}{NT}\sum_{s=1}^{T}\left\|\sum_{i=1}^{N}\lambda_{i0}'H'^{-1}(\tilde{F}_s - H'F_s)e_{is}\right\|^2\right\}^{1/2} \\
&\leq c_{1NT}h^{1/2}\left\{\frac{1}{T}\sum_{s=1}^{T}\left\|\tilde{F}_s\right\|^2\right\}^{1/2}\left\{\frac{1}{NT}\sum_{s=1}^{T}\left\|\sum_{i=1}^{N}\lambda_{i0}'H'^{-1}(\tilde{F}_s - H'F_s)e_{is}\right\|^2\right\}^{1/2} \\
&= O_P(h^{-1/2})O_P(1)O_P(C_{NT}^{-1}) = o_P(1).
\end{aligned}$$

In addition,

$$\begin{aligned}
B_{2,3} &= \frac{a_{NT}h^{1/2}}{T^2N^{1/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\operatorname{tr}(L_{st}\tilde{F}_s\tilde{F}_s')F_s'g_{is}e_{is} = \frac{N^{1/4}h^{1/4}}{T^{5/2}}\sum_{t=1}^{T}\sum_{s=1}^{T}\operatorname{tr}(L_{st}\tilde{F}_s\tilde{F}_s')F_s'\left(\frac{1}{N}\sum_{i=1}^{N}g_{is}e_{is}\right) \\
&\leq c_{1NT}\frac{N^{1/4}h^{1/4}}{T^{1/2}}\max_s\left|\frac{1}{N}\sum_{i=1}^{N}g_{is}e_{is}\right|\left\{\frac{1}{T}\sum_{s=1}^{T}\left\|\tilde{F}_s\right\|^4\frac{1}{T}\sum_{s=1}^{T}\|F_s\|^2\right\}^{1/2} \\
&= N^{1/4}h^{-3/4}T^{-1/2}O_P(N^{-1/2}\ln T)O_P(1) = o_P(1).
\end{aligned}$$

Thus $B_2 = o_P(1)$.

Now, we show (i3). For B_3, we use the definition of L_{st} and make the following decomposition:

$$\begin{aligned}
B_3 &= \frac{h^{1/2}}{T^2N^{1/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\operatorname{tr}\left[L_{st}(\tilde{F}_s\tilde{F}_s' - H'F_sF_s'H)\right]e_{is}^2 \\
&= \frac{h^{1/2}}{T^2N^{1/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}k_{h,st}^2\operatorname{tr}\left[S_{Tt}^{-1}S_{Tt}^{-1}(\tilde{F}_s\tilde{F}_s' - H'F_sF_s'H)\right]e_{is}^2 \\
&\quad -\frac{2h^{1/2}}{T^2N^{1/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}k_{h,st}\operatorname{tr}\left[S_{Tt}^{-1}(\tilde{F}_s\tilde{F}_s' - H'F_sF_s'H)\right]e_{is}^2 + \frac{h^{1/2}}{T^2N^{1/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\operatorname{tr}(\tilde{F}_s\tilde{F}_s' - H'F_sF_s'H)e_{is}^2 \\
&\equiv B_{3,1} + B_{3,2} + B_{3,3}, \text{ say.}
\end{aligned}$$

Using $\tilde{F}_s\tilde{F}_s' - H'F_sF_s'H = (\tilde{F}_s - H'F_s)(\tilde{F}_s - H'F_s)' + (\tilde{F}_s - H'F_s)F_s'H + H'F_s(\tilde{F}_s - H'F_s)'$, we can decompose $B_{3,1}$ as follows

$$\begin{aligned}
|B_{3,1}| &\leq \frac{h^{1/2}}{T^2N^{1/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}k_{h,st}^2\operatorname{tr}\left[S_{Tt}^{-1}S_{Tt}^{-1}(\tilde{F}_s - H'F_s)(\tilde{F}_s - H'F_s)'\right]e_{is}^2 \\
&\quad +\frac{2h^{1/2}}{T^2N^{1/2}}\left|\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}k_{h,st}^2\operatorname{tr}\left[S_{Tt}^{-1}S_{Tt}^{-1}(\tilde{F}_s - H'F_s)F_s'H\right]e_{is}^2\right| \equiv B_{3,1}^{(1)} + 2B_{3,1}^{(2)}.
\end{aligned}$$

By Lemma A.2(i) and the fact that $\max_i \frac{1}{N}\sum_{i=1}^N e_{is}^2 = O_P(1)$, we have

$$
\begin{aligned}
B_{3,1}^{(1)} &\leq c_{1NT} N^{1/2} h^{1/2} \left\{ \max_i \frac{1}{N}\sum_{i=1}^N e_{is}^2 \right\} \left\{ \frac{1}{T}\sum_{s=1}^T \left\| \tilde{F}_s - H' F_s \right\|^2 \right\} \\
&= O_P(N^{1/2}h^{-1/2}) O_P(1) O_P(C_{NT}^{-2}) = o_P(1).
\end{aligned}
$$

In addition, by Lemmas A.3(iv), A.6 (iii) and (vii), we can readily show that

$$
\begin{aligned}
B_{3,1}^{(2)} &= \frac{h^{1/2}}{T^2 N^{1/2}} \frac{1}{T}\sum_{t=1}^T k_{h,st}^2 \operatorname{tr}\left(S_{Tt}^{-1} S_{Tt}^{-1} \sum_{i=1}^N \sum_{s=1}^T (\tilde{F}_s - H'F_s) F_s' H e_{is}^2 \right) \\
&\leq N^{1/2} h^{1/2} \left\{ \max_t \left\| S_{Tt}^{-1} \right\|^2 \right\} \left\{ \max_s \frac{1}{T}\sum_{t=1}^T k_{h,st}^2 \right\} \left\| \frac{1}{NT} \sum_{i=1}^N \sum_{s=1}^T (\tilde{F}_s - H'F_s) F_s' H e_{is}^2 \right\| \\
&\leq N^{1/2} h^{1/2} O_P(1) O\left(h^{-1}\right) O_P(1) O_P(a_{NT}) = O_P(T^{-1/2} N^{-1/4} h^{-3/4}) = o_P(1).
\end{aligned}
$$

Thus $B_{3,1} = o_P(1)$. Similarly, we have $B_{3,l} = o_P(1)$ for $l = 2, 3$. Then $B_3 = o_P(1)$. This completes the proof of Proposition A.11. ■

Proposition A.12 *Suppose that the conditions in Theorem 3.2 hold. Then* $\hat{\mathbb{V}}_{NT} = \mathbb{V}_{NT} + o_P(1)$ *under* $\mathbb{H}_1(a_{NT})$.

Proof. Let $\bar{k}_{sr} = \bar{K}\left(\frac{s-r}{Th}\right)$. Observe that $\hat{\mathbb{V}}_{NT} - \mathbb{V}_{NT} = \mathbb{V}_{1NT} + \mathbb{V}_{2NT}$, where

$$
\begin{aligned}
\mathbb{V}_{1NT} &= 2T^{-2}N^{-1}h^{-1} \sum_{1 \leq s \neq r \leq T} \bar{k}_{sr}^2 \left[\tilde{F}_r' \tilde{\Sigma}_F^{-1} \tilde{\Sigma}_F^{-1} \tilde{F}_s \tilde{F}_s' \tilde{\Sigma}_F^{-1} \tilde{\Sigma}_F^{-1} \tilde{F}_r \left(\tilde{e}_r' \tilde{e}_s\right)^2 - F_r' \mathbb{S} F_s F_s' \mathbb{S} F_r \left(e_r' e_s\right)^2 \right], \\
\mathbb{V}_{2NT} &= 2T^{-2}N^{-1}h^{-1} \sum_{1 \leq s \neq r \leq T} \bar{k}_{sr}^2 \left[F_r' \mathbb{S} F_s F_s' \mathbb{S} F_r \left(e_r' e_s\right)^2 - E(F_r' \mathbb{S} F_s F_s' \mathbb{S} F_r \left(e_r' e_s\right)^2) \right].
\end{aligned}
$$

Using $\left(\tilde{e}_r'\tilde{e}_s\right)^2 - \left(e_r' e_s\right)^2 = \left(\tilde{e}_r'\tilde{e}_s - e_r' e_s\right)^2 + 2\left(\tilde{e}_r'\tilde{e}_s - e_r' e_s\right) e_r' e_s$ we can decompose $\mathbb{V}_{1NT}$ as follows:

$$
\begin{aligned}
\mathbb{V}_{1NT} &= 2T^{-2}N^{-1}h^{-1} \sum_{1 \leq s \neq r \leq T} \bar{k}_{sr}^2 \tilde{F}_r' \tilde{\Sigma}_F^{-1} \tilde{\Sigma}_F^{-1} \tilde{F}_s \tilde{F}_s' \tilde{\Sigma}_F^{-1} \tilde{\Sigma}_F^{-1} \tilde{F}_r \left(\tilde{e}_r'\tilde{e}_s - e_r' e_s\right)^2 \\
&\quad + 4T^{-2}N^{-1}h^{-1} \sum_{1 \leq s \neq r \leq T} \bar{k}_{sr}^2 \tilde{F}_r' \tilde{\Sigma}_F^{-1} \tilde{\Sigma}_F^{-1} \tilde{F}_s \tilde{F}_s' \tilde{\Sigma}_F^{-1} \tilde{\Sigma}_F^{-1} \tilde{F}_r \left(\tilde{e}_r'\tilde{e}_s - e_r' e_s\right) e_r' e_s \\
&\quad + 2T^{-2}N^{-1}h^{-1} \sum_{1 \leq s \neq r \leq T} \bar{k}_{sr}^2 (\tilde{F}_r' \tilde{\Sigma}_F^{-1} \tilde{\Sigma}_F^{-1} \tilde{F}_s \tilde{F}_s' \tilde{\Sigma}_F^{-1} \tilde{\Sigma}_F^{-1} \tilde{F}_r - F_r' \mathbb{S} F_s F_s' \mathbb{S} F_r) \left(e_r' e_s\right)^2 \\
&\equiv 2\mathbb{V}_{1NT,1} + 4\mathbb{V}_{1NT,2} + 2\mathbb{V}_{1NT,3}, \text{ say.}
\end{aligned}
$$

Using (A.4) and following the analysis in proving (i), we can readily show that $\mathbb{V}_{1NT,l} = o_P(1)$ for $l = 1, 2$. For $\mathbb{V}_{1NT,3}$, using $\tilde{a}'\tilde{a} - a'a = (\tilde{a} - a)'(\tilde{a} - a) + (\tilde{a} - a)' a + a'(\tilde{a} - a)$, we decompose it as follows:

$$
\begin{aligned}
\mathbb{V}_{1NT,3} &= T^{-2}N^{-1}h^{-1} \sum_{1 \leq s \neq r \leq T} \bar{k}_{sr}^2 (\tilde{F}_r' \tilde{\Sigma}_F^{-1} \tilde{\Sigma}_F^{-1} \tilde{F}_s - F_r' \mathbb{S} F_s)(\tilde{F}_r' \tilde{\Sigma}_F^{-1} \tilde{\Sigma}_F^{-1} \tilde{F}_s - F_r' \mathbb{S} F_s)' \left(e_r' e_s\right)^2 \\
&\quad + 2T^{-2}N^{-1}h^{-1} \sum_{1 \leq s \neq r \leq T} \bar{k}_{sr}^2 (\tilde{F}_r' \tilde{\Sigma}_F^{-1} \tilde{\Sigma}_F^{-1} \tilde{F}_s - F_r' \mathbb{S} F_s) F_r' \mathbb{S} F_s \left(e_r' e_s\right)^2 \\
&\equiv \mathbb{V}_{1NT,3}^{(1)} + 2\mathbb{V}_{1NT,3}^{(2)}, \text{ say.}
\end{aligned}
$$

Noting that $\tilde{F}_r'\tilde{\Sigma}_F^{-1}\tilde{\Sigma}_F^{-1}\tilde{F}_s - F_r'\mathbb{S}F_s = \tilde{F}_r'\tilde{\Sigma}_F^{-1}\tilde{\Sigma}_F^{-1}(\tilde{F}_s - H'F_s) + \tilde{F}_r'H^{-1}(H\tilde{\Sigma}_F^{-1}\tilde{\Sigma}_F^{-1}H' - \mathbb{S})F_s + (\tilde{F}_r'H^{-1} - F_r)\mathbb{S}F_s$, we have

$$\begin{aligned}
\mathbb{V}_{1NT,3}^{(1)} \leq\ & 3T^{-2}N^{-1}h^{-1}\sum_{1\leq s\neq r\leq T}\bar{k}_{sr}^2\left\|\tilde{F}_r'\tilde{\Sigma}_F^{-1}\tilde{\Sigma}_F^{-1}(\tilde{F}_s - H'F_s)\right\|^2(e_r'e_s)^2 \\
&+3T^{-2}N^{-1}h^{-1}\sum_{1\leq s\neq r\leq T}\bar{k}_{sr}^2\left\|\tilde{F}_r'H^{-1}(H\tilde{\Sigma}_F^{-1}\tilde{\Sigma}_F^{-1}H' - \mathbb{S})F_s\right\|^2(e_r'e_s)^2 \\
&+3T^{-2}N^{-1}h^{-1}\sum_{1\leq s\neq r\leq T}\bar{k}_{sr}^2\left\|(\tilde{F}_r'H^{-1} - F_r)\mathbb{S}F_s\right\|^2(e_r'e_s)^2.
\end{aligned}$$

Using Lemma A.2, we can readily show that each term in the last expression is $o_P(1)$. Then we have $\mathbb{V}_{1NT,3}^{(1)} = o_P(1)$. Similarly, we can show $\mathbb{V}_{1NT,3}^{(2)} = o_P(1)$. So $\mathbb{V}_{1NT,3} = o_P(1)$ and $\mathbb{V}_{1NT} = o_P(1)$.

In addition, noting that $E(\mathbb{V}_{2NT}) = 0$ and $\mathrm{Var}(\mathbb{V}_{2NT}) = o(1)$, we have $\mathbb{V}_{2NT} = o_P(1)$. Thus, $\hat{\mathbb{V}}_{NT} = \mathbb{V}_{NT} + o_P(1)$. ■

Proof of Theorem 3.3. Let P^* denote the probability measure induced by the modified parametric bootstrap conditional on the original sample $\mathcal{W}_{NT}$. Let E^* and Var^* denote the expectation and variance under P^*. Let $O_{P^*}(\cdot)$ and $o_{P^*}(\cdot)$ denote the probability order under P^*, e.g., $b_{NT} = o_{P^*}(1)$ if for any $\epsilon > 0$, $P^*(\|b_{NT}\| > \epsilon) = o_P(1)$. The proof is similar to but much simpler than that of Theorem 3.2 for three reasons: (1) the null hypothesis is satisfied in the bootstrap world, (2) e_t^*'s are independent over t conditional on $\mathcal{W}_{NT}$, and (3) both $\tilde{\lambda}_{i0}$ and $\tilde{F}_t$ are fixed given $\mathcal{W}_{NT}$. Even though $\tilde{\lambda}_{i0}$ and $\tilde{F}_t$ are not uniformly bounded over i or t, we can use arguments as used in the proof of Lemma A.6(i) to demonstrate that $\frac{1}{T}\sum_{t=1}^T||\tilde{F}_t||^8 = O_P(1) + O_P(T^3C_{NT}^{-8}) = O_P(1)$ and that $\frac{1}{N}\sum_{i=1}^N||\tilde{\lambda}_{i0}||^8 = O_P(1)$. These are sufficient for the analysis of $\hat{J}_{NT}^*$.

Let $\tilde{\lambda}_{i0}^*$, $\tilde{F}_t^*$, and $\tilde{\lambda}_{it}^*$ denote the bootstrap analogue of $\tilde{\lambda}_{i0}$, $\tilde{F}_t$, and $\tilde{\lambda}_{it}$, respectively. Let $\hat{M}^*$, J_{NT}^*, $\mathbb{B}_{NT}^*$, $\mathbb{V}_{NT}^*$, $\hat{J}_{NT}^*$, $\hat{\mathbb{B}}_{NT}^*$, and $\hat{\mathbb{V}}_{NT}^*$ denote the bootstrap analogue of $\hat{M}$, J_{NT}, $\mathbb{B}_{NT}$, $\mathbb{V}_{NT}$, $\hat{J}_{NT}$, $\hat{\mathbb{B}}_{NT}$, and $\hat{\mathbb{V}}_{NT}$, respectively. Then $J_{NT}^* \equiv (TN^{1/2}h^{1/2}\hat{M}^* - \mathbb{B}_{NT}^*)/\sqrt{\mathbb{V}_{NT}^*}$ and $\hat{J}_{NT}^* \equiv (N^{-1/2}\hat{M}^* - \hat{\mathbb{B}}_{NT}^*)/\sqrt{\hat{\mathbb{V}}_{NT}^*}$. Following the proof of Theorem 3.2, we can show that $TN^{1/2}h^{1/2}\hat{M}^* - \mathbb{B}_{NT}^* = \sum_{s=2}^T Z_{NT,s}^* + o_{P^*}(1)$, where $Z_{NT,s}^* = 2T^{-1}N^{-1/2}h^{-1/2}\sum_{r=1}^{s-1}\bar{k}_{sr}\tilde{F}_s'\mathbb{S}^*\tilde{F}_r e_s^{*\prime}e_r^*$, $e_s^* = (e_{N1}^*, ..., e_{Ns}^*)'$, and $\mathbb{S}^* = HS_{Tt}^{-1}S_{Tt}^{-1}H'$. Then we can prove the theorem by showing that: (i) $\sum_{s=2}^T Z_{NT,s}^*/\sqrt{\mathbb{V}_{NT}^*} \xrightarrow{D^*} N(0,1)$, (ii) $\hat{\mathbb{B}}_{NT}^* = \mathbb{B}_{NT}^* + o_{P^*}(1)$, and (iii) $\hat{\mathbb{V}}_{NT}^* = \mathbb{V}_{NT}^* + o_{P^*}(1)$.

We only outline the proof of (i) as those of other parts are analogous to the corresponding parts in the proof of Theorem 3.2. Noting that $\{Z_{NT,t}^*, \mathcal{F}_{NT,t}^*\}$ is an m.d.s., we can continue to apply the martingale CLT by showing that

$$\mathcal{Z}^* \equiv \sum_{t=2}^T E_{\mathcal{F}_{NT,t-1}^*}^*\left|Z_{NT,t}^*\right|^4 = o_{P^*}(1), \text{ and } \sum_{t=2}^T Z_{NT,t}^{*2} - \mathbb{V}_{NT}^* = o_{P^*}(1). \tag{A.5}$$

As in the proof of Proposition A.8,

$$\begin{aligned}
E^*(\mathcal{Z}^*) =\ & \frac{16}{T^4N^2h^2}\sum_{s=2}^T E^*\Bigg[\sum_{r=1}^{s-1}\bar{k}_{sr}^4\phi_{sr}^{*4} + 2\sum_{1\leq r_1<r_2\leq s-1}\bar{k}_{sr_1}^2\bar{k}_{sr_2}^2\phi_{sr_1}^{*2}\phi_{sr_2}^{*2} \\
&+4\sum_{t=1}^{s-1}\sum_{1\leq r_1<r_2\leq s-1}\bar{k}_{st}^2\bar{k}_{sr_1}\bar{k}_{sr_2}\phi_{st}^{*2}\phi_{sr_1}^*\phi_{sr_2}^* + 4\sum_{1\leq r_1<r_2\leq s-1, 1\leq t_1<t_2\leq s-1}\bar{k}_{r_1s}\bar{k}_{r_2s}\bar{k}_{t_1s}\bar{k}_{t_2s}\phi_{sr_1}^*\phi_{sr_2}^*\phi_{st_1}^*\phi_{st_2}^*\Bigg] \\
\equiv\ & \mathcal{Z}_1^* + \mathcal{Z}_2^* + \mathcal{Z}_3^* + \mathcal{Z}_4^*,
\end{aligned}$$

where $\phi_{sr}^{*}=\tilde{F}_s'\mathbb{S}^{*}\tilde{F}_r e_s^{*\prime}e_r^{*}$. Using the *i.i.d.* property of ς_{it} and the conditions in Theorem 3.3, we can readily verify that $\mathcal{Z}_l^{*}=o_P(1)$ for $l=1,2,3,4$. For example, noting that $E\left[\varsigma_{i_1s}\varsigma_{i_2s}\varsigma_{i_3s}\varsigma_{i_4s}\right]=3$ if $i_1=i_2=i_3=i_4$, $=1$ if $i_1=i_2\neq i_3=i_4$, $i_1=i_3\neq i_2=i_4$, or $i_1=i_4\neq i_2=i_4$, and zero otherwise, we have for any $s\neq r$,

$$
\begin{aligned}
E^{*}[(e_s^{*\prime}e_r^{*})^4] &= E^{*}[(\varsigma_s'\tilde{\Sigma}\varsigma_r)^4]=\sum_{i_1,\ldots,i_4,j_1,\ldots,j_4}\tilde{\sigma}_{i_1j_1}\tilde{\sigma}_{i_2j_2}\tilde{\sigma}_{i_3j_3}\tilde{\sigma}_{i_4j_4}E\left[\varsigma_{i_1s}\varsigma_{i_2s}\varsigma_{i_3s}\varsigma_{i_4s}\right]E\left[\varsigma_{j_1r}\varsigma_{j_2r}\varsigma_{j_3r}\varsigma_{j_4r}\right]\\
&= 9\sum_{i,j}\tilde{\sigma}_{ij}^4+9\sum_i\sum_{j_1\neq j_2}\tilde{\sigma}_{ij_1}^2\tilde{\sigma}_{ij_2}^2+9\sum_{i_1\neq i_2}\sum_j\tilde{\sigma}_{i_1j}^2\tilde{\sigma}_{i_2j}^2\\
&\quad+\sum_{i_1\neq i_2}\sum_{j_1\neq j_2}[\tilde{\sigma}_{i_1j_1}\tilde{\sigma}_{i_1j_1}\tilde{\sigma}_{i_2j_2}\tilde{\sigma}_{i_2j_2}+\tilde{\sigma}_{i_1j_1}\tilde{\sigma}_{i_1j_2}\tilde{\sigma}_{i_2j_1}\tilde{\sigma}_{i_2j_2}+\tilde{\sigma}_{i_1j_1}\tilde{\sigma}_{i_1j_2}\tilde{\sigma}_{i_2j_2}\tilde{\sigma}_{i_2j_1}\\
&\quad+\tilde{\sigma}_{i_1j_1}\tilde{\sigma}_{i_2j_1}\tilde{\sigma}_{i_1j_2}\tilde{\sigma}_{i_2j_2}+\tilde{\sigma}_{i_1j_1}\tilde{\sigma}_{i_2j_2}\tilde{\sigma}_{i_1j_1}\tilde{\sigma}_{i_2j_2}+\tilde{\sigma}_{i_1j_1}\tilde{\sigma}_{i_2j_2}\tilde{\sigma}_{i_1j_2}\tilde{\sigma}_{i_2j_1}\\
&\quad+\tilde{\sigma}_{i_1j_1}\tilde{\sigma}_{i_2j_1}\tilde{\sigma}_{i_2j_2}\tilde{\sigma}_{i_1j_2}+\tilde{\sigma}_{i_1j_1}\tilde{\sigma}_{i_2j_2}\tilde{\sigma}_{i_2j_1}\tilde{\sigma}_{i_1j_2}+\tilde{\sigma}_{i_1j_1}\tilde{\sigma}_{i_2j_2}\tilde{\sigma}_{i_2j_2}\tilde{\sigma}_{i_1j_1}]\\
&= 9\sum_{i,j}\tilde{\sigma}_{ij}^4+18\sum_i\sum_{j_1\neq j_2}\tilde{\sigma}_{ij_1}^2\tilde{\sigma}_{ij_2}^2+3\sum_{i_1\neq i_2}\sum_{j_1\neq j_2}[\tilde{\sigma}_{i_1j_1}^2\tilde{\sigma}_{i_2j_2}^2+2\tilde{\sigma}_{i_1j_1}\tilde{\sigma}_{i_1j_2}\tilde{\sigma}_{i_2j_1}\tilde{\sigma}_{i_2j_2}]\\
&= O_P\left(\xi_{NT}^3N+N\xi_{NT}^2+N^2\xi_{NT}^2\right)=O_P\left(N^2\xi_{NT}^2\right).
\end{aligned}
$$

Then

$$
\begin{aligned}
\mathcal{Z}_1^{*} &= \frac{16}{T^4N^2h^2}\sum_{s=2}^{T}\sum_{r=1}^{s-1}\bar{k}_{sr}^4\left(\tilde{F}_s'\mathbb{S}^{*}\tilde{F}_r\right)^4E^{*}\left(e_s^{*\prime}e_r^{*}\right)^4\\
&= \frac{16}{T^4N^2h^2}\sum_{s=2}^{T}\sum_{r=1}^{s-1}\bar{k}_{sr}^4\left(\tilde{F}_s'\mathbb{S}^{*}\tilde{F}_r\right)^4O_P\left(N^2\xi_{NT}^2\right)=O_P\left(T^{-2}h^{-2}\xi_{NT}^2\right),
\end{aligned}
$$

where we use the fact that $\frac{1}{T^2h}\sum_{s=2}^{T}\sum_{r=1}^{s-1}\bar{k}_{sr}^4||\tilde{F}_s||^8=O_P(1)$ under Assumption A.3 and the extra conditions in the theorem. Similarly, noting that for any $r_1<r_2<s$,

$$
\begin{aligned}
E^{*}\left[\left(e_s^{*\prime}e_{r_1}^{*}\right)^2\left(e_s^{*\prime}e_{r_2}^{*}\right)^2\right] &= E^{*}\left[(\varsigma_s'\tilde{\Sigma}\varsigma_{r_1}\varsigma_{r_1}'\tilde{\Sigma}\varsigma_s)(\varsigma_s'\tilde{\Sigma}\varsigma_{r_2}\varsigma_{r_1}'\tilde{\Sigma}\varsigma_s)\right]=E^{*}[\varsigma_s'\tilde{\Sigma}\tilde{\Sigma}\varsigma_s\varsigma_s'\tilde{\Sigma}\tilde{\Sigma}\varsigma_s]\\
&= \sum_{i_1,\ldots,i_4,j_1,j_2}\tilde{\sigma}_{i_1j_1}\tilde{\sigma}_{j_1i_2}\tilde{\sigma}_{i_3j_2}\tilde{\sigma}_{j_2i_4}E\left[\varsigma_{i_1s}\varsigma_{i_2s}\varsigma_{i_3s}\varsigma_{i_4s}\right]\\
&= 3\sum_{i,j_1,j_2}\tilde{\sigma}_{ij_1}^2\tilde{\sigma}_{ij_2}^2+\sum_{i_1,i_2,j_1,j_2}[\tilde{\sigma}_{i_1j_1}^2\tilde{\sigma}_{i_2j_2}^2+2\tilde{\sigma}_{i_1j_1}\tilde{\sigma}_{j_1i_2}\tilde{\sigma}_{i_1j_2}\tilde{\sigma}_{j_2i_2}]\\
&= O_P\left(N\xi_{NT}^3\right)+O_P\left(N^2\xi_{NT}^2\right)=O_P\left(N^2\xi_{NT}^2\right),
\end{aligned}
$$

where we use the fact that $\tilde{\sigma}_{ij}=\tilde{\sigma}_{ji}$ and $\xi_{NT}=o\left(T^{1/2}\right)=o(N)$, we have

$$
\begin{aligned}
\mathcal{Z}_4^{*} &= \frac{64}{T^4N^2h^2}\sum_{s=2}^{T}\sum_{1\leq r_1<r_2\leq s-1,1\leq t_1<t_2\leq s-1}\bar{k}_{r_1s}\bar{k}_{r_2s}\bar{k}_{t_1s}\bar{k}_{t_2s}\tilde{F}_s'\mathbb{S}^{*}\tilde{F}_{r_1}\tilde{F}_s'\mathbb{S}^{*}\tilde{F}_{r_2}\tilde{F}_s'\mathbb{S}^{*}\tilde{F}_{t_1}\tilde{F}_s'\mathbb{S}^{*}\tilde{F}_{t_2}\\
&\quad\times E^{*}\left[\left(e_s^{*\prime}e_{r_1}^{*}\right)\left(e_s^{*\prime}e_{r_2}^{*}\right)\left(e_s^{*\prime}e_{t_1}^{*}\right)\left(e_s^{*\prime}e_{t_2}^{*}\right)\right]\\
&= \frac{64}{T^4N^2h^2}\sum_{s=2}^{T}\sum_{1\leq r_1<r_2\leq s-1}\bar{k}_{r_1s}^2\bar{k}_{r_2s}^2\left(\tilde{F}_s'\mathbb{S}^{*}\tilde{F}_{r_1}\right)^2\left(\tilde{F}_s'\mathbb{S}^{*}\tilde{F}_{r_2}\right)^2O_P\left(N^2\xi_{NT}^2\right)=O_P\left(T^{-2}h^{-2}\xi_{NT}^2\right),
\end{aligned}
$$

Then $\mathcal{Z}^{*}=o_{P^{*}}(1)$ by the conditional Markov inequality. Now $\sum_{t=2}^{T}E^{*}(Z_{NT,t}^{*2})=4T^{-2}N^{-1}h^{-1}E^{*}[\sum_{r=1}^{s-1}\bar{k}_{sr}\times\tilde{F}_s'\mathbb{S}^{*}\tilde{F}_re_s^{*\prime}e_r^{*}]^2=\mathbb{V}_{NT}^{*}$. Straightforward moment calculations yield that $E^{*}(\sum_{t=2}^{T}Z_{NT,t}^{*2})^2=\mathbb{V}_{NT}^{*2}+o_P(1)$. Thus $\mathrm{Var}^{*}(\sum_{t=2}^{T}Z_{NT,t}^{*2})=o_P(1)$ and $\sum_{t=2}^{T}Z_{NT,t}^{*2}-\mathbb{V}_{NT}^{*}=o_P(1)$. This completes the proof of (i). ■

A.3 Proofs of the Technical Lemmas

Recall that $\max_i$, $\max_t$, and $\max_{s,t}$ denote $\max_{1\leq i\leq N}$, $\max_{1\leq t\leq T}$, and $\max_{1\leq s,t\leq T}$, respectively. Let $\|A\|_q = \{E\|A\|^q\}^{1/q}$ for $q \geq 1$.

Proof of Lemma A.1. (i) From the principal component analysis, we have the identity $(NT)^{-1} XX'\tilde{F} = \tilde{F}V_{NT}$. Pre-multiplying both sides by $T^{-1}\tilde{F}'$ and using the normalization $T^{-1}\tilde{F}'\tilde{F} = \mathbb{I}_R$ yields $T^{-1}\tilde{F}'(NT)^{-1} XX'\tilde{F} = V_{NT}$. By Bai (2003, Lemma A.3) and following the proof of (ii) below, V_{NT} has probability limit V_0 that is a diagonal matrix consisting of the R eigenvalues of $\Sigma_{\Lambda_0}\Sigma_F$ under Assumptions A.1-A.3 and A.5.

(ii) Noting that $X = F\Lambda_0' + e^\dagger$, where $e^\dagger = e + a_{NT}g^\dagger$, $g^\dagger = (g_1^\dagger, ..., g_T^\dagger)'$, $g_t^\dagger = (F_t'g_{1t}, ..., F_t'g_{Nt})'$ and $g_{it} = g_i(t/T)$, (i) implies that

$$(T^{-1}\tilde{F}'F)\left(N^{-1}\Lambda_0'\Lambda_0\right)(T^{-1}F'\tilde{F}) + d_{NT} = V_{NT} \xrightarrow{P} V_0, \tag{A.6}$$

where $d_{NT} = N^{-1}T^{-2}\tilde{F}'e^\dagger e^{\dagger\prime}\tilde{F} + (T^{-1}\tilde{F}'F)(N^{-1}T^{-1}\Lambda_0'e^{\dagger\prime}\tilde{F}) + (N^{-1}T^{-1}\tilde{F}'e^\dagger\Lambda_0)(T^{-1}F'\tilde{F})$. Noting that

$$\begin{aligned}
N^{-1}T^{-2}||\tilde{F}'e^{\dagger\prime}e^\dagger\tilde{F}|| &\leq 2N^{-1}T^{-1}\{T^{-1}||\tilde{F}||^2\}\left(R\|e\|_{\text{sp}}^2 + a_{NT}^2\left\|g^\dagger\right\|^2\right) \\
&= O_P(T^{-1} + N^{-1} + a_{NT}^2), \\
N^{-1}T^{-1}||\Lambda_0'e^{\dagger\prime}\tilde{F}|| &\leq N^{-1}T^{-1/2}\{T^{-1/2}||\tilde{F}||\}\left(\|e\Lambda_0\| + a_{NT}\|\bar{g}\Lambda_0\|\right) \\
&= N^{-1}T^{-1/2}O_P(1)\left(N^{1/2}T^{1/2} + a_{NT}NT^{1/2}\right) = O_P(N^{-1/2} + a_{NT}),
\end{aligned}$$

and $T^{-1}||F'\tilde{F}|| = O_P(1)$ under Assumptions A.1-A.3 and A.5, we have

$$\|d_{NT}\| = O_P(N^{-1} + T^{-1} + a_{NT}^2 + N^{-1/2} + a_{NT}) = o_P(1). \tag{A.7}$$

It follows that $(\tilde{F}'F/T)(\Lambda_0'\Lambda_0/N)(F'\tilde{F}/T) \xrightarrow{P} V_0$. ■

Proof of Lemma A.2. (i) Let $e_t^\dagger = (e_{1t}^\dagger, ..., e_{Nt}^\dagger)'$ and $\Lambda_0 = (\lambda_{10}, ..., \lambda_{N0})'$. Noting that $(NT)^{-1}XX'\tilde{F} = \tilde{F}V_{NT}$ and $X_{it} = \lambda_{it}'F_t + e_{it} = \lambda_{i0}'F_t + e_{it}^\dagger$ with $e_{it}^\dagger = e_{it} + a_{NT}F_t'g_{it}$, we can decompose $\tilde{F}_t - H'F_t$ as follows:

$$\begin{aligned}
\tilde{F}_t - H'F_t &= V_{NT}^{-1}\frac{1}{NT}\sum_{s=1}^T \tilde{F}_sX_s'X_t - H'F_t \\
&= V_{NT}^{-1}\frac{1}{NT}\sum_{s=1}^T \tilde{F}_s\left[\Lambda_0F_s + e_s^\dagger\right]'\left[\Lambda_0F_t + e_t^\dagger\right] - H'F_t \\
&= V_{NT}^{-1}\left\{\frac{1}{T}\sum_{s=1}^T \tilde{F}_sE(e_s^{\dagger\prime}e_t^\dagger/N) + \frac{1}{T}\sum_{s=1}^T \tilde{F}_s\left[e_s^{\dagger\prime}e_t^\dagger/N - E(e_s^{\dagger\prime}e_t^\dagger/N)\right]\right. \\
&\quad \left. + \frac{1}{T}\sum_{s=1}^T \tilde{F}_sF_s'\Lambda_0'e_t^\dagger/N + \frac{1}{T}\sum_{s=1}^T \tilde{F}_sF_t'\Lambda_0'e_s^\dagger/N\right\} \\
&\equiv A_1(t) + A_2(t) + A_3(t) + A_4(t), \quad \text{say.}
\end{aligned} \tag{A.8}$$

By (A.8) and the inequality $(a+b+c+d)^2 \leq 4(a^2+b^2+c^2+d^2)$, we have

$$\frac{1}{T}\sum_{t=1}^T\left\|\tilde{F}_t - H'F_t\right\|^2 \leq \frac{4}{T}\left\|V_{NT}^{-1}\right\|^2\sum_{t=1}^T\left[\|V_{NT}A_1(t)\|^2 + \|V_{NT}A_2(t)\|^2 + \|V_{NT}A_3(t)\|^2 + \|V_{NT}A_4(t)\|^2\right].$$

By Lemma A.1(i), it suffices to bound $\frac{1}{T}\sum_{t=1}^T\|V_{NT}A_l(t)\|^2$ for $l = 1,2,3,4$. Let $g_t^\dagger \equiv (F_t'g_{1t}, ..., F_t'g_{Nt})'$. Using $e_s^{\dagger\prime}e_t^\dagger = (e_s + a_{NT}g_s^\dagger)'(e_t + a_{NT}g_t^\dagger) = e_s'e_t + a_{NT}e_s'g_t^\dagger + a_{NT}g_s^{\dagger\prime}e_t + a_{NT}^2g_s^{\dagger\prime}g_t^\dagger$ and Cauchy-Schwarz (CS hereafter) inequality, we have

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}\left[E(e_s^{\dagger\prime}e_t^{\dagger}/N)\right]^2 \leq \frac{4}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}\left[E\left(e_t'e_s/N\right)\right]^2+\frac{8a_{NT}^2}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}\left[E(e_s'g_t^{\dagger}/N)\right]^2+\frac{4a_{NT}^4}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}\left[E(g_s^{\dagger\prime}g_t^{\dagger}/N)\right]^2.$$

As in Bai (2003), the first term is bounded above by $4\max_{s,t}\gamma_N(s,t)\max_t\sum_{t=1}^{T}|\gamma_N(s,t)| = O(1)$ by Assumption A.3(iv). By Davydov inequality and Assumptions A.1(ii), A.3(i) and (iii) and A.5(i)

$$\begin{aligned}
\frac{a_{NT}^2}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}\left[E(e_s'g_t^{\dagger}/N)\right]^2 &= \frac{a_{NT}^2}{TN^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{i=1}^{N}\sum_{j=1}^{N}E\left(e_{is}F_t'g_{it}\right)E\left(e_{js}F_t'g_{jt}\right)\\
&\leq C\max_{i,s}E\left(e_{is}F_t'g_{it}\right)\|e_{is}\|_{2+\delta}\|F_t\|_{2+\delta}\frac{8a_{NT}^2}{TN^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_j\left(|t-s|\right)^{\delta/(2+\delta)}\\
&\leq C\max_{i,s}E\left(e_{is}F_t'g_{it}\right)\|e_{is}\|_{2+\delta}\|F_t\|_{2+\delta}\frac{8a_{NT}^2}{N}\sum_{j=1}^{N}\sum_{s=1}^{\infty}\alpha_j\left(s\right)^{\delta/(2+\delta)} = O(a_{NT}^2).
\end{aligned}$$

In addition, we can show that $\frac{a_{NT}^4}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}\left[E(g_s^{\dagger\prime}g_t^{\dagger}/N)\right]^2 = O(a_{NT}^4T) = o(1)$ under Assumptions A.1(ii) and A.5(i). It follows that $\frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}\left[E(e_s^{\dagger\prime}e_t^{\dagger}/N)\right]^2 \leq \frac{4}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}\left[E\left(e_t'e_s/N\right)\right]^2+o(1) = O(1)$ under Assumption A.3(i)-(ii). Then by the submultiplicative property of the Frobenius norm, CS inequality, and the fact that $T^{-1}\tilde{F}'\tilde{F} = \mathbb{I}_R$, we have

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T}\|V_{NT}A_1(t)\|^2 &= \frac{1}{T}\sum_{t=1}^{T}\left\|\frac{1}{T}\sum_{s=1}^{T}\tilde{F}_sE(e_s^{\dagger\prime}e_t^{\dagger}/N)\right\|^2 \leq \frac{1}{T}\sum_{t=1}^{T}\left\{\frac{1}{T}\sum_{s=1}^{T}\left\|\tilde{F}_sE(e_s^{\dagger\prime}e_t^{\dagger}/N)\right\|\right\}^2\\
&\leq \frac{1}{T}\sum_{r=1}^{T}\left\|\tilde{F}_r\right\|^2\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\left[E\left(e_s^{\dagger\prime}e_t^{\dagger}/N\right)\right]^2 = O_P(1)O(T^{-1}) = O_P(T^{-1}).
\end{aligned}$$

Now, we consider the second term. Recall that $\xi_{st} = e_s'e_t/N - E\left(e_s'e_t/N\right)$. Let $\xi_{st}^{\dagger} = e_s^{\dagger\prime}e_t^{\dagger}/N - E(e_s^{\dagger\prime}e_t^{\dagger}/N)$. Then

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T}\|V_{NT}A_2(t)\|^2 &= \frac{1}{T}\sum_{t=1}^{T}\left\|\frac{1}{T}\sum_{s=1}^{T}\tilde{F}_s\xi_{st}^{\dagger}\right\|^2 = \frac{1}{T^3}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{r=1}^{T}\tilde{F}_s'\tilde{F}_r\xi_{st}^{\dagger}\xi_{rt}^{\dagger}\\
&\leq \frac{1}{T}\left[\frac{1}{T^2}\sum_{s=1}^{T}\sum_{r=1}^{T}(\tilde{F}_s'\tilde{F}_r)^2\right]^{1/2}\left[\frac{1}{T^2}\sum_{s=1}^{T}\sum_{l=1}^{T}\left(\sum_{t=1}^{T}\xi_{st}^{\dagger}\xi_{lt}^{\dagger}\right)^2\right]^{1/2}\\
&= \frac{1}{T}\left[\frac{1}{T}\sum_{s=1}^{T}\left\|\tilde{F}_s\right\|^2\right]\left[\frac{1}{T^2}\sum_{s=1}^{T}\sum_{l=1}^{T}\left(\sum_{t=1}^{T}\xi_{st}^{\dagger}\xi_{rt}^{\dagger}\right)^2\right]^{1/2}.
\end{aligned}$$

In addition, using $\xi_{st}^{\dagger} = \xi_{st}+a_{NT}N^{-1}[e_s'g_t^{\dagger}-E(e_s'g_t^{\dagger})]+a_{NT}N^{-1}[g_s^{\dagger\prime}e_t-E\left(g_s^{\dagger\prime}e_t\right)]+a_{NT}^2N^{-1}[g_s^{\dagger\prime}g_t^{\dagger}-E(g_s^{\dagger\prime}g_t^{\dagger})]$, we can readily show that $\frac{1}{T^2}\sum_{s=1}^{T}\sum_{l=1}^{T}E(\sum_{t=1}^{T}\xi_{st}^{\dagger}\xi_{rt}^{\dagger})^2 = O(T^2N^{-2})$ under Assumptions A.3 and A.5. It follows that $\frac{1}{T}\sum_{t=1}^{T}\|V_{NT}A_2(t)\|^2 = \frac{1}{T}O_P\left(T/N\right) = O_P(N^{-1})$. For the third term, noting that $E(e_{it}g_{jt}^{\dagger}) = 0$, we have by Assumptions A.1(i)-(ii), A.2(i), A.3(vi) and A.5(i)

$$
\begin{aligned}
(NT)^{-1}\sum_{t=1}^{T}E\left\|\Lambda_0'e_t^{\dagger}\right\|^2 &= (NT)^{-1}\sum_{t=1}^{T}\sum_{i=1}^{N}\sum_{j=1}^{N}\lambda_{i0}'\lambda_{j0}E(e_{it}^{\dagger}e_{jt}^{\dagger}) \\
&= (NT)^{-1}\sum_{t=1}^{T}\sum_{i=1}^{N}\sum_{j=1}^{N}\lambda_{i0}'\lambda_{j0}E\left(e_{it}e_{jt}\right)+a_{NT}^2(NT)^{-1}\sum_{t=1}^{T}\sum_{i=1}^{N}\sum_{j=1}^{N}\lambda_{i0}'\lambda_{j0}E(g_{it}^{\dagger}g_{jt}^{\dagger}) \\
&\leq C(NT)^{-1}\sum_{t=1}^{T}\sum_{i=1}^{N}\sum_{j=1}^{N}\left|E\left(e_{it}e_{jt}\right)\right|+O(T^{-1}N^{1/2}h^{-1/2})=O\left(1\right)
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T}\left\|V_{NT}A_3(t)\right\|^2 &= \frac{1}{T}\sum_{t=1}^{T}\left\|\frac{1}{T}\sum_{s=1}^{T}\tilde{F}_sF_s'\Lambda_0'e_t^{\dagger}/N\right\|^2\leq\frac{1}{N}\frac{1}{T}\sum_{t=1}^{T}\left\|\frac{\Lambda_0'e_t^{\dagger}}{\sqrt{N}}\right\|^2\left[\frac{1}{T}\sum_{s=1}^{T}\left\|\tilde{F}_s\right\|^2\right]\left[\frac{1}{T}\sum_{s=1}^{T}\left\|F_s\right\|^2\right] \\
&= O_P(N^{-1}).
\end{aligned}
$$

For the fourth term, we have

$$
\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T}\left\|V_{NT}A_4(t)\right\|^2 &= \frac{1}{T}\sum_{t=1}^{T}\left\|\frac{1}{T}\sum_{s=1}^{T}\tilde{F}_sF_t'\Lambda_0'e_s^{\dagger}/N\right\|^2\leq\frac{1}{N^2T^3}\sum_{t=1}^{T}\left\{\sum_{s=1}^{T}\left\|\tilde{F}_sF_t'\right\|\left\|\Lambda_0'e_s^{\dagger}\right\|\right\}^2 \\
&\leq\frac{1}{N}\left\{\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\left\|\tilde{F}_sF_t'\right\|^2\right\}\frac{1}{NT}\sum_{s=1}^{T}\left\|\Lambda_0'e_s^{\dagger}\right\|^2=N^{-1}O_P\left(1\right)O_P\left(1\right)=O_P\left(N^{-1}\right),
\end{aligned}
$$

as we have shown that $\frac{1}{NT}\sum_{s=1}^{T}E\left\|\Lambda_0'e_s^{\dagger}\right\|^2=O(1+N^{1/2}T^{-1}h^{-1/2})=O\left(1\right).$ Combining these results, we have $\frac{1}{T}\sum_{t=1}^{T}\left\|\tilde{F}_t-H'F_t\right\|^2=O_P(C_{NT}^{-2}).$

(ii) By (A.8), we have $\frac{1}{T}(\tilde{F}-FH)'FH=\frac{1}{T}\sum_{t=1}^{T}[A_1(t)+A_2(t)+A_3(t)+A_4(t)]F_t'H.$ We first decompose $V_{NT}\frac{1}{T}\sum_{t=1}^{T}\left[A_1(t)+A_2\left(t\right)\right]F_t'H$ as follows

$$
\begin{aligned}
V_{NT}\frac{1}{T}\sum_{t=1}^{T}\left[A_1(t)+A_2\left(t\right)\right]F_t'H &= \frac{1}{NT^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\tilde{F}_se_s^{\dagger\prime}e_t^{\dagger}F_t'H \\
&= \frac{1}{NT^2}\sum_{t=1}^{T}\sum_{s=1}^{T}(\tilde{F}_s-H'F_s)e_s^{\dagger\prime}e_t^{\dagger}F_t'H+\frac{1}{NT^2}H'\sum_{t=1}^{T}\sum_{s=1}^{T}F_se_s^{\dagger\prime}e_t^{\dagger}F_t'H \\
&\equiv \bar{A}_1+\bar{A}_2,\ \text{say.}
\end{aligned}
$$

For $\bar{A}_1$, we apply CS inequality and the result in part (i) to obtain

$$
\begin{aligned}
\left\|\bar{A}_1\right\| &\leq \left\|H\right\|\left\{\frac{1}{T}\sum_{s=1}^{T}\left\|\tilde{F}_s-H'F_s\right\|^2\right\}^{1/2}\left\{\frac{1}{N^2T^3}\sum_{s=1}^{T}\left\|\sum_{t=1}^{T}e_s^{\dagger\prime}e_t^{\dagger}F_t'\right\|^2\right\}^{1/2} \\
&= O_P\left(C_{NT}^{-1}\right)O_P(N^{-1/2})=O_P(C_{NT}^{-1}N^{-1/2}),
\end{aligned}
$$

provided that $\frac{1}{N^2T^3}\sum_{s=1}^T \left\|\sum_{t=1}^T e_s^{\dagger\prime} e_t^\dagger F_t^\prime\right\|^2 = O_P\left(N^{-1}\right)$. To see why the last claim is true, note that

$$\begin{aligned}
\frac{1}{N^2T^3}\sum_{s=1}^T E\left\|\sum_{t=1}^T e_s^{\dagger\prime} e_t^\dagger F_t^\prime\right\|^2 &\leq \frac{4}{N^2T^3}\sum_{s=1}^T E\left\|\sum_{t=1}^T e_s^\prime e_t F_t^\prime\right\|^2 + \frac{4a_{NT}}{N^2T^3}\sum_{s=1}^T E\left\|\sum_{t=1}^T e_s^\prime g_t^\dagger F_t^\prime\right\|^2 \\
&\quad + \frac{4a_{NT}}{N^2T^3}\sum_{s=1}^T E\left\|\sum_{t=1}^T g_s^{\dagger\prime} e_t F_t^\prime\right\|^2 + \frac{4a_{NT}^2}{N^2T^3}\sum_{s=1}^T E\left\|\sum_{t=1}^T g_s^{\dagger\prime} g_t^\dagger F_t^\prime\right\|^2 \\
&\equiv 4A_{1,1} + 4A_{1,2} + 4A_{1,3} + 4A_{1,4}.
\end{aligned}$$

For $A_{1,1}$, we have by Assumptions A.1(ii) and A.3(iv),

$$\begin{aligned}
A_{1,1} &\leq \frac{2}{N^2T^3}\sum_{s=1}^T E\left\|\sum_{t=1}^T E(e_s^\prime e_t)F_t^\prime\right\|^2 + \frac{2}{T^2}\sum_{s=1}^T E\| \left\|\sum_{t=1}^T \xi_{st}F_t^\prime\right\|^2 \\
&= \frac{2}{N^2T^3}\sum_{s=1}^T\sum_{t=1}^T\sum_{l=1}^T E(e_s^\prime e_t)E(e_s^\prime e_l)E(F_t^\prime F_l) + \frac{2}{T^3}\sum_{s=1}^T\sum_{t=1}^T\sum_{l=1}^T E(\xi_{st}\xi_{sl}F_t^\prime F_l) \\
&\leq 2T^{-2}\left(\max_s \sum_{t=1}^T \gamma_N(s,t)\right)^2 \max_t \|F_t\|_2^2 + 2N^{-1}\max_{s,t}\|N^{1/2}\xi_{st}\|^2 \max_t \|F_t\|_4^2 \\
&= O(T^{-2} + N^{-1}).
\end{aligned}$$

For $A_{1,2}$, we have under Assumptions A.2-A.3 and A.5(i),

$$\begin{aligned}
A_{1,2} &= \frac{a_{NT}}{N^2T^3}\sum_{s=1}^T E\left\|e_s^\prime \sum_{t=1}^T g_t^\dagger F_t^\prime\right\|^2 = \frac{a_{NT}}{N^2T^3}E\left[\sum_{s=1}^T e_s^\prime e_s \left\|\sum_{t=1}^T g_t^\dagger F_t^\prime\right\|^2\right] \\
&\leq a_{NT}N^{-1}E\left\|\frac{1}{NT}\sum_{s=1}^T e_s^\prime e_s\right\|_2 \left\{\frac{1}{T^4}E\left[\left\|\sum_{t=1}^T g_t^\dagger F_t^\prime\right\|^4\right]\right\}^{1/2} = O\left(a_{NT}N^{-1}\right).
\end{aligned}$$

Similarly, we can show that $A_{1,3} = O\left(a_{NT}(N^{-1} + T^{-1})\right)$ and $A_{1,4} = O\left(a_{NT}^2\right)$. As a result, $\frac{1}{N^2T^3}\sum_{s=1}^T E\left\|\sum_{t=1}^T e_s^{\dagger\prime} e_t^\dagger F_t^\prime\right\|^2 = O\left(N^{-1}\right)$ and $\bar{A}_1 = O_P\left(C_{NT}^{-1}N^{-1/2}\right)$. Now, let $\check{A}_2 = \frac{1}{NT^2}\sum_{t=1}^T\sum_{s=1}^T F_s e_s^{\dagger\prime} e_t^\dagger F_t^\prime$. Let $a_{2,mn}$ denote the (m,n)th element of $\check{A}_2$ for $m,n = 1,...,R$. By CS inequality, it is easy to see that $|a_{2,mn}| \leq \{a_{2,mm}a_{2,nn}\}^{1/2}$. This, in conjunction with the Markov inequality, implies that it suffices to show that $\check{A}_2 = O_P\left(T^{-1}\right)$ by showing that $E\left|a_{2,mm}\right| = O\left(T^{-1}\right)$. In fact, by CS inequality, Assumptions A.1(ii), A.3(v) and A.5(i), we can readily show that

$$\begin{aligned}
E\left|a_{2,mm}\right| &= E\left(a_{2,mm}\right) = \frac{1}{NT^2}\sum_{t=1}^T\sum_{s=1}^T \iota_m^\prime E\left(F_s e_s^{\dagger\prime} e_t^\dagger F_t^\prime\right)\iota_m \\
&\leq \frac{2}{NT^2}\sum_{t=1}^T\sum_{s=1}^T \iota_m^\prime E\left(F_s e_s^\prime e_t F_t^\prime\right)\iota_m + \frac{2a_{NT}^2}{NT^2}\sum_{t=1}^T\sum_{s=1}^T \iota_m^\prime E\left(F_s g_s^{\dagger\prime} g_t^\dagger F_t^\prime\right)\iota_m \\
&\leq \frac{2}{T}\max_s\left(\sum_{t=1}^T \gamma_{N,FF}(s,t)\right) + \frac{2a_{NT}^2}{NT^2}\sum_{t=1}^T\sum_{s=1}^T \iota_m^\prime E\left(F_s g_s^{\dagger\prime} g_t^\dagger F_t^\prime\right)\iota_m \\
&= O\left(T^{-1}\right) + O\left(a_{NT}^2\right) = O\left(T^{-1}\right),
\end{aligned}$$

where ι_m is the mth column of R-dimensional identity matrix $\mathbb{I}_R$. It follows that $\bar{A}_2 = O_P\left(T^{-1}\right)$ and $\frac{1}{T}\sum_{t=1}^{T} V_{NT}\left[A_1(t) + A_2(t)\right] F_t' H = O_P\left(C_{NT}^{-2}\right).$

Now, we consider $V_{NT}\frac{1}{T}\sum_{t=1}^{T} A_3(t) F_t' H$. Note that

$$\begin{aligned}\left\|V_{NT}\frac{1}{T}\sum_{t=1}^{T} A_3(t) F_t' H\right\| &= \frac{1}{NT^2}\left\|\sum_{s=1}^{T}\tilde{F}_s F_s' \sum_{t=1}^{T}\Lambda_0' e_t^{\dagger} F_t' H\right\| \le \|H\| \left\|\frac{1}{T}\sum_{s=1}^{T}\tilde{F}_s F_s'\right\| \left\|\frac{1}{NT}\sum_{t=1}^{T}\Lambda_0' e_t^{\dagger} F_t'\right\| \\ &= o\left(a_{NT}\right)\end{aligned}$$

provided $\left\|\frac{1}{NT}\sum_{t=1}^{T}\Lambda_0' e_t^{\dagger} F_t'\right\| = o\left(a_{NT}\right)$. To see this, we write

$$\begin{aligned}E\left\|\frac{1}{NT}\sum_{t=1}^{T}\Lambda_0' e_t^{\dagger} F_t'\right\|^2 &= E\left\|\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\lambda_{i0}\left(e_{it} + a_{NT} F_t' g_{it}\right) F_t'\right\|^2 \\ &\le \frac{2}{N^2T^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T} E\left(e_{it}e_{js}F_t'F_s\right)\lambda_{j0}'\lambda_{i0} \\ &\quad + \frac{2a_{NT}^2}{N^2T^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T} g_{js}' E\left(F_sF_s'F_tF_t'\right) g_{it}\lambda_{j0}'\lambda_{i0} \\ &\equiv 2A_{3,1} + 2A_{3,2}, \text{ say.}\end{aligned}$$

It is easy to show that $A_{3,1} = O\left(N^{-1}T^{-1}\right)$ under Assumptions A.2(i) and A.3(v). For $A_{3,2}$, using $E\left(F_sF_s'F_tF_t'\right) = E\left(F_sF_s'\right)E\left(F_tF_t'\right) + \mathrm{Cov}(F_sF_s', F_tF_t')$, we have

$$\begin{aligned}A_{3,2} &= \frac{a_{NT}^2}{N^2T^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T} g_{js}' E\left(F_sF_s'F_tF_t'\right) g_{it}\lambda_{j0}'\lambda_{i0} \\ &= \frac{a_{NT}^2}{N^2T^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T} g_{js}'\Sigma_F\Sigma_F g_{it}\lambda_{j0}'\lambda_{i0} + \frac{a_{NT}^2}{N^2T^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T} g_{js}'\mathrm{Cov}\left(F_sF_s', F_tF_t'\right) g_{it}\lambda_{j0}'\lambda_{i0} \\ &\equiv A_{3,2}^{(1)} + A_{3,2}^{(2)}.\end{aligned}$$

By local normalization $\int_0^1 g_i(u)du = 0$, $\lambda_{i0}'\frac{1}{T}\sum_{t=1}^{T} g_{it} = \lambda_{i0}'\frac{1}{T}\sum_{t=1}^{T} g_{NT,i}\left(t/T\right) = \lambda_{i0}'\int_0^1 g_{NT,i}\left(\tau\right)d\tau + O\left(\frac{1}{T}\right) = o\left(1\right)$ uniformly in i. Thus $A_{3,2}^{(1)} = o(a_{NT}^2)$. By Davydov inequality and Assumptions A.1(ii), A.2(i), A.3(iii), and A.5(i), we can readily show that $A_{3,2}^{(2)} = O\left(a_{NT}^2T^{-1}\right)$. It follows that $A_{3,2} = o\left(a_{NT}^2\right)$ and $V_{NT}\frac{1}{T}\sum_{t=1}^{T} A_3(t)F_t'H = o_P\left(a_{NT}\right).$

Now, we consider $V_{NT}\frac{1}{T}\sum_{t=1}^{T} A_4(t)F_t'H$:

$$\begin{aligned}V_{NT}\frac{1}{T}\sum_{t=1}^{T} A_4(t)F_t'H &= \frac{1}{NT^2}\sum_{t=1}^{T}\sum_{s=1}^{T}(\tilde{F}_s - H'F_s)F_t'\Lambda_0 e_s^{\dagger}F_t'H + \frac{1}{NT^2}H'\sum_{t=1}^{T}\sum_{s=1}^{T} F_sF_t'\Lambda_0 e_s^{\dagger}F_t'H \\ &\equiv A_{4,1} + A_{4,2}, \text{ say.}\end{aligned}$$

For $A_{4,1}$, we apply CS inequality and the result in part (i) to obtain

$$\begin{aligned}\left\|A_{4,1}\right\| &\le \|H\|\frac{1}{NT^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\left\|\tilde{F}_s - H'F_s\right\|\left\|F_t\right\|^2\left\|\Lambda_0 e_s^{\dagger}\right\| \\ &\le \|H\|\left\{\frac{1}{T}\sum_{t=1}^{T}\left\|F_t\right\|^2\right\}\left\{\frac{1}{T}\sum_{s=1}^{T}\left\|\tilde{F}_s - H'F_s\right\|^2\right\}^{1/2}\left\{\frac{1}{N^2T}\sum_{s=1}^{T}\left\|\Lambda_0' e_s^{\dagger}\right\|^2\right\}^{1/2} \\ &= O_P\left(1\right)O_P\left(C_{NT}^{-1}\right)O_P(N^{-1/2} + a_{NT}) = O_P\left(C_{NT}^{-2}\right)\end{aligned}$$

as $\frac{1}{N^2T}\sum_{s=1}^{T}E\left\|\Lambda_0' e_s^{\dagger}\right\|^2 = O\left(N^{-1}+a_{NT}^2\right)$ under Assumptions A.2, A.3 and A.5. Let $\bar{A}_{4,2} = \frac{1}{NT^2}\sum_{t=1}^{T}\sum_{s=1}^{T} F_s F_t' \Lambda_0 e_s^{\dagger} F_t'$. Let $a_{4,2,mn}$ denote the (m,n)th element of $\bar{A}_{4,2}$. Then

$$\begin{aligned} |a_{4,2,mn}| &= \left|\frac{1}{NT^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\iota_m' F_s F_t' \Lambda_0 e_s^{\dagger} F_t' \iota_n\right| = \left|\frac{1}{NT^2}\sum_{t=1}^{T}\left(F_t'\iota_n\right)F_t'\sum_{s=1}^{T}\Lambda_0' e_s^{\dagger}\iota_m' F_s\right| \\ &\leq \frac{1}{T}\sum_{t=1}^{T}\|F_t\|^2\frac{1}{NT}\left\|\sum_{s=1}^{T}\Lambda_0' e_s^{\dagger}\iota_m' F_s\right\| = o_P\left(a_{NT}\right), \end{aligned}$$

because

$$\begin{aligned} E\left\|\frac{1}{NT}\sum_{s=1}^{T}\Lambda_0' e_s^{\dagger}\iota_m' F_s\right\|^2 &\leq \frac{2}{N^2T^2}E\left\|\sum_{i=1}^{N}\sum_{t=1}^{T}\lambda_{ir}e_{it}\iota_m' F_t\right\|^2 + \frac{2a_{NT}^2}{N^2T^2}E\left\|\sum_{i=1}^{N}\sum_{t=1}^{T}\lambda_{ir}F_t' g_{it}\iota_m' F_t\right\|^2 \\ &= O(N^{-1}T^{-1}) + o(a_{NT}^2) = o(a_{NT}^2) \end{aligned}$$

by using arguments as used in the analysis of $A_{3,2}$. It follows that $\bar{A}_{4,2} = o_P\left(a_{NT}\right)$ and $V_{NT}\frac{1}{T}\sum_{t=1}^{T}A_4(t)F_t'H = o_P\left(a_{NT}\right)$. Combining the above results yields the claim in part (ii) of the lemma.

(iii) This follows from the results in (i) and (ii) and the triangle inequality.

(iv) Observing that $\frac{1}{T}(\tilde{F}'\tilde{F} - HF'FH) = \frac{1}{T}(\tilde{F}-FH)'(\tilde{F}-FH) + \frac{1}{T}(\tilde{F}-FH)'FH + \frac{1}{T}\left(FH\right)'\left(\tilde{F}-FH\right)$, the results follows from (i) and (ii).

(v) By (A.6) in the proof of Lemma A.1(ii),

$$(T^{-1}\tilde{F}'F)\left(N^{-1}\Lambda_0'\Lambda_0\right)(T^{-1}F'\tilde{F}) + d_{NT} = V_{NT}.$$

Premultiplying both sides by $\left(N^{-1}\Lambda_0'\Lambda_0\right)^{1/2}(T^{-1}F'\tilde{F})$ and using the fact that $T^{-1}\tilde{F}'\tilde{F} = \mathbb{I}_R$, we have

$$\left(N^{-1}\Lambda_0'\Lambda_0\right)^{1/2}(T^{-1}F'F)\left(N^{-1}\Lambda_0'\Lambda_0\right)(T^{-1}F'\tilde{F}) + \bar{d}_{NT} = \left(N^{-1}\Lambda_0'\Lambda_0\right)^{1/2}(T^{-1}F'\tilde{F})V_{NT}$$

where $\bar{d}_{NT} = (T^{-1}\tilde{F}'F)\left(N^{-1}\Lambda_0'\Lambda_0\right)d_{NT} = O_P(N^{-1/2})$. Let $D_{NT} = \left(N^{-1}\Lambda_0'\Lambda_0\right)^{1/2}(T^{-1}F'F)\left(N^{-1}\Lambda_0'\Lambda_0\right)^{1/2}$, $R_{NT} = \left(N^{-1}\Lambda_0'\Lambda_0\right)^{1/2}(T^{-1}F'\tilde{F})$, and $D_0 = \Sigma_{\Lambda_0}^{1/2}\Sigma_F\Sigma_{\Lambda_0}^{1/2}$. Then as in Bai (2003, p.161),

$$\left[D_{NT} + \bar{d}_{NT}R_{NT}\right]\Upsilon_{NT} = \Upsilon_{NT}V_{NT}$$

where $\Upsilon_{NT} = R_{NT}V_{NT}^{*-1/2}$ with V_{NT}^{*} being a diagonal matrix that contains the diagonal elements of $R_{NT}'R_{NT}$. That is, V_{NT} contains the eigenvalues of $D_{NT} + \bar{d}_{NT}R_{NT}$ with the corresponding normalized eigenvectors contained in Υ_{NT}. It is trivial to show that

$$\left\|D_{NT} + \bar{d}_{NT}R_{NT} - D_0\right\| = O_P\left(C_{NT}^{-1}\right). \tag{A.9}$$

By the perturbation theory for eigenvalue problem,

$$\left|\mu_j\left(D_{NT} + \bar{d}_{NT}R_{NT}\right) - \mu_j\left(D_0\right)\right| \leq \left\|D_{NT} + \bar{d}_{NT}R_{NT} - D_0\right\| = O_P\left(C_{NT}^{-1}\right),$$

where $\mu_j\left(A\right)$ denotes the jth largest eigenvalue of a symmetric matrix A. That is, $V_{NT} - V_0 = O_P\left(C_{NT}^{-1}\right)$.

(vi) Let Υ_0 denote the probability limit of Υ_{NT}. By (A.9) and the eigenvector perturbation theory that requires distinctness of eigenvalues (see, e.g., Steward and Sun (1990), $\|\Upsilon_{NT} - \Upsilon_0\| = O_P\left(C_{NT}^{-1}\right)$. [Let $\left(\phi_j, \mu_j\right)$ and $(\tilde{\phi}_j, \tilde{\mu}_j)$ be the eigenvector-eigenvalue pairs of a symmetric matrix A and its symmetric perturbation version $\tilde{A} = A + \Delta A$, respectively, where the eigenvectors are properly normalized. Then (i) $\tilde{\mu}_j = \mu_j + \phi_j'\Delta A\phi_j + o(\|\Delta A\|^2)$, and (ii) $\tilde{\phi}_j = \phi_j + \sum_{j\neq i}\left[\phi_j'\Delta A\phi_i/(\phi_j - \phi_i)\right]\phi_j + o(\|\Delta A\|^2)$ if $\mu_j \neq \mu_i$ for all $j \neq i$.] This, in conjunction with the definition of R_{NT}, implies that

$$T^{-1}F'\tilde{F} = (N^{-1}\Lambda_0'\Lambda_0)^{-1/2}\Upsilon_{NT}V_{NT}^{*1/2} = \Sigma_{\Lambda_0}^{-1/2}\Upsilon_0 V_0^{1/2} + O_P\left(C_{NT}^{-1}\right).$$

It follows that $H = (N^{-1}\Lambda_0'\Lambda_0)(T^{-1}F'\tilde{F})V_{NT}^{-1} = \Sigma_{\Lambda_0}\Sigma_{\Lambda_0}^{-1/2}\Upsilon_0 V_0^{1/2}V_0^{-1} + O_P\left(C_{NT}^{-1}\right) = \Sigma_{\Lambda_0}^{1/2}\Upsilon_0 V_0^{-1/2}$ $+O_P\left(C_{NT}^{-1}\right) = Q_0^{-1} + O_P\left(C_{NT}^{-1}\right)$, where $Q_0 = V_0^{1/2}\Upsilon_0^{-1}\Sigma_{\Lambda_0}^{-1/2}$. ■

Proof of Lemma A.3. (i) The proof parallels that of Lemma A.2(i) and we only sketch it. By (A.8)

$$\frac{1}{T}\sum_{t=1}^{T} k_{h,tr}\left\|\tilde{F}_t - H'F_t\right\|^2 \leq 4\left\|V_{NT}^{-1}\right\|^2 \sum_{l=1}^{4}\frac{1}{T}\sum_{t=1}^{T} k_{h,tr}\left\|V_{NT}A_l(t)\right\|^2 \equiv 4\left\|V_{NT}^{-1}\right\|^2\sum_{l=1}^{4} A_{lNT}(r), \text{ say.}$$

We prove (i) by find the bound for $A_{lNT}(r)$, $l = 1,2,3,4$, uniformly in r. Using the fact that $\max_s \sum_{s=1}^{T}[E(e_s^{\dagger\prime} e_t^{\dagger}/N)]^2 = O(1)$ and that

$$\frac{1}{T}\sum_{t=1}^{T} k_{h,tr} = \frac{1}{Th}\sum_{t=1}^{T} K_r\left(\frac{t-r}{Th}\right) = 1 + O\left(\frac{1}{Th}\right) \text{ uniformly in } r \text{ under Assumption A.4,}$$

we have

$$\begin{aligned}
\max_r A_{1NT}(r) &= \max_r \frac{1}{T}\sum_{t=1}^{T} k_{h,tr}\left\|V_{NT}A_1(t)\right\|^2 \leq \max_r \frac{1}{T}\sum_{t=1}^{T} k_{h,tr}\left\{\frac{1}{T}\sum_{s=1}^{T}\left\|\tilde{F}_s E(e_s^{\dagger\prime}e_t^{\dagger}/N)\right\|\right\}^2 \\
&\leq \frac{1}{T}\left\{\frac{1}{T}\sum_{l=1}^{T}\left\|\tilde{F}_l\right\|^2\right\}\max_t \sum_{s=1}^{T}\left[E(e_s^{\dagger\prime}e_t^{\dagger}/N)\right]^2 \max_r\left\{\frac{1}{T}\sum_{t=1}^{T}k_{h,tr}\right\} \\
&= T^{-1}O(1)O(1)O(1) = O_P(T^{-1}).
\end{aligned}$$

For $A_{2NT}(r)$, using notations defined in the proof of Lemma A.2(i) and by CS inequality, we have

$$\begin{aligned}
A_{2NT}(r) &= \frac{1}{T}\sum_{t=1}^{T}k_{h,tr}\left\|V_{NT}A_2(t)\right\|^2 = \frac{1}{T}\sum_{t=1}^{T}k_{h,tr}\left\|\frac{1}{T}\sum_{s=1}^{T}\tilde{F}_s\xi_{st}^{\dagger}\right\|^2 \\
&\leq \frac{2}{T}\sum_{t=1}^{T}k_{h,tr}\left\|\frac{1}{T}\sum_{s=1}^{T}(\tilde{F}_s - H'F_s)\xi_{st}^{\dagger}\right\|^2 + \frac{2}{T}\sum_{t=1}^{T}k_{h,tr}\left\|\frac{1}{T}\sum_{s=1}^{T}H'F_s\xi_{st}^{\dagger}\right\|^2 \\
&\equiv 2A_{2NT,1}(r) + 2A_{2NT,2}(r), \text{ say.}
\end{aligned}$$

For $A_{2NT,1}(r)$ we apply Lemma A.2(i) to obtain the rough bound

$$\begin{aligned}
\max_r A_{2NT,1}(r) &\leq \frac{1}{T}\sum_{s=1}^{T}\left\|\tilde{F}_s - H'F_s\right\|^2 \max_r \frac{1}{T}\sum_{t=1}^{T}k_{h,tr}\max_t \frac{1}{T}\sum_{s=1}^{T}\left\|\xi_{st}^{\dagger}\right\|^2 \\
&= O_P\left(C_{NT}^{-2}\right)O(1)O_P(1) = O_P\left(C_{NT}^{-2}\right)
\end{aligned}$$

as we can readily show that $\max_t \frac{1}{T}\sum_{s=1}^{T}||\xi_{st}^{\dagger}||^2 \leq \max_t \frac{1}{T}\sum_{s=1}^{T}E||\xi_{st}^{\dagger}||^2 + \max_t|\frac{1}{T}\sum_{s=1}^{T}(||\xi_{st}^{\dagger}||^2 - E||\xi_{st}^{\dagger}||^2)| = O(1) + o_P(1)$ by a simple application of Bernstein inequality for strong mixing processes. Let $\bar{A}_{2NT,2}(r) = \frac{1}{T^3}\sum_{t=1}^{T}k_{h,tr}\sum_{s=1}^{T}\sum_{l=1}^{T}F_sF_l'\xi_{st}^{\dagger}\xi_{lt}^{\dagger}$. Observing that $A_{2NT,2}(r) = \text{tr}(HH'\bar{A}_{2NT,2}(r))$, we can bound $A_{2NT,2}(r)$ by bounding each element of $\bar{A}_{2NT,2}(r)$. Let $a_{mn}(r)$ denote the (m,n)th element of $\bar{A}_{2NT,2}(r)$. Noting that $a_{mn}(r) \leq \{a_{mm}(r)a_{nn}(r)\}^{1/2}$, it suffices to bound $a_{mm}(r)$ for $m = 1,...,R$. Observe that

$$\begin{aligned}
a_{mm}(r) &= \frac{1}{T^3}\sum_{t=1}^{T}k_{h,tr}\sum_{s=1}^{T}\sum_{l=1}^{T}\iota_m'F_sF_l'\iota_m\xi_{st}^{\dagger}\xi_{lt}^{\dagger} = \frac{1}{T}\sum_{t=1}^{T}k_{h,tr}\left\{\frac{1}{T}\sum_{s=1}^{T}\iota_m'F_s\xi_{st}^{\dagger}\right\}^2 \\
&\leq \max_t\left|\frac{1}{T}\sum_{s=1}^{T}\iota_m'F_s\xi_{st}^{\dagger}\right|^2 \max_r \frac{1}{T}\sum_{t=1}^{T}k_{h,tr} = O_P\left(T^{-1}\ln T + a_{NT}^2\right)O(1).
\end{aligned}$$

It follows that $\max_r A_{2NT,2}(r) = O_P\left(T^{-1}\ln T\right)$ and $\max_r A_{2NT}(r) = O_P\left(T^{-1}\ln T + N^{-1}\right)$. To study $A_{3NT}(r)$, we first study $\bar{A}_{3NT}(r) \equiv \frac{1}{T}\sum_{t=1}^{T} k_{h,tr}\left\|\frac{\Lambda_0' e_t^{\dagger}}{\sqrt{N}}\right\|^2$.

$$\begin{aligned}\bar{A}_{3NT}(r) &= \frac{1}{NT}\sum_{t=1}^{T} k_{h,tr}\left\|\Lambda_0' e_t^{\dagger}\right\|^2 \leq \frac{2}{NT}\sum_{t=1}^{T} k_{h,tr}\left\|\Lambda_0' e_t\right\|^2 + \frac{2a_{NT}^2}{NT}\sum_{t=1}^{T} k_{h,tr}\left\|\Lambda_0' g_t^{\dagger}\right\|^2 \\ &\equiv 2\bar{A}_{3NT,1}(r) + \bar{A}_{3NT,2}(r), \text{ say.}\end{aligned}$$

For $\bar{A}_{3NT,1}(r)$, we have under Assumptions A.2-A.4

$$\begin{aligned}\max_r \bar{A}_{3NT,1}(r) &\leq \max_t \left[N^{-1}E\left\|\Lambda_0' e_t\right\|^2 + N^{-1}\left(\left\|\Lambda_0' e_t\right\|^2 - E\left\|\Lambda_0' e_t\right\|^2\right)\right]\max_r \frac{1}{T}\sum_{t=1}^{T} k_{h,tr} \\ &= O\left(1 + o_P(1)\right)O(1) = O_P(1).\end{aligned}$$

Similarly, $\max_r \bar{A}_{3NT,2}(r) = O_P\left(a_{NT}^2 N\right)$. Then $\max_r \bar{A}_{3NT}(r) = O_P(1)$ and

$$\begin{aligned}\max_r A_{3NT}(r) &= \max_r \frac{1}{T}\sum_{t=1}^{T} k_{h,tr}\left\|\frac{1}{T}\sum_{s=1}^{T}\tilde{F}_s F_s' \Lambda_0' e_t^{\dagger}/N\right\|^2 \\ &\leq \frac{1}{N}\max_r \bar{A}_{3NT}(r)\left[\frac{1}{T}\sum_{s=1}^{T}\left\|\tilde{F}_s\right\|^2\right]\left[\frac{1}{T}\sum_{s=1}^{T}\left\|F_s\right\|^2\right] = O_P(N^{-1}).\end{aligned}$$

For the fourth term, we have

$$\begin{aligned}\max_r A_{4NT}(r) &= \max_r \frac{1}{T}\sum_{t=1}^{T} k_{h,tr}\left\|\frac{1}{T}\sum_{s=1}^{T}\tilde{F}_s F_t' \Lambda_0' e_s^{\dagger}/N\right\|^2 \leq \max_r \frac{1}{N^2T^3}\sum_{t=1}^{T} k_{h,tr}\left\{\sum_{s=1}^{T}\left\|\tilde{F}_s F_t'\right\|\left\|\Lambda_0' e_s^{\dagger}\right\|\right\}^2 \\ &\leq \frac{1}{N}\max_r\left\{\frac{1}{T}\sum_{t=1}^{T} k_{h,tr}\left\|F_t\right\|^2\right\}\frac{1}{T}\sum_{s=1}^{T}\left\|\tilde{F}_s\right\|^2\frac{1}{NT}\sum_{s=1}^{T}\left\|\Lambda_0' e_s^{\dagger}\right\|^2 \\ &= N^{-1}O_P(1)\,O_P(1)\,O_P(1) = O_P(N^{-1}),\end{aligned}$$

as we can show that $\max_r \frac{1}{T}\sum_{t=1}^{T} k_{h,tr}\left\|F_t\right\|^2 \leq \max_r \frac{1}{T}\sum_{t=1}^{T} k_{h,tr}E\left\|F_t\right\|^2 + \max_r |\frac{1}{T}\sum_{t=1}^{T} k_{h,tr}[\left\|F_t\right\|^2 - E\left\|F_t\right\|^2]\} = O_P(1)$. Combining these results, we have $\max_r ||\frac{1}{T}\sum_{r=1}^{T} k_{h,tr}(\tilde{F}_r - H'F_r)\left(\tilde{F}_r - H'F_r\right)'|| = O_P(T^{-1}\ln T + N^{-1})$.

(ii) The proof of (ii) is analogous to that of (i) with some modifications similar to those used in the proof of Lemma A.2(ii).

(iii) Write $S_{Tr}^{(0)} = \frac{1}{T}\sum_{t=1}^{T} k_{h,tr}E\left(F_tF_t'\right) + \frac{1}{T}\sum_{t=1}^{T} k_{h,tr}\left[F_tF_t' - E\left(F_tF_t'\right)\right] \equiv S_{T,1}^{(r,0)} + S_{T,2}^{(r,0)}$, say. Using $E\left(F_tF_t'\right) = \Sigma_F$ and the Riemann sum approximation of integral, we have

$$\max_r\left\|S_{T,1}^{(r,0)} - \Sigma_F\right\| = \max_r\left|\frac{1}{Th}\sum_{t=1}^{T} K_r\left(\frac{t-r}{Th}\right) - 1\right|\left\|\Sigma_F\right\| = O\left(\frac{1}{Th}\right).$$

By Bernstein inequality for strong mixing processes, we can readily show that $\max_r ||S_{T,2}^{(r,0)}|| = O_P\left(T^{-1/2}(\ln T)^{1/2}\right)$. It follows that $\max_r ||S_{Tr}^{(0)} - \Sigma_F|| = O_P\left(T^{-1/2}(\ln T)^{1/2}\right)$.

(iv) Using $\tilde{F}_s = H'F_s + (\tilde{F}_s - H'F_s)$, we make the following decomposition:

$$\begin{aligned} S_{Tr} &= \frac{1}{T}\sum_{t=1}^{T} k_{h,tr}\tilde{F}_t\tilde{F}_t' \\ &= H'S_{Tr}^{(0)}H + \frac{1}{T}\sum_{t=1}^{T} k_{h,tr}H'F_t(\tilde{F}_t - H'F_t)' + \frac{1}{T}\sum_{t=1}^{T} k_{h,tr}(\tilde{F}_t - H'F_t)F_t'H \\ &\quad + \frac{1}{T}\sum_{t=1}^{T} k_{h,tr}(\tilde{F}_t - H'F_t)(\tilde{F}_t - H'F_t)' \\ &\equiv S_{Tr,1} + S_{Tr,2} + S_{Tr,3} + S_{Tr,4}, \text{ say.} \end{aligned}$$

By Lemmas A.2(vi) and A.3(iii), $\max_r ||S_{Tr,1} - \left(Q_0^{-1}\right)' \Sigma_F Q_0^{-1}|| = O_P((Th)^{-1/2} (\ln T)^{1/2} + N^{-1/2})$. By Lemma A.3(i)-(ii), $\max_r ||S_{Tr,2}|| = \max_r ||S_{Tr,3}|| = O_P(T^{-1}\ln T + N^{-1} + a_{NT})$ and $\max_r ||S_{Tr,4}|| = O_P(T^{-1}\ln T + N^{-1})$. Combining these results yield the desired result.

(v) This follows from Lemmas A.2(vi) and A.3 (iv) above. ■

Proof Lemma A.4. (i) First, using $\tilde{F}_tF_t' = H'F_tF_t' + (\tilde{F}_t - H'F_t)F_t' = H'\Sigma_F + H'\left(F_tF_t' - \Sigma_F\right) + (\tilde{F}_t - H'F_t)F_t'$, we have:

$$\begin{aligned} \frac{1}{T}\sum_{t=1}^{T} k_{h,tr}\tilde{F}_tF_t'g_{it} &= H'\Sigma_F \frac{1}{T}\sum_{t=1}^{T} k_{h,tr}g_{it} + H'\frac{1}{T}\sum_{t=1}^{T} k_{h,tr}\left(F_tF_t' - \Sigma_F\right)g_{it} + \frac{1}{T}\sum_{t=1}^{T} k_{h,tr}(\tilde{F}_t - H'F_t)F_t'g_{it} \\ &\equiv L_1(i,r) + L_2(i,r) + L_3(i,r), \text{ say.} \end{aligned}$$

By the uniform approximation property for Riemann integral and Bernstein inequality for mixing processes, we have that under Assumptions A.1 and A.3-A.5

$$\begin{aligned} \max_{i,r} \|L_1(i,r)\| &\leq \|H'\Sigma_F\| \bar{c}_g \max_r \frac{1}{T}\sum_{t=1}^{T} k_{h,tr} = \|H'\Sigma_F\| \bar{c}_g \left\{1 + O\left(\frac{1}{Th}\right)\right\} = O(1), \\ \max_{i,r} \|L_2(i,r)\| &\leq \|H'\| \max_r \left|\frac{1}{T}\sum_{t=1}^{T} k_{h,tr}\left(F_tF_t' - \Sigma_F\right)g_{it}\right| = o_p(1). \end{aligned}$$

In addition, by arguments as used in the proof of Lemma A.3(i), we can readily show $\max_{i,r} \|L_3(i,r)\| = o_P(1)$. Alternatively, we can apply CS inequality and Lemma A.2(i)

$$\begin{aligned} \max_{i,r} \|L_3(i,r)\| &\leq \max_{i,r} \left\{\frac{1}{T}\sum_{t=1}^{T} k_{h,tr}^2 \|F_tg_{it}\|^2\right\}^{1/2} \left\{\frac{1}{T}\sum_{t=1}^{T} \left\|\tilde{F}_t - H'F_t\right\|^2\right\}^{1/2} \\ &= O_P(h^{-1/2})O_P\left(C_{NT}^{-1}\right) = o_P(1). \end{aligned}$$

(ii) It is standard to show that $\max_{i,r} \left\|\frac{1}{T}\sum_{t=1}^{T} k_{h,tr}F_t'e_{it}\right\| = O_P(T^{-1/2}h^{-1/2}\ln(NT))$ by using Bernstein inequality for strong mixing processes.

(iii) Using $e_{is}^{\dagger} = e_{is} + a_{NT}F_s'g_{is}$ and CS inequality,

$$\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{1}{T}\sum_{s=1}^{T}(\tilde{F}_s - HF_s)e_{is}^{\dagger}\right\|^2 \leq \frac{2}{N}\sum_{i=1}^{N}\left\|\frac{1}{T}\sum_{s=1}^{T}(\tilde{F}_s - HF_s)e_{is}\right\|^2 + \frac{2a_{NT}^2}{N}\sum_{i=1}^{N}\left\|\frac{1}{T}\sum_{s=1}^{T}(\tilde{F}_s - HF_s)F_s'g_{is}\right\|^2.$$

It is standard to show that the first term is $O_P\left(C_{NT}^{-4}\right)$ and $\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{1}{T}\sum_{s=1}^{T}(\tilde{F}_s - HF_s)F_s'g_{is}\right\|^2 = O_P\left(C_{NT}^{-4}\right) + o_P\left(a_{NT}^2\right)$. It follows that $\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{1}{T}\sum_{s=1}^{T}(\tilde{F}_s - HF_s)e_{is}^{\dagger}\right\|^2 = O_P\left(C_{NT}^{-4}\right)$.

(iv) By (A.8) and CS inequality

$$\frac{1}{NT}\sum_{i=1}^{N}\sum_{r=1}^{T}\left\|\frac{1}{T}\sum_{t=1}^{T}\left(\tilde{F}_t-H'F_t\right)e_{it}k_{h,tr}\right\|^2\leq 4\sum_{l=1}^{4}\frac{1}{NT}\sum_{i=1}^{N}\sum_{r=1}^{T}\left\|\frac{1}{T}\sum_{t=1}^{T}A_l\left(t\right)e_{it}k_{h,tr}\right\|^2\equiv 4\sum_{l=1}^{4}I_l,\text{ say.}$$

Using $\tilde{F}_s=(\tilde{F}_s-H'F_s)+H'F_s$ and CS inequality,

$$\begin{aligned}
I_1 &= \frac{1}{NT}\sum_{i=1}^{N}\sum_{r=1}^{T}\left\|V_{NT}^{-1}\frac{1}{T}\sum_{t=1}^{T}\left\{\frac{1}{T}\sum_{s=1}^{T}\tilde{F}_sE(e_s^{\dagger\prime}e_t^{\dagger}/N)\right\}e_{it}k_{h,tr}\right\|^2\\
&\leq 2\left\|V_{NT}^{-1}\right\|^2\frac{1}{NT}\sum_{i=1}^{N}\sum_{r=1}^{T}\left\|\frac{1}{T}\sum_{t=1}^{T}\left\{\frac{1}{T}\sum_{s=1}^{T}(\tilde{F}_s-H'F_s)E(e_s^{\dagger\prime}e_t^{\dagger}/N)\right\}e_{it}k_{h,tr}\right\|^2\\
&\quad+2\left\|V_{NT}^{-1}H'\right\|^2\frac{1}{NT}\sum_{i=1}^{N}\sum_{r=1}^{T}\left\|\frac{1}{T}\sum_{t=1}^{T}\left\{\frac{1}{T}\sum_{s=1}^{T}F_sE(e_s^{\dagger\prime}e_t^{\dagger}/N)\right\}e_{it}k_{h,tr}\right\|^2\equiv 2I_{1,1}+2I_{1,2},\text{ say.}
\end{aligned}$$

For $I_{1,1}$, we have

$$\begin{aligned}
I_{1,1} &= \left\|V_{NT}^{-1}\right\|^2\frac{1}{NT}\sum_{i=1}^{N}\sum_{r=1}^{T}\left\|\frac{1}{T}\sum_{s=1}^{T}(\tilde{F}_s-H'F_s)\left(\frac{1}{T}\sum_{t=1}^{T}E(e_s^{\dagger\prime}e_t^{\dagger}/N)e_{it}k_{h,tr}\right)\right\|^2\\
&\leq \left\|V_{NT}^{-1}\right\|^2\frac{1}{T}\sum_{s=1}^{T}\left\|\tilde{F}_s-H'F_s\right\|^2\frac{1}{NT}\sum_{i=1}^{N}\sum_{r=1}^{T}\frac{1}{T}\sum_{s=1}^{T}\left\{\frac{1}{T}\sum_{t=1}^{T}E(e_s^{\dagger\prime}e_t^{\dagger}/N)e_{it}k_{h,tr}\right\}^2\\
&= O_P\left(C_{NT}^{-2}\right)O_P\left(N^{-1}\right)=O_P\left(N^{-1}C_{NT}^{-2}\right),
\end{aligned}$$

as one can readily show that $\frac{1}{NT}\sum_{i=1}^{N}\sum_{r=1}^{T}\frac{1}{T}\sum_{s=1}^{T}E[\frac{1}{T}\sum_{t=1}^{T}E(e_s^{\dagger\prime}e_t^{\dagger}/N)e_{it}k_{h,tr}]^2=O\left(N^{-1}+a_{NT}^2\right).$ For $I_{1,2}$, by straightforward moment calculations, we can show that

$$\begin{aligned}
&\frac{1}{NT}\sum_{i=1}^{N}\sum_{r=1}^{T}E\left\|\frac{1}{T}\sum_{t=1}^{T}\left\{\frac{1}{T}\sum_{s=1}^{T}F_sE(e_s^{\dagger\prime}e_t^{\dagger}/N)\right\}e_{it}k_{h,tr}\right\|^2\\
=&\ \frac{1}{N^3T^5}\sum_{i=1}^{N}\sum_{r=1}^{T}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{t_1=1}^{T}\sum_{s_1=1}^{T}E(e_s^{\dagger\prime}e_t^{\dagger})E(e_{s_1}^{\dagger\prime}e_{t_1}^{\dagger})E\left(F_s'F_{s_1}e_{it}e_{it_1}\right)k_{h,tr}k_{h,t_1r}\\
=&\ \frac{1}{N^3T^5}\sum_{i=1}^{N}\sum_{r=1}^{T}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{s_1=1}^{T}E(e_s^{\dagger\prime}e_t^{\dagger})E(e_{s_1}^{\dagger\prime}e_t^{\dagger})E\left(F_s'F_{s_1}e_{it}^2\right)k_{h,tr}^2=O\left(T^{-1}N^{-2}h^{-1}+a_{NT}^2T^{-1}h^{-1}\right).
\end{aligned}$$

So $I_{1,2}=O\left(T^{-1}N^{-2}h^{-1}+a_{NT}^2T^{-1}h^{-1}\right)$ and $I_1=O_P\left(N^{-1}C_{NT}^{-2}+T^{-1}N^{-2}h^{-1}+a_{NT}^2T^{-1}h^{-1}\right).$ For I_2, we have

$$\begin{aligned}
I_2 &= \frac{1}{NT}\sum_{i=1}^{N}\sum_{r=1}^{T}\left\| V_{NT}^{-1}\frac{1}{T}\sum_{t=1}^{T}\left\{\frac{1}{T}\sum_{s=1}^{T}\tilde{F}_s\left[e_s^{\dagger\prime}e_t^{\dagger}/N-E(e_s^{\dagger\prime}e_t^{\dagger}/N)\right]\right\}e_{it}k_{h,tr}\right\|^2 \\
&\leq \frac{4}{NT}\sum_{i=1}^{N}\sum_{r=1}^{T}\left\| V_{NT}^{-1}\frac{1}{T}\sum_{t=1}^{T}\left\{\frac{1}{T}\sum_{s=1}^{T}\tilde{F}_s\left[e_s'e_t/N-E(e_s'e_t/N)\right]\right\}e_{it}k_{h,tr}\right\|^2 \\
&\quad +\frac{4a_{NT}^2}{NT}\sum_{i=1}^{N}\sum_{r=1}^{T}\left\| V_{NT}^{-1}\frac{1}{T}\sum_{t=1}^{T}\left\{\frac{1}{T}\sum_{s=1}^{T}\tilde{F}_s\left[g_s^{\dagger\prime}g_t^{\dagger}/N-E(g_s^{\dagger\prime}g_t^{\dagger}/N)\right]\right\}e_{it}k_{h,tr}\right\|^2 \\
&\quad +\frac{4a_{NT}^2}{NT}\sum_{i=1}^{N}\sum_{r=1}^{T}\left\| V_{NT}^{-1}\frac{1}{T}\sum_{t=1}^{T}\left\{\frac{1}{T}\sum_{s=1}^{T}\tilde{F}_s\left[e_s'g_t^{\dagger}/N-E(e_s'g_t^{\dagger}/N)\right]\right\}e_{it}k_{h,tr}\right\|^2 \\
&\quad +\frac{4a_{NT}^4}{NT}\sum_{i=1}^{N}\sum_{r=1}^{T}\left\| V_{NT}^{-1}\frac{1}{T}\sum_{t=1}^{T}\left\{\frac{1}{T}\sum_{s=1}^{T}\tilde{F}_s\left[g_s^{\dagger\prime}g_t^{\dagger}/N-E(g_s^{\dagger\prime}g_t^{\dagger}/N)\right]\right\}e_{it}k_{h,tr}\right\|^2 \\
&\equiv 4I_{2,1}+4I_{2,2}+4I_{2,3}+4I_{2,4},\ \text{say.}
\end{aligned}$$

It is trivial to show that $I_{2,j}=o_P\left(a_{NT}^2\right)$ for $j=2,3$ and $I_{2,4}=O_P\left(a_{NT}^4\right)$. To bound $I_{2,1}$, notice that

$$\begin{aligned}
I_{2,1} &\leq \frac{2}{NT}\sum_{i=1}^{N}\sum_{r=1}^{T}\left\| V_{NT}^{-1}\frac{1}{T}\sum_{t=1}^{T}\left\{\frac{1}{T}\sum_{s=1}^{T}(\tilde{F}_s-H'F_s)\xi_{st}\right\}e_{it}k_{h,tr}\right\|^2 \\
&\quad +\frac{2}{NT}\sum_{i=1}^{N}\sum_{r=1}^{T}\left\| V_{NT}^{-1}H'\frac{1}{T}\sum_{t=1}^{T}\left\{\frac{1}{T}\sum_{s=1}^{T}F_s\xi_{st}\right\}e_{it}k_{h,tr}\right\|^2 \equiv 2I_{2,1}^{(1)}+2I_{2,1}^{(2)}.
\end{aligned}$$

One can readily show that $I_{2,1}^{(1)}=O_P\left(C_{NT}^{-2}N^{-1}\right)$. Noting that under Assumptions A.1(ii), A.3(iv) and A.4,

$$\begin{aligned}
&\frac{1}{NT}\sum_{i=1}^{N}\sum_{r=1}^{T}E\left\|\frac{1}{T}\sum_{t=1}^{T}\left\{\frac{1}{T}\sum_{s=1}^{T}F_s\xi_{st}\right\}e_{it}k_{h,tr}\right\|^2 \\
&= \frac{1}{N^3T^5}\sum_{r=1}^{T}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{t_1=1}^{T}\sum_{s_1=1}^{T}E\left(\xi_{st}\xi_{s_1t_1}e_t'e_{t_1}F_s'F_{s_1}\right)k_{h,tr}k_{h,t_1r} \\
&= \frac{1}{T^5}\sum_{r=1}^{T}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{t_1=1}^{T}\sum_{s_1=1}^{T}\left[E\left(\xi_{st}\xi_{s_1t_1}\xi_{tt_1}F_s'F_{s_1}\right)+E\left(\xi_{st}\xi_{s_1t_1}F_s'F_{s_1}\right)\gamma_N\left(t,t_1\right)\right]k_{h,tr}k_{h,t_1r} \\
&\leq N^{-3/2}\max_{s,t}\left\|N^{1/2}\xi_{st}\right\|_4^3\frac{1}{T^5}\sum_{r=1}^{T}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{t_1=1}^{T}\sum_{s_1=1}^{T}\left\|F_s'F_{s_1}\right\|_4k_{h,tr}k_{h,t_1r} \\
&\quad +N^{-1}\max_{s,t}\left\|N^{1/2}\xi_{st}\right\|_4^2\frac{1}{T^5}\sum_{r=1}^{T}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{t_1=1}^{T}\sum_{s_1=1}^{T}\gamma_N\left(t,t_1\right)\left\|F_s'F_{s_1}\right\|_2k_{h,tr}k_{h,t_1r} \\
&= O(N^{-3/2}+N^{-1}T^{-1}),
\end{aligned}$$

we have

$$I_{2,1}^{(2)}\leq\left\|V_{NT}^{-1}H'\right\|\frac{1}{NT}\sum_{i=1}^{N}\sum_{r=1}^{T}\left\|\frac{1}{T}\sum_{t=1}^{T}\left\{\frac{1}{T}\sum_{s=1}^{T}F_s\left[e_s'e_t/N-E(e_s'e_t/N)\right]\right\}e_{it}k_{h,tr}\right\|^2=O_P(N^{-3/2}+N^{-1}T^{-1}).$$

It follows that $I_{2,1} = O_P(C_{NT}^{-2}N^{-1})$ and $I_2 = O_P(C_{NT}^{-2}N^{-1}) + o_P(a_{NT}^2)$. In addition, we can readily show that

$$\begin{aligned}
I_3 &= \frac{1}{NT}\sum_{i=1}^{N}\sum_{r=1}^{T}\left\| V_{NT}^{-1}\frac{1}{T}\sum_{t=1}^{T}\left\{\frac{1}{T}\sum_{s=1}^{T}\tilde{F}_s F_s'(\Lambda_0' e_t^{\dagger}/N)\right\} e_{it}k_{h,tr}\right\|^2 \\
&\leq 2\left\|V_{NT}^{-1}\right\|\left\|\frac{1}{T}\sum_{s=1}^{T}\tilde{F}_s F_s'\right\|\frac{1}{NT}\sum_{i=1}^{N}\sum_{r=1}^{T}\left\|\frac{1}{NT}\sum_{t=1}^{T}\Lambda_0' e_t e_{it}k_{h,tr}\right\|^2 \\
&\quad +2a_{NT}^2\left\|V_{NT}^{-1}\right\|\left\|\frac{1}{T}\sum_{s=1}^{T}\tilde{F}_s F_s'\right\|\frac{1}{NT}\sum_{i=1}^{N}\sum_{r=1}^{T}\left\|\frac{1}{NT}\sum_{t=1}^{T}\Lambda_0' g_t^{\dagger} e_{it}k_{h,tr}\right\|^2 \\
&= O_P(N^{-3/2} + a_{NT}^2 T^{-1}h^{-1}),
\end{aligned}$$

and

$$\begin{aligned}
I_4 &= \frac{1}{NT}\sum_{i=1}^{N}\sum_{r=1}^{T}\left\| V_{NT}^{-1}\sum_{t=1}^{T}\left\{\frac{1}{T}\sum_{s=1}^{T}\tilde{F}_s F_t'\Lambda_0' e_s^{\dagger}/N\right\} e_{it}k_{h,tr}\right\|^2 \\
&\leq \frac{2}{NT}\sum_{i=1}^{N}\sum_{r=1}^{T}\left\| V_{NT}^{-1}\sum_{t=1}^{T}\left\{\frac{1}{T}\sum_{s=1}^{T}\left(\tilde{F}_s - H'F_s\right) e_s^{\dagger\prime}\Lambda_0 F_t/N\right\} e_{it}k_{h,tr}\right\|^2 \\
&\quad +\frac{2}{NT}\sum_{i=1}^{N}\sum_{r=1}^{T}\left\| V_{NT}^{-1}H'\sum_{t=1}^{T}\left\{\frac{1}{T}\sum_{s=1}^{T}F_s e_s^{\dagger\prime}\Lambda_0 F_t/N\right\} e_{it}k_{h,tr}\right\|^2 \\
&= O_P(C_{NT}^{-4}) + O_P(N^{-3/2} + a_{NT}^2 N^{-1/2})
\end{aligned}$$

It follows that $\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\|\frac{1}{T}\sum_{s=1}^{T}(\tilde{F}_s - H'F_s)e_{is}k_{h,st}\|^2 = O_P(N^{-3/2} + T^{-1}) + o_P(a_{NT}^2)$. ■

Proof Lemma A.5. (i) Using $F_tF_t' = E(F_tF_t') + F_tF_t' - E(F_tF_t')$, we have

$$\frac{1}{T}\sum_{t=1}^{T}k_{h,tr}F_tF_t'g_{it} = \Sigma_F\frac{1}{T}\sum_{t=1}^{T}k_{h,tr}g_{it} + \frac{1}{T}\sum_{t=1}^{T}k_{h,tr}\left(F_tF_t' - \Sigma_F\right)g_{it}.$$

The second term is $o_P(1)$ uniformly in (i, r) under Assumptions A.1(ii), A.3(iii), A.4 and A.5(i). For the first term, we consider three cases: (i1) $r \in (Th, T(1-h)]$, (i2) $r \in (1, Th]$, and (i3) $r \in (T(1-h), T]$. In case (i1), by the fact that the kernel function K has compact support on $[-1, 1]$, the uniform approximation of Riemann integral, and the dominated convergence theorem, we have

$$\begin{aligned}
\Sigma_F\frac{1}{T}\sum_{t=1}^{T}k_{h,tr}g_{it} &= \Sigma_F\frac{1}{Th}\sum_{t=1}^{T}K_r\left(\frac{t-r}{Th}\right)g_i\left(\frac{t-r+r}{T}\right) \\
&= \Sigma_F\frac{1}{Th}\sum_{t=r-Th}^{r+Th}K\left(\frac{t-r}{Th}\right)g_i\left(\frac{t-r+r}{T}\right) \\
&= \Sigma_F\int_0^1 K(u)\,g_i\left(uh + \frac{r}{T}\right)du + O\left(\frac{1}{Th}\right) \\
&= \Sigma_F g_i\left(\frac{r}{T}\right) + o(1).
\end{aligned}$$

In case (i2),

$$\begin{aligned}
\Sigma_F \frac{1}{T}\sum_{t=1}^{T} k_{h,tr} g_{it} &= \Sigma_F \frac{1}{Th}\sum_{t=1}^{T} K_r\left(\frac{t-r}{Th}\right) g_i\left(\frac{t-r+r}{T}\right) \\
&= \Sigma_F \frac{1}{Th\int_{-r/(Th)}^{1} K(u)\,du}\sum_{t=1}^{r+Th} K\left(\frac{t-r}{Th}\right) g_i\left(\frac{t-r+r}{T}\right) \\
&= \Sigma_F \frac{1}{\int_{-r/(Th)}^{1} K(u)\,du}\int_{-r/(Th)}^{1} K(u)\, g_i\left(uh+\frac{r}{T}\right) du + O\left(\frac{1}{Th}\right) \\
&= \Sigma_F g_i\left(\frac{r}{T}\right) + o(1).
\end{aligned}$$

Similar result holds in case (i3). It follows that $\frac{1}{T}\sum_{t=1}^{T} k_{h,tr} F_t F_t' g_{it} = \Sigma_F g_i\left(\frac{r}{T}\right) + o_P(1)$ uniformly in (i,r).

(ii) As in the above analysis, $\frac{1}{T}\sum_{t=1}^{T} F_t F_t' g_{it} = \Sigma_F \frac{1}{T}\sum_{t=1}^{T} g_{it} + o_P(1) = o_P(1)$. ■

Proof Lemma A.6. (i) By (A.3) and CS inequality,

$$\frac{1}{N}\sum_{i=1}^{N}\left\|\tilde{\lambda}_{i0} - H^{-1}\lambda_{i0}\right\|^2 \le \frac{3}{N}\sum_{i=1}^{N}\left\{\|D_5(i)\|^2 + \|D_6(i)\|^2 + \|D_7(i)\|^2\right\}.$$

By straightforward moment calculations and Chebyshev inequality,

$$\begin{aligned}
\frac{1}{N}\sum_{i=1}^{N}\|D_5(i)\|^2 &= \frac{1}{NT^2}\sum_{i=1}^{N}\mathrm{tr}\left(HH'\sum_{s=1}^{T}\sum_{r=1}^{T} F_s e_{is}^{\dagger} e_{ir}^{\dagger\prime} F_r'\right) \\
&\le \|H\|^2 \frac{1}{NT^2}\left\|\sum_{i=1}^{N}\sum_{s=1}^{T}\sum_{r=1}^{T} F_s e_{is}^{\dagger} e_{ir}^{\dagger\prime} F_r'\right\| = O_P\left(C_{NT}^{-2} + a_{NT}^2\right).
\end{aligned}$$

Using $e_{is}^{\dagger} = e_{is} + a_{NT} F_s' g_{is}$, we can readily show that

$$\begin{aligned}
\frac{1}{N}\sum_{i=1}^{N}\|D_6(i)\|^2 &= \frac{1}{N}\sum_{i=1}^{N}\left\|\frac{1}{T}\sum_{s=1}^{T}(\tilde{F}_s - HF_s)e_{is}^{\dagger}\right\|^2 \\
&\le \frac{2}{N}\sum_{i=1}^{N}\left\|\frac{1}{T}\sum_{s=1}^{T}(\tilde{F}_s - HF_s)e_{is}\right\|^2 + \frac{2a_{NT}^2}{N}\sum_{i=1}^{N}\left\|\frac{1}{T}\sum_{s=1}^{T}(\tilde{F}_s - HF_s)F_s' g_{is}\right\|^2. \\
&= O_P\left(C_{NT}^{-2} + a_{NT}^2\right) = O_P\left(C_{NT}^{-2}\right).
\end{aligned}$$

By Lemma A.2(i), $\frac{1}{N}\sum_{i=1}^{N}\|D_7(i)\|^2 \le \frac{1}{T}\left\|\tilde{F}'(\tilde{F}H^{-1} - F)\right\| \frac{1}{N}\sum_{i=1}^{N}\|\lambda_{i0}\|^2 = O_P\left(C_{NT}^{-2}\right)$. It follows that $\frac{1}{N}\sum_{i=1}^{N}\left\|\tilde{\lambda}_{i0} - H^{-1}\lambda_{i0}\right\|^2 = O_P\left(C_{NT}^{-2}\right)$.

(ii) By the CS inequality, $\frac{1}{T}\sum_{t=1}^{T}\left\|\tilde{F}_t\right\|^4 \le \frac{8}{T}\sum_{t=1}^{T}\left\|\tilde{F}_t - H'F_t\right\|^4 + \frac{8}{T}\sum_{t=1}^{T}\|H'F_s\|^4$. Apparently, the second term is bounded from above by $8\|H\|^4 \frac{1}{T}\sum_{s=1}^{T}\|F_s\|^4 = O_P(1)$ under Assumption A.1(ii). For the first term, we apply Lemma A.2(i) to obtain a rough bound

$$\begin{aligned}
\frac{8}{T}\sum_{t=1}^{T}\left\|\tilde{F}_t - H'F_t\right\|^4 &\le \max_s\left\|\tilde{F}_s - H'F_s\right\|^2 \frac{1}{T}\sum_{t=1}^{T}\left\|\tilde{F}_t - H'F_t\right\|^2 \\
&\le T\left\{\frac{1}{T}\sum_{t=1}^{T}\left\|\tilde{F}_t - H'F_t\right\|^2\right\}^2 = O_P\left(TC_{NT}^{-4}\right).
\end{aligned}$$

It follows that $\frac{1}{T}\sum_{t=1}^{T}\left\|\tilde{F}_t\right\|^4 = O_P\left(TC_{NT}^{-4}\right)+O_P\left(1\right)=O_P\left(1\right).$ Similarly,

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T}\left\|\tilde{F}_t\right\|^6 &\leq \frac{32}{T}\sum_{t=1}^{T}\left\|\tilde{F}_t-H'F_t\right\|^6+\frac{32}{T}\sum_{t=1}^{T}\left\|H'F_s\right\|^6\\ &\leq \max_t\left\|\tilde{F}_t-H'F_t\right\|^2\frac{32}{T}\sum_{t=1}^{T}\left\|\tilde{F}_t-H'F_t\right\|^4+O_P\left(1\right)\\ &= T\left\{\frac{1}{T}\sum_{t=1}^{T}\left\|\tilde{F}_t-H'F_t\right\|^2\right\}\frac{32}{T}\sum_{t=1}^{T}\left\|\tilde{F}_t-H'F_t\right\|^4+O_P\left(1\right)\\ &= TO_P\left(C_{NT}^{-2}\right)TO_P\left(C_{NT}^{-4}\right)+O_P\left(1\right)=O_P\left(1\right).\end{aligned}$$

(iii) Noting that $\int_{-t/(Th)}^{1}K\left(u\right)du\geq\int_0^1K\left(u\right)du=\frac{1}{2}$ for any $t\in\left(0,\lfloor Th\rfloor\right)$ and $\int_{-1}^{1-t/(Th)}K\left(u\right)du\geq\int_{-1}^{0}K\left(u\right)du=\frac{1}{2}$ for any $t\in\left(\lfloor T(1-h)\rfloor,T\right]$, we have

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T}k_{h,st}^2 &= \frac{1}{Th^2}\sum_{t=1}^{T}K_t\left(\frac{s-t}{Th}\right)^2\\ &= \frac{1}{Th^2}\left\{\sum_{t=\lfloor Th\rfloor}^{\lfloor T(1-h)\rfloor}+\sum_{t=1}^{\lfloor Th\rfloor-1}+\sum_{t=\lfloor T(1-h)\rfloor+1}^{T}\right\}K_t\left(\frac{s-t}{Th}\right)^2\\ &= \frac{1}{Th^2}\sum_{t=\lfloor Th\rfloor}^{\lfloor T(1-h)\rfloor}K\left(\frac{t-s}{Th}\right)^2+\frac{1}{Th^2}\sum_{t=1}^{\lfloor Th\rfloor-1}\frac{1}{\left(\int_{-t/(Th)}^{1}K\left(u\right)du\right)^2}K\left(\frac{s-t}{Th}\right)^2\\ &\quad+\frac{1}{Th^2}\sum_{t=\lfloor T(1-h)\rfloor+1}^{T}\frac{1}{\left(\int_{-1}^{1-t/(Th)}K\left(u\right)du\right)^2}K_t\left(\frac{s-t}{Th}\right)^2\\ &\leq \frac{1}{Th^2}\sum_{t=\lfloor Th\rfloor}^{\lfloor T(1-h)\rfloor}K\left(\frac{t-s}{Th}\right)^2+\frac{4}{Th^2}\sum_{t=1}^{\lfloor Th\rfloor-1}K\left(\frac{t-s}{Th}\right)^2+\frac{4}{Th^2}\sum_{t=\lfloor T(1-h)\rfloor+1}^{T}K\left(\frac{t-s}{Th}\right)^2.\end{aligned}$$

By the uniform approximation property of Riemann integral, $\frac{1}{Th}\sum_{t=\lfloor Th\rfloor}^{\lfloor T(1-h)\rfloor}K\left(\frac{t-s}{Th}\right)^2=\int_{-1}^{1}K\left(u\right)^2du+O\left(\frac{1}{Th}\right)$ uniformly in s under Assumption A.4. So the first term is $O\left(\frac{1}{Th}\right)$ uniformly in s. Similar results hold for the other two terms. Thus $\max_s\frac{1}{T}\sum_{t=1}^{T}k_{h,st}^2=O\left(h^{-1}\right).$

(iv) Observe that $L_{st}=\left(k_{h,st}S_{Tt}^{-1}-\mathbb{I}_R\right)\left(k_{h,st}S_{Tt}^{-1}-\mathbb{I}_R\right)\leq k_{h,st}^2S_{Tt}^{-1}S_{Tt}^{-1}+\mathbb{I}_R$, we have

$$\begin{aligned}\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\tilde{F}_s'L_{st}\tilde{F}_s\tilde{F}_s'\tilde{F}_s &\leq \frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}k_{h,st}^2\tilde{F}_s'S_{Tt}^{-1}S_{Tt}^{-1}\tilde{F}_s\tilde{F}_s'\tilde{F}_s+\frac{1}{T}\sum_{s=1}^{T}\tilde{F}_s'\tilde{F}_s\tilde{F}_s'\tilde{F}_s\\ &\leq \max_t\left\|S_{Tt}^{-1}\right\|^2\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}k_{h,st}^2\left\|\tilde{F}_s\right\|^4+\frac{1}{T}\sum_{s=1}^{T}\left\|\tilde{F}_s\right\|^4\\ &\leq \left[\max_t\left\|S_{Tt}^{-1}\right\|^2\max_s\left(\frac{1}{T}\sum_{t=1}^{T}k_{h,st}^2\right)+1\right]\frac{1}{T}\sum_{s=1}^{T}\left\|\tilde{F}_s\right\|^4\\ &= \left[O_P\left(1\right)O\left(h^{-1}\right)+1\right]O_P\left(1\right)=O\left(h^{-1}\right).\end{aligned}$$

(v) First $\frac{1}{T}\sum_{t=1}^{T}||(\tilde{F}_t-H'F_t)(\tilde{F}_t-H'F_t)'||^2\leq T\{\frac{1}{T}\sum_{t=1}^{T}||\tilde{F}_t-H'F_t||^2\}^2=O\left(TC_{NT}^{-4}\right).$ By (A.8), $\frac{1}{T}\sum_{t=1}^{T}||(\tilde{F}_t-H'F_t)F_t'H||^2\leq\left\|V_{NT}^{-1}\right\|\left\|H\right\|\{\frac{3}{T}\sum_{t=1}^{T}\left\|V_{NT}\left(A_{1t}+A_{2t}\right)F_t'\right\|^2+\frac{3}{T}\sum_{t=1}^{T}\left\|V_{NT}A_{3t}F_t'\right\|^2+\frac{3}{T}\sum_{t=1}^{T}$

$\|V_{NT}A_{4t}F_t'\|^2\}$. We bound each term in the last pair of curly brackets. The first term satisfies

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T}\|V_{NT}\left(A_{1t}+A_{2t}\right)F_t'\|^2 &= \frac{1}{T^3N^2}\sum_{t=1}^{T}\left\|\sum_{s=1}^{T}\tilde{F}_s e_s^{\dagger\prime}e_t^{\dagger}F_t'\right\|^2 \\
&\leq \frac{2}{T^3N^2}\sum_{t=1}^{T}\left\|\sum_{s=1}^{T}(\tilde{F}_s-H'F_s)e_s^{\dagger\prime}e_t^{\dagger}F_t\right\|^2+\frac{2}{T^3N^2}\sum_{t=1}^{T}\left\|H'\sum_{s=1}^{T}F_s e_s^{\dagger\prime}e_t^{\dagger}F_t\right\|^2 \\
&\equiv 2A_1+2A_2,\ \text{say.}
\end{aligned}$$

For A_1, we only consider its rough bound. Noting that $\frac{1}{T^2N^2}\sum_{t=1}^{T}\sum_{s=1}^{T}E\|e_s^{\dagger\prime}e_t^{\dagger}F_t'\|^2=O\left(1\right)$, we have by Lemma A.2(i),

$$A_1\leq\left\{\frac{1}{T}\sum_{s=1}^{T}\left\|\tilde{F}_s-H'F_s\right\|^2\right\}\left\{\frac{1}{T^2N^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\left\|e_s^{\dagger\prime}e_t^{\dagger}F_t\right\|^2\right\}=O_P\left(C_{NT}^{-2}\right)O_P\left(1\right)=O_P\left(C_{NT}^{-2}\right).$$

For A_2, we observe that $A_2\leq 2\left\|H\right\|^2\bar{A}_2$ where $\bar{A}_2=\frac{1}{T^3N^2}\sum_{t=1}^{T}\left\|\sum_{s=1}^{T}F_s e_s^{\dagger\prime}e_t^{\dagger}F_t\right\|^2$. By CS inequality,

$$\begin{aligned}
E\left(\bar{A}_2\right) &\leq \frac{4}{T^3N^2}\sum_{t=1}^{T}E\left\|\sum_{s=1}^{T}F_s e_s' e_t F_t'\right\|^2+\frac{4\gamma_{NT}^2}{T^3N^2}\sum_{t=1}^{T}E\left\|\sum_{s=1}^{T}F_s g_s^{\dagger\prime}e_t F_t'\right\|^2 \\
&\quad+\frac{4\gamma_{NT}^2}{T^3N^2}\sum_{t=1}^{T}E\left\|\sum_{s=1}^{T}F_s e_s' g_t^{\dagger}F_t'\right\|^2+\frac{4\gamma_{NT}^4}{T^3N^2}\sum_{t=1}^{T}E\left\|\sum_{s=1}^{T}F_s g_s^{\dagger\prime}g_t^{\dagger}F_t'\right\|^2 \\
&\equiv 4A_{2,1}+4A_{2,2}+4A_{2,3}+4A_{2,4}.
\end{aligned}$$

Noting that under Assumptions A.1(ii) and A.3(iv)

$$\begin{aligned}
\frac{1}{TN^2}\sum_{t=1}^{T}E\left\|\frac{1}{T}\sum_{s=1}^{T}F_s E\left(e_s'e_t\right)F_t'\right\|^2 &= \frac{1}{T^3N^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{r=1}^{T}E\left(e_s'e_t\right)E\left(e_r'e_t\right)E\left(F_t'F_tF_r'F_s\right) \\
&\leq \frac{1}{T}\max_t\|F_t\|_4^4\left\{\max_t\sum_{s=1}^{T}\gamma_N\left(s,t\right)\right\}^2=O\left(T^{-1}\right)
\end{aligned}$$

and

$$\begin{aligned}
\frac{1}{TN^2}\sum_{t=1}^{T}E\left\|\frac{1}{T}\sum_{s=1}^{T}F_s\left[e_s'e_t-E\left(e_s'e_t\right)\right]F_t'\right\|^2 &= \frac{1}{T^3}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{r=1}^{T}E(\xi_{st}\xi_{rt}F_t'F_tF_r'F_s) \\
&\leq N^{-1}\left\|N^{1/2}\xi_{st}\right\|_4^2\frac{1}{T^3}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{r=1}^{T}\|F_t'F_t\|_4\|F_r'F_s\|_4=O\left(N^{-1}\right)
\end{aligned}$$

we have

$$\begin{aligned}
A_{2,1} &\leq \frac{2}{TN^2}\sum_{t=1}^{T}E\left\|\frac{1}{T}\sum_{s=1}^{T}F_s E\left(e_s'e_t\right)F_t'\right\|^2+\frac{2}{TN^2}\sum_{t=1}^{T}E\left\|\frac{1}{T}\sum_{s=1}^{T}F_s\left[e_s'e_t-E\left(e_s'e_t\right)\right]F_t'\right\|^2 \\
&= O\left(T^{-1}\right)+O\left(N^{-1}\right)=O\left(C_{NT}^{-2}\right).
\end{aligned}$$

For $A_{1,2}$, $A_{1,3}$, and $A_{1,4}$, one can readily obtain their rough bounds given by $O(a_{NT}^2)$, $O(a_{NT}^2)$, and $O(a_{NT}^4)$, respectively. It follows that $A_2=O_P(C_{NT}^{-2})$ and $\frac{1}{T}\sum_{t=1}^{T}\|V_{NT}\left(A_{1t}+A_{2t}\right)F_t'\|^2=O_P(C_{NT}^{-2})$. In addition, noting that $\frac{1}{TN^2}\sum_{t=1}^T E||\Lambda_0' e_t^\dagger F_t'||^2 = O(N^{-1}+a_{NT}^2) = O(N^{-1})$, we have

$$\frac{1}{T}\sum_{t=1}^T \|V_{NT}A_{3t}F_t'\|^2 = \frac{1}{T^3N^2}\sum_{t=1}^T \left\|\sum_{s=1}^T \tilde{F}_s F_s' \Lambda_0' e_t^\dagger F_t'\right\|^2 \le \left\{\frac{1}{T}\sum_{s=1}^T \left\|\tilde{F}_s F_s'\right\|^2\right\} \frac{1}{TN^2}\sum_{t=1}^T \left\|\Lambda_0' e_t^\dagger F_t'\right\|^2 = O_P(N^{-1}).$$

Similarly,

$$\frac{1}{T}\sum_{t=1}^T \|V_{NT}A_{4t}F_t'\|^2 = \frac{1}{T^3N^2}\sum_{t=1}^T \left\|\sum_{s=1}^T \tilde{F}_s e_s^{\dagger\prime}\Lambda_0 F_t F_t'\right\|^2 \le \frac{1}{T^2N^2}\left\|\sum_{s=1}^T \tilde{F}_s e_s^{\dagger\prime}\Lambda_0\right\|^2 \frac{1}{T}\sum_{t=1}^T \|F_t\|^2 = O_P(N^{-1}),$$

because

$$\begin{aligned}
\frac{1}{T^2N^2}\left\|\sum_{s=1}^T \tilde{F}_s e_s^{\dagger\prime}\Lambda_0\right\|^2 &\le \frac{2}{T^2N^2}\left\|\sum_{s=1}^T \left(\tilde{F}_s - H'F_s\right) e_s^{\dagger\prime}\Lambda_0\right\|^2 + \frac{2}{T^2N^2}\left\|H'\sum_{s=1}^T F_s e_s^{\dagger\prime}\Lambda_0\right\|^2 \\
&\le \frac{2}{N}\frac{1}{T}\sum_{s=1}^T \left\|\tilde{F}_s - H'F_s\right\|^2 \frac{1}{TN}\sum_{s=1}^T \left\|e_s^{\dagger\prime}\Lambda_0\right\|^2 + \|H\|^2 \frac{2}{T^2N^2}\left\|\sum_{s=1}^T F_s e_s^{\dagger\prime}\Lambda_0\right\| \\
&= N^{-1}O_P(C_{NT}^{-2}) + O_P(N^{-1}) = O_P(N^{-1}).
\end{aligned}$$

It follows that $\frac{1}{T}\sum_{t=1}^T ||(\tilde{F}_t - H'F_t)\tilde{F}_t'||^2 \le \frac{2}{T}\sum_{t=1}^T ||(\tilde{F}_t - H'F_t)(\tilde{F}_t - H'F_t)'||^2 + \frac{1}{T}\sum_{t=1}^T ||(\tilde{F}_t - H'F_t)F_t H||^2$ $= O_P(TC_{NT}^{-4}) + O_P(C_{NT}^{-2}) = O_P(C_{NT}^{-2} + TN^{-2})$.

(vi) By (A.8), $\frac{1}{NT}\sum_{t=1}^T \left\|\sum_{i=1}^N \lambda_{i0}' H'^{-1}(\tilde{F}_t - H'F_t)e_{it}\right\|^2 \le \sum_{j=1}^4 \frac{1}{NT}\sum_{t=1}^T \left\|\sum_{i=1}^N \lambda_{i0}' H'^{-1} A_j(t) e_{it}\right\|^2 \equiv$ $4\sum_{j=1}^4 II_j$, say. For II_1, we have

$$\begin{aligned}
II_1 &= \frac{1}{NT}\sum_{t=1}^T \left\|\sum_{i=1}^N \lambda_{i0}' H'^{-1} V_{NT}^{-1}\frac{1}{T}\sum_{s=1}^T \tilde{F}_s E(e_s^{\dagger\prime}e_t^\dagger/N)e_{it}\right\|^2 \\
&= \frac{1}{NT}\sum_{t=1}^T \left\|\mathrm{tr}\left(H'^{-1}V_{NT}^{-1}\frac{1}{T}\sum_{s=1}^T \tilde{F}_s E(e_s^{\dagger\prime}e_t^\dagger/N)\sum_{i=1}^N e_{it}\lambda_{i0}'\right)\right\|^2 \\
&\le \left\|H'^{-1}V_{NT}^{-1}\right\|^2 \frac{1}{NT}\sum_{t=1}^T \left\|\frac{1}{T}\sum_{s=1}^T \tilde{F}_s E(e_s^{\dagger\prime}e_t^\dagger/N)\sum_{i=1}^N e_{it}\lambda_{i0}'\right\|^2.
\end{aligned}$$

Note that

$$\begin{aligned}
&\frac{1}{NT}\sum_{t=1}^T \left\|\frac{1}{T}\sum_{s=1}^T \tilde{F}_s E(e_s^{\dagger\prime}e_t^\dagger/N)\sum_{i=1}^N e_{it}\lambda_{i0}'\right\|^2 \\
\le\ &\frac{4}{NT}\sum_{t=1}^T \left\|\frac{1}{T}\sum_{s=1}^T \tilde{F}_s E(e_s'e_t/N)\sum_{i=1}^N e_{it}\lambda_{i0}'\right\|^2 + \frac{4a_{NT}}{NT}\sum_{t=1}^T \left\|\frac{1}{T}\sum_{s=1}^T \tilde{F}_s E(e_s'\breve{g}_t/N)\sum_{i=1}^N e_{it}\lambda_{i0}'\right\|^2 \\
&+\frac{4a_{NT}}{NT}\sum_{t=1}^T \left\|\frac{1}{T}\sum_{s=1}^T \tilde{F}_s E(\breve{g}_s'e_t/N)\sum_{i=1}^N e_{it}\lambda_{i0}'\right\|^2 + \frac{a_{NT}^2}{NT}\sum_{t=1}^T \left\|\frac{1}{T}\sum_{s=1}^T \tilde{F}_s E(\breve{g}_s'\breve{g}_t/N)\sum_{i=1}^N e_{it}\lambda_{i0}'\right\|^2 \\
\equiv\ &4II_{1,1} + 4II_{1,2} + 4II_{1,3} + 4II_{1,4}, \text{ say.}
\end{aligned}$$

One can readily show that $II_{1,1} = O_P(T^{-1})$, $II_{1,2} = O_P(a_{NT}N^{-1}\ln(NT))$, $II_{1,3} = O_P(a_{NT}N^{-1}\ln(NT))$, and $II_{1,4} = O_P(a_{NT}^2)$. It follows that $II_1 = O_P(C_{NT}^{-1})$. Similarly,

$$\begin{aligned}
II_2 &= \frac{1}{NT}\sum_{t=1}^{T}\left\|\sum_{i=1}^{N}\lambda_{i0}' H'^{-1} V_{NT}^{-1}\frac{1}{T}\sum_{s=1}^{T}\tilde{F}_s\left[e_s^{\dagger\prime}e_t^{\dagger}/N - E(e_s^{\dagger\prime}e_t^{\dagger}/N)\right]e_{it}\right\|^2 \\
&= \frac{1}{NT}\sum_{t=1}^{T}\left\|\operatorname{tr}\left(H'^{-1} V_{NT}^{-1}\frac{1}{T}\sum_{s=1}^{T}\tilde{F}_s\left[e_s^{\dagger\prime}e_t^{\dagger}/N - E(e_s^{\dagger\prime}e_t^{\dagger}/N)\right]\sum_{i=1}^{N}e_{it}\lambda_{i0}'\right)\right\|^2 \\
&\leq \left\|H'^{-1} V_{NT}^{-1}\right\|\frac{1}{NT}\sum_{t=1}^{T}\left\|\frac{1}{T}\sum_{s=1}^{T}\tilde{F}_s\left[e_s^{\dagger\prime}e_t^{\dagger}/N - E(e_s^{\dagger\prime}e_t^{\dagger}/N)\right]\sum_{i=1}^{N}e_{it}\lambda_{i0}'\right\|^2 \\
&\leq \left\|H'^{-1} V_{NT}^{-1}\right\|\frac{1}{NT}\sum_{t=1}^{T}\left\|\frac{1}{T}\sum_{s=1}^{T}\tilde{F}_s\left[e_s^{\dagger\prime}e_t^{\dagger}/N - E(e_s^{\dagger\prime}e_t^{\dagger}/N)\right]\right\|^2\left\|\sum_{i=1}^{N}e_{it}\lambda_{i0}'\right\|^2 \\
&\leq \left\|H'^{-1} V_{NT}^{-1}\right\|\frac{1}{T}\sum_{s=1}^{T}\left\|\tilde{F}_s\right\|^2\frac{1}{NT}\sum_{t=1}^{T}\left\{\frac{1}{T}\sum_{s=1}^{T}\left[e_s^{\dagger\prime}e_t^{\dagger}/N - E(e_s^{\dagger\prime}e_t^{\dagger}/N)\right]^2\left\|\sum_{i=1}^{N}e_{it}\lambda_{i0}'\right\|^2\right\} \\
&= O_P(1)\,O_P(1)\,O_P\left(N^{-1}\right) = O_P\left(N^{-1}\right),
\end{aligned}$$

$$\begin{aligned}
II_3 &= \frac{1}{NT}\sum_{t=1}^{T}\left\|\sum_{i=1}^{N}\lambda_{i0}' H'^{-1} V_{NT}^{-1}\frac{1}{T}\sum_{s=1}^{T}\tilde{F}_s F_s'\left(\Lambda_0' e_t^{\dagger}/N\right)e_{it}\right\|^2 \\
&= \frac{1}{NT}\sum_{t=1}^{T}\left\|\operatorname{tr}\left(H'^{-1} V_{NT}^{-1}\frac{1}{T}\sum_{s=1}^{T}\tilde{F}_s F_s'\left(\Lambda_0' e_t^{\dagger}/N\right)\sum_{i=1}^{N}e_{it}\lambda_{i0}'\right)\right\|^2 \\
&\leq \left\|H'^{-1} V_{NT}^{-1}\right\|\left\|\frac{1}{T}\sum_{s=1}^{T}\tilde{F}_s F_s'\right\|^2\frac{1}{NT}\sum_{t=1}^{T}\left\|\left(\Lambda_0' e_t^{\dagger}/N\right)\sum_{i=1}^{N}e_{it}\lambda_{i0}'\right\|^2 \\
&= O_P(1)\,O_P(1)\,O_P\left(N^{-1}\right) = O_P\left(N^{-1}\right),
\end{aligned}$$

and

$$\begin{aligned}
II_4 &= \frac{1}{NT}\sum_{t=1}^{T}\left\|\sum_{i=1}^{N}\lambda_{i0}' H'^{-1} V_{NT}^{-1}\frac{1}{T}\sum_{s=1}^{T}\tilde{F}_s F_t'\left(\Lambda_0' e_s^{\dagger}/N\right)e_{it}\right\|^2 \\
&= \frac{1}{NT}\sum_{t=1}^{T}\left\|\operatorname{tr}\left(H'^{-1} V_{NT}^{-1}\frac{1}{T}\sum_{s=1}^{T}\tilde{F}_s\left(e_s^{\dagger\prime}\Lambda_0/N\right)\sum_{i=1}^{N}F_t e_{it}\lambda_{i0}'\right)\right\|^2 \\
&\leq \left\|H'^{-1} V_{NT}^{-1}\right\|\left\|\frac{1}{T}\sum_{s=1}^{T}\tilde{F}_s\left(e_s^{\dagger\prime}\Lambda_0/N\right)\right\|^2\frac{1}{NT}\sum_{t=1}^{T}\left\|\sum_{i=1}^{N}F_t e_{it}\lambda_{i0}'\right\|^2 \\
&= O_P(1)\,O_P(N^{-1})O_P(1) = O_P(N^{-1}).
\end{aligned}$$

Consequently, $\frac{1}{NT}\sum_{t=1}^{T}||\sum_{i=1}^{N}\lambda_{i0}' H'^{-1}(\tilde{F}_t - H' F_t)e_{it}||^2 = O_P\left(C_{NT}^{-2}\right)$.

(vii) The proof is analogous to that of Lemma A.2(ii) and thus omitted. ■

B Some Additional Simulation and Applications Results

In this appendix, we report some additional simulation and applications results.

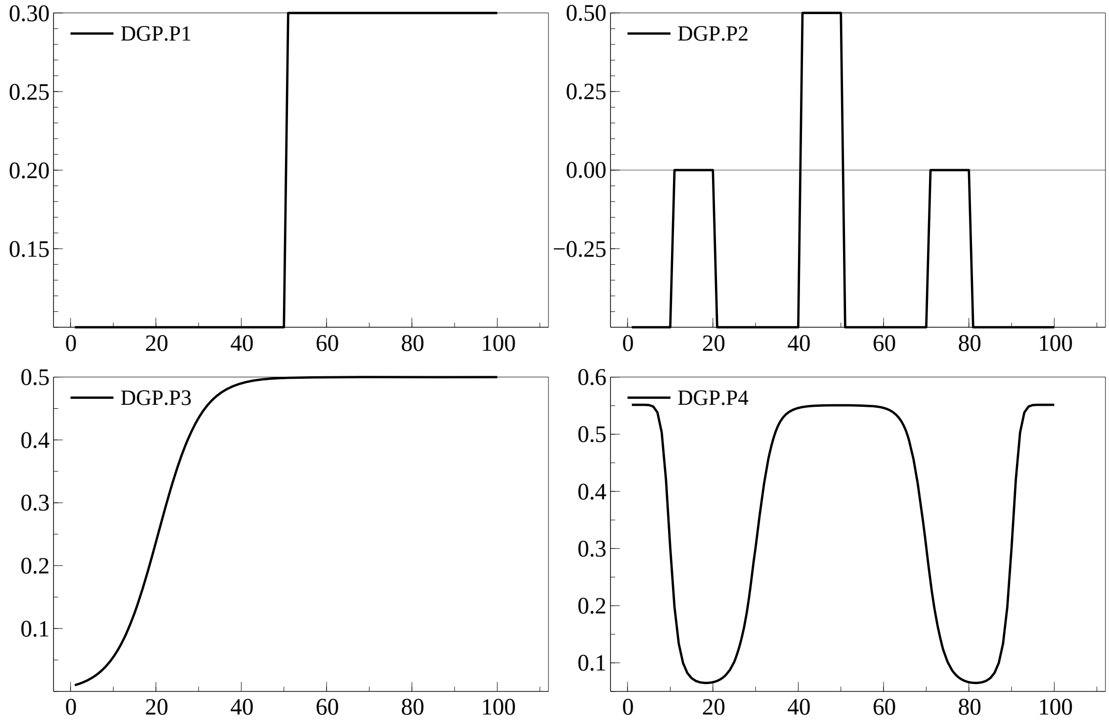

Figure 1: The factor loadings' paths for DGP.P1-P4 when $T = 100$

B.1 Some additional simulation results

First, following the suggestion of an anonymous referee, we plot λ_{it} for DGP.P1-P4 as a function of t for a representative cross-sectional unit. As we mentioned in the paper, DGP.P1-P2 have a single and multiple structural breaks, respectively, while DGP.P3-P4 describe two kinds of smooth structural changes. Among them, the factor loadings given by DGP.P3 are monotonic functions of t/T (or t), while the factor loadings given by DGP.P4 are smooth transition functions of t/T with multiple regime shifts. Figure 1 plots the paths of factor loadings under DGP.P1-P4 as functions of t when $T = 100$.

Second, to examine the sensitivity of our nonparametric test to the choice of the bandwidth parameter h, we set

$$h = c \cdot \frac{2.35}{\sqrt{12}} T^{-1/5}$$

for $c = 0.5$, 1 and 1.5. Tables A.1 and A.2 report the empirical rejection rates of our test at the 5% and 10% significance levels when the number of common factors is fixed as the true value and determined by BN's information criterion, respectively. As shown in Table A.1, the size of our test is robust to the choice of bandwidth. However, the power of our test reported in Table A.2 is a bit sensitive to the choice of bandwidth. For DGPs P1, P3, P5, and P7, the larger the bandwidth, the higher the power. In contrast, the power of the test for DGPs P2, P4, P6 and P8 tends to decrease as the bandwidth increases. Moreover, the power increases quickly as either N or T increases.

Third, we consider the tests when the number of factors are estimated by using Su and Wang's (2017) local-PCA-based information criterion. As mentioned in the paper, Su and Wang's (2017) information criterion can consistently estimate the true number of breaks under both the null and alternative hypotheses.

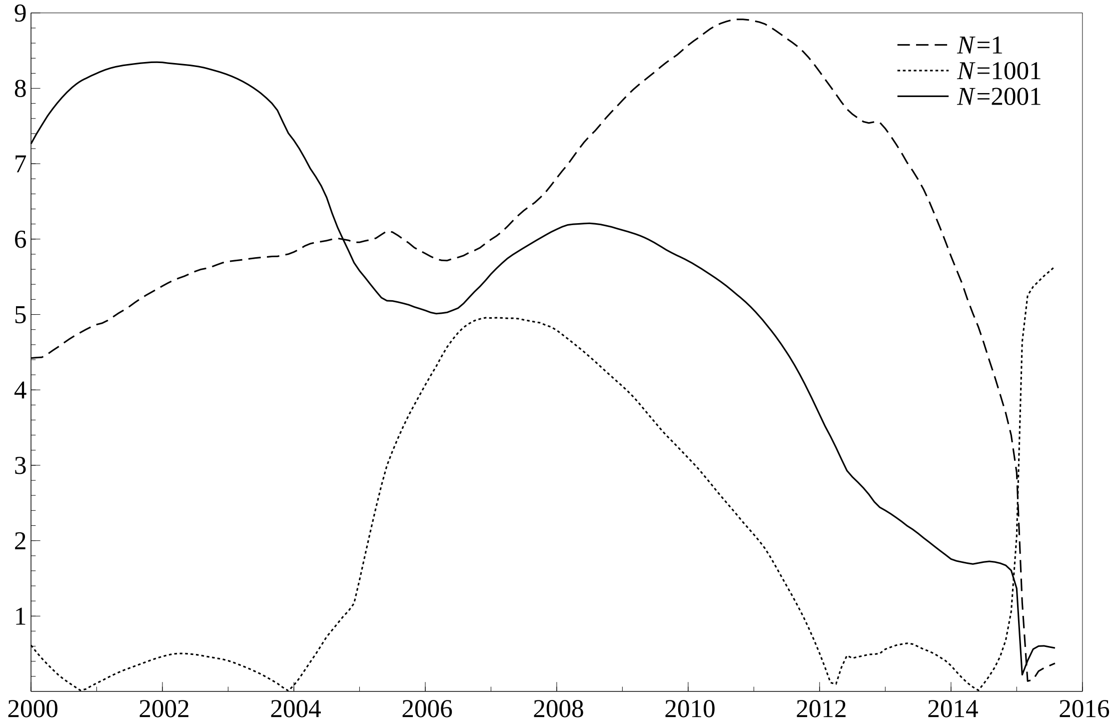

Figure 2: Some representative factor loadings estimated by local PCA

Tables A.3 and A.4 report the empirical rejection rates of several tests considered in the paper. As expected, the results in these two tables are quite similar to those in Tables 1 and 2.

B.2 Some additional application results

Following the suggestion of an anonymous referee, we use Su and Wang's (2017) local PCA to estimate the time-varying factor loadings in the empirical study. Since there are $N = 2684$ stocks and the factor loadings for these stocks are quite different from each other, it is impossible to plot them one by one. For this reason, we only plot the estimates of some representative factor loadings in Figure 2. From the figure we can see that the estimated factor loadings show significant structural changes that very likely appear to be smooth structural changes.

References

Bai, J. (2003) Inferential theory for factor models of large dimensions. *Econometrica* 71, 135-172.

Pollard, D. (1984) *Convergence of Stochastic Processes.* Springer-Verlag, New York.

Stewart, G.W. & J.-G. Sun (1990) *Matrix Perturbation Theory.* Academic Press, Boston.

Sun, S. & C.Y. Chiang (1997) Limiting behavior of the perturbed empirical distribution functions evaluated at U-statistics for strongly mixing sequences of random variables. *International Journal of Stochastic Analysis* 10, 3-20.

Table A.1 The size of our test with different bandwidth sequences under DGP.S1-S4

DGP	N	T	R is fixed to the true value						R is determined from the data					
			$c = 0.5$		$c = 1$		$c = 1.5$		$c = 0.5$		$c = 1$		$c = 1.5$	
			5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
S1	100	100	6.6	12.0	5.0	10.8	4.4	10.8	6.6	12.0	5.0	10.8	4.4	10.8
	100	200	4.8	10.2	5.8	12.4	5.8	13.2	4.8	10.2	5.8	12.4	5.8	13.2
	200	100	4.6	10.4	4.8	8.8	4.4	10.6	4.6	10.4	4.8	8.8	4.4	10.6
	200	200	5.0	10.0	5.4	10.8	5.4	10.4	5.0	10.0	5.4	10.8	5.4	10.4
S2	100	100	6.2	11.6	5.2	9.6	4.8	10.4	6.2	11.6	5.2	9.6	4.8	10.4
	100	200	6.4	9.6	4.6	9.8	5.0	10.0	6.4	9.6	4.6	9.8	5.0	10.0
	200	100	6.8	13.2	5.4	10.6	6.6	10.6	6.8	13.2	5.4	10.6	6.6	10.6
	200	200	6.0	11.4	6.6	11.2	6.4	12.0	6.0	11.4	6.6	11.2	6.4	12.0
S3	100	100	4.4	10.0	5.6	10.8	4.6	10.6	4.4	10.0	5.6	10.8	4.6	10.6
	100	200	5.0	9.8	4.8	9.8	5.6	10.8	5.0	9.8	4.8	9.8	5.6	10.8
	200	100	4.8	11.0	6.8	12.2	7.2	13.0	4.8	11.0	6.8	12.2	7.2	13.0
	200	200	5.6	11.6	7.4	13.4	7.8	13.4	5.6	11.6	7.4	13.4	7.8	13.4
S4	100	100	6.8	11.0	5.2	12.0	5.6	10.0	6.8	11.0	5.2	12.0	5.6	10.0
	100	200	6.4	12.4	5.2	10.4	4.8	11.2	6.4	12.4	5.2	10.4	4.8	11.2
	200	100	6.2	13.4	6.0	12.0	5.8	11.6	6.2	13.4	6.0	12.0	5.8	11.6
	200	200	5.2	9.2	5.0	9.6	4.6	10.8	5.2	9.2	5.0	9.6	4.6	10.8

Note: (i) The results are obtained by setting $h = c(2.35/\sqrt{12})T^{-1/5}$ for $c = 0.5, 1$, and 1.5; (ii) R is the number of common factors.

Table A.2 The power of our test with different bandwidth sequences under DGP.P1-P8

DGP	N	T	R is fixed to the true value						R is determined from the data					
			$c = 0.5$		$c = 1$		$c = 1.5$		$c = 0.5$		$c = 1$		$c = 1.5$	
			5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
P1	100	100	47.8	60.4	72.2	81.4	81.4	87.8	46.0	60.8	72.8	80.8	81.2	87.6
	100	200	91.6	97.4	98.4	99.6	99.4	100	91.2	96.6	98.8	99.6	99.4	99.8
	200	100	73.0	84.4	94.0	97.2	98.0	98.8	72.8	83.0	94.2	97.2	98.0	98.8
	200	200	98.8	99.6	100	100	100	100	98.6	99.8	100	100	100	100
P2	100	100	52.2	62.0	29.4	41.4	10.0	18.0	51.2	61.6	29.8	41.6	9.6	17.0
	100	200	94.6	97.6	82.2	86.8	37.2	48.4	94.0	98.4	85.2	88.6	28.8	42.6
	200	100	66.4	77.6	41.0	51.8	12.2	22.0	65.4	77.6	37.8	50.0	12.8	19.8
	200	200	99.8	99.8	93.0	95.8	55.8	67.0	99.6	99.6	92.4	94.6	53.8	65.4
P3	100	100	30.6	41.8	37.2	47.8	49.2	60.0	29.0	44.0	36.0	41.0	51.0	60.0
	100	200	54.0	68.2	64.8	73.8	78.4	86.8	55.2	67.4	65.2	74.8	77.8	87.2
	200	100	28.6	38.6	42.4	53.8	71.2	79.4	29.4	40.0	42.4	52.0	71.6	78.4
	200	200	60.0	67.8	76.0	82.2	95.8	97.2	60.8	67.8	76.0	81.6	95.8	97.2
P4	100	100	59.4	71.8	25.0	38.0	11.0	19.4	60.0	72.4	25.4	35.8	11.2	19.0
	100	200	99.8	100	74.2	83.6	34.6	46.0	99.8	100	73.4	83.6	34.2	46.2
	200	100	82.6	88.6	40.6	52.8	16.2	24.6	81.0	88.6	40.0	52.6	15.6	24.4
	200	200	100	100	92.0	94.4	51.4	62.0	100	100	91.0	94.6	51.6	62.8
P5	100	100	42.8	55.0	67.8	79.8	79.6	86.2	42.0	54.0	69.0	80.2	78.4	87.0
	100	200	90.0	95.0	97.4	99.2	100	100	89.8	94.4	97.6	99.4	100	100
	200	100	69.0	78.6	90.0	94.2	95.6	97.4	69.4	79.4	90.4	94.8	94.8	97.0
	200	200	99.6	99.6	100	100	100	100	99.2	99.8	100	100	100	100
P6	100	100	48.8	59.8	29.6	38.6	11.6	19.8	48.2	60.6	27.8	39.8	11.0	19.4
	100	200	95.6	97.4	81.2	86.0	36.6	48.8	95.2	97.6	80.8	85.2	36.0	49.2
	200	100	67.8	78.6	38.4	52.6	15.2	22.2	69.2	79.8	40.4	52.6	15.8	23.0
	200	200	99.6	99.8	92.4	95.8	53.6	65.0	99.8	99.8	92.4	95.4	54.0	63.2
P7	100	100	29.4	38.2	34.0	45.8	50.8	62.6	30.4	38.0	33.2	46.0	49.0	62.8
	100	200	57.4	63.6	62.4	72.2	77.2	85.8	58.2	64.2	63.6	72.2	77.8	85.6
	200	100	32.4	42.2	44.0	53.0	69.2	77.2	31.8	41.8	44.8	52.2	70.2	77.6
	200	200	62.4	73.2	78.8	85.0	96.0	98.2	62.2	72.4	78.2	84.4	95.6	98.2
P8	100	100	64.4	76.0	38.2	50.6	11.8	19.2	65.0	76.8	38.6	51.2	12.0	20.6
	100	200	99.0	99.6	91.2	94.8	48.6	60.6	98.2	99.6	90.6	95.4	48.4	60.4
	200	100	78.6	85.4	49.4	60.8	15.4	23.0	78.2	86.2	49.4	60.8	14.4	23.8
	200	200	100	100	97.8	99.0	66.4	76.2	100	100	97.8	99.2	66.0	76.0

Note: See the note in Table A.1.

Table A.3 Size of tests under DGP.S1-S4 when the number of factors is determined by Su and Wang's (2017) IC

DGP	N	T	SM_B		$SW17$		HI_{LM}		CDG_{LM}		BE_{LM}	
			5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
S1	100	100	5.0	10.8	6.6	13.4	0.6	3.8	3.4	8.2	2.8	6.5
	100	200	5.8	12.4	7.4	13.0	2.4	6.8	4.8	7.4	3.4	7.5
	200	100	4.8	8.8	5.2	10.2	1.6	4.4	1.4	7.0	2.7	6.3
	200	200	5.4	10.8	5.8	12.0	1.6	6.8	3.6	8.8	3.4	7.5
S2	100	100	5.2	9.6	7.4	12.0	0.4	2.4	2.0	8.2	2.8	6.4
	100	200	4.6	9.8	5.0	11.4	1.0	5.8	2.0	6.6	3.7	7.8
	200	100	5.4	10.6	6.4	14.0	0.4	1.8	1.0	4.6	2.8	6.4
	200	200	6.6	11.2	7.0	14.0	0.6	5.4	2.6	6.8	3.6	7.7
S3	100	100	5.6	10.8	7.2	11.2	0.4	2.2	2.2	8.8	11.9	20.3
	100	200	4.8	9.8	6.0	11.4	1.6	5.2	2.0	6.0	15.3	24.7
	200	100	6.8	12.2	7.8	11.6	0.4	1.8	1.2	5.0	11.9	20.2
	200	200	7.4	13.4	8.2	13.0	0.8	5.2	2.4	7.0	15.3	24.8
S4	100	100	5.2	12.0	6.2	12.2	0.4	4.6	2.8	8.0	2.8	6.4
	100	200	5.2	10.4	4.2	10.4	2.0	6.8	4.6	8.6	3.4	7.5
	200	100	6.0	12.0	6.8	12.0	1.6	3.2	2.6	6.6	2.8	6.3
	200	200	5.0	9.6	5.6	10.2	2.2	7.0	4.0	8.4	3.4	7.4

Note: (i) SM_B denotes the results of our $\widehat{SM}_{NT}$ test using bootstrap critical values; (ii) $SW17$ denotes the results of Su and Wang's (2017) bootstrap-based test; (iii) HI_{LM} denotes Han and Inoue's (2014) sup-LM test; (iv) CDG_{LM} denotes Chen et al.'s (2014) sup-LM test; (v) BE_{LM} denotes Breitung and Eickmeier's (2011) N variable-specific sup-LM test. The main entries report the average percentage of rejection.

Table A.4 Power of tests under DGP.P1-P8 when the number of factors is determined by Su and Wang's (2017) IC

DGP	N	T	SM_B		$SW17$		HI_{LM}		CDG_{LM}		BE_{LM}	
			5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
P1	100	100	72.2	81.2	67.8	79.4	0.8	4.4	2.4	7.2	5.9	11.1
	100	200	98.4	99.6	98.4	99.4	4.2	10.6	2.0	6.8	11.2	17.8
	200	100	94.0	97.2	92.2	96.4	0.8	4.0	2.4	6.6	5.7	10.7
	200	200	100	100	100	100	5.0	12.2	2.2	6.6	11.1	17.5
P2	100	100	29.6	41.4	26.2	40.4	0.6	2.0	2.2	8.6	3.8	8.3
	100	200	82.6	87.0	77.2	84.2	1.6	6.4	2.2	6.4	6.7	12.7
	200	100	40.8	51.8	27.6	40.6	0.8	2.8	1.8	8.6	3.7	8.1
	200	200	93.0	95.8	85.2	91.6	1.6	5.8	1.8	7.6	6.5	12.4
P3	100	100	37.0	46.8	46.2	56.2	35.6	66.4	6.8	16.8	4.9	10.3
	100	200	65.0	74.2	76.8	86.4	97.4	99.8	10.2	18.4	9.8	17.2
	200	100	42.4	53.4	45.2	60.2	37.4	71.4	6.6	15.4	5.2	10.7
	200	200	76.0	82.2	84.2	92.0	99.2	100	10.2	20.0	9.8	17.7
P4	100	100	25.2	38.0	25.8	36.4	0.4	1.6	1.0	4.0	3.5	7.9
	100	200	74.0	83.6	72.2	81.4	0.6	4.0	3.0	5.6	5.4	10.6
	200	100	40.6	52.8	34.2	45.2	0.4	1.4	1.0	5.8	3.5	7.8
	200	200	92.0	94.4	86.8	92.8	0.2	3.8	3.2	6.4	5.5	10.7
P5	100	100	68.0	79.8	63.0	75.8	1.4	5.8	3.2	8.8	4.9	10.1
	100	200	97.4	99.2	96.8	99.0	6.0	12.8	4.4	8.4	9.8	16.6
	200	100	90.0	94.2	88.0	92.0	2.0	6.6	1.2	6.6	4.9	9.9
	200	200	100	100	99.6	99.8	3.8	11.4	4.8	10.6	9.4	15.8
P6	100	100	29.4	38.8	27.2	36.0	0.8	5.0	3.6	9.2	3.7	8.1
	100	200	81.0	85.8	75.8	82.6	3.2	10.4	5.6	10.8	6.2	12.1
	200	100	38.6	52.6	27.6	38.2	1.4	4.6	1.6	7.6	3.6	7.9
	200	200	92.4	95.8	85.2	90.8	3.0	9.8	4.6	11.0	6.2	11.9
P7	100	100	33.8	45.8	36.6	54.6	32.4	65.0	7.4	14.6	5.0	10.5
	100	200	62.6	72.2	74.4	86.2	98.2	99.6	12.0	18.0	9.5	16.9
	200	100	44.2	53.0	43.8	60.0	36.6	68.8	7.0	15.2	5.0	10.5
	200	200	78.6	85.0	86.4	92.6	99.0	99.8	10.8	19.6	9.7	17.5
P8	100	100	38.2	50.8	35.4	47.4	0.4	2.4	2.0	8.6	4.2	10.0
	100	200	91.0	94.8	88.4	92.2	1.4	6.0	2.2	6.6	10.5	18.0
	200	100	49.4	60.8	32.8	43.8	0.8	2.8	1.8	8.8	4.7	9.5
	200	200	97.8	99.0	93.2	95.4	1.6	5.8	2.0	7.6	9.3	15.9

Note: See the note in Table A.3.