Estimation and Inference on Time-Varying FAVAR Models *

Zhonghao Fu[a,b], Liangjun Su[c], Xia Wang[d]

[a] School of Economics, Fudan University, China

[b] Shanghai Institute of Finance and Economics, China

[c] School of Economics and Management, Tsinghua University, China

[d] School of Economics, Renmin University of China, China

Abstract

We introduce a time-varying (TV) factor-augmented vector autoregressive (FAVAR) model to capture the TV behavior in the factor loadings and the VAR coefficients. To consistently estimate the TV parameters, we first obtain the unobserved common factors via the local principal component analysis (PCA) and then estimate the TV-FAVAR model via a local smoothing approach. The limiting distribution of the proposed estimators is established. To gauge possible sources of TV features in the FAVAR model, we propose three L^2-distance-based test statistics and study their asymptotic properties under the null and local alternatives. Simulation studies demonstrate the excellent finite sample performance of the proposed estimators and tests. In an empirical application to the U.S. macroeconomic dataset, we document overwhelming evidence of structural changes in the FAVAR model and show that the TV-FAVAR model outperforms the conventional time-invariant FAVAR model in predicting certain key macroeconomic series.

*Fu acknowledges financial support from the National Science Foundation of China (NSFC) (No.71903032). Su gratefully acknowledges the support from NSFC (No. 72133002). Wang acknowledges financial supports from NSFC (No.71873151, No.71401160). Please address correspondence to: Xia Wang, School of Economics, Renmin University of China, Beijing, 100872, China. Email: *wxia@ruc.edu.cn*.

JEL Classification: C12, C14, C22.

Key Words: FAVAR, Local smoothing, Specification test, Structural change, Time-varying modelling

1 Introduction

Factor-augmented vector autoregressive (FAVAR) models have drawn increasing attention in the macroeconomic literature. As pointed out by Sims (1992) and Christiano et al. (1999), there is a dilemma between incorporating sufficiently large information and controlling the degree of freedom of a standard VAR model, as the number of parameters increases rapidly with the number of variables. By introducing the unobservable common factors into the VAR structure, the FAVAR model can summarize large-dimensional information and achieve dimension reduction. Empirical studies show that one can improve the economic prediction and get a more reasonable interpretation of the economic relationship by incorporating the latent common factors into regressions (e.g., Stock and Watson, 2002).

The FAVAR models were initially proposed by Bernanke et al. (2005) to identify the monetary transmission mechanism. Bai and Ng (2006) provide the asymptotic theory for factor-augmented regressions. Recently, Bai et al. (2016) studied the identification restrictions and proposed a likelihood-based two-step approach to estimate the FAVAR model. The literature mentioned above establishes a substantial theoretical foundation for the FAVAR model. However, most of the existing literature on the FAVAR model assumes that both the factor loadings and the VAR coefficients are time-invariant. In fact, since the FAVAR model is widely used in macroeconomic analysis and the datasets usually have a long time span, it is unsuitable to assume that the factor loadings and the VAR coefficients are time-invariant. Driving forces such as economic transition and technological progress could significantly influence the relationship among economic and financial variables.

Structural change has drawn much attention in the literature. Examples in the linear regression models include Bai and Perron (1998) and Qu and Perron (2007). Besides, there also exists extensive literature on testing for structural changes in factor models; see, e.g., Breitung and Eickmeier (2011), Chen et al. (2014), Corradi and Swanson (2014), Han and Inoue (2015), Su and Wang (2017, 2020a) and Baltagi et al. (2021). Despite the vast literature on structural changes in factor models, little attention has been paid to the instability of FAVAR models. Eickmeier et al. (2015) introduce a TV-FAVAR model by modeling the TV factor loadings and coefficients as random walk processes. Li et al. (2020) introduce a functional-coefficient predictive model with latent factor regressors. Wei and Zhang (2020) propose a TV diffusion index model. Yan and Cheng (2022) also consider a factor-augmented predictive regression and introduce a threshold structure to capture parameter instability. Cai and Liu (2021) allow the coefficients of the FAVAR models to vary with certain state variables. Most of these papers assume the factor loadings are time-invariant and only impose instability on the VAR coefficients. Moreover, they only consider the estimation problem and do not establish the limiting distribution of the estimated coefficients.

In this paper, we propose a novel TV-FAVAR model that allows the unknown factor loadings and VAR coefficients to change smoothly over time. We suggest a two-stage procedure to estimate the model. In the first stage, we estimate a TV factor model by the local principal component analysis (PCA) proposed by Su and Wang (2017, 2020a) to obtain consistent estimates for the common factors. In the second stage, we augment a VAR model using the estimated common factors and then estimate the TV VAR coefficients by a local smoothing approach. We further establish the limiting distributions of the estimated VAR coefficients under the standard large N and large T setting. Besides, we propose three test statistics to gauge the possible sources of TV features in the FAVAR model. Let $\{X_{it}\}$ be the dataset to obtain the estimates of the common factors and factor loadings, and $\{Y_t\}$ be the vectors used in the VAR model. To construct our tests, we first estimate a time-invariant FAVAR to collect the estimates for the time-invariant factor loadings and VAR coefficients and run two additional regressions using the local smoothing approach. In the first regression, we regress $\{X_{it}\}$ on the estimated common factors via local smoothing to obtain the estimates for TV factor loadings. In the second regression, we replace the unobserved common factors with the estimated common factors and estimate the VAR coefficients via local smoothing. The test statistics are then constructed by measuring the L^2-distances between these two sets of estimators. In an empirical application, we use the proposed TV-FAVAR model to check whether the U.S. economy suffers from structural changes and evaluate the forecasting performance of our approach for some key variables. We find strong evidence of structural changes in both the factor structure and the VAR dynamics and show that the TV-FAVAR model delivers a superior forecasting performance than Bai and Ng's (2006) conventional time-invariant FAVAR model.

The rest of this paper is organized as follows. In Section 2, we introduce the TV-FAVAR model. In Section 3, we propose a two-stage estimation procedure to estimate the TV-FAVAR model and study the asymptotic properties of the estimators. In Section 4, we construct three test statistics to test for the TV factor loadings and/or VAR coefficients and study the asymptotic null distributions and asymptotic local power properties. Section 5 studies the finite sample performance of our estimators and tests via simulations, and Section 6 provides empirical applications. Section 7 concludes. All proofs are contained in the online supplement.

NOTATION. For a 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 $\|A\|_{\mathrm{sp}} \equiv [\mu_1(A'A)]^{1/2}$. $\mu_s(\cdot)$ and $\mu_{\min}(\cdot)$ denote the sth largest and the smallest eigenvalues of a real symmetric matrix, respectively. For an $m \times n$ functional matrix $A(\tau) = [A_{ij}(\tau)]_{i=1,...,m;j=1,...,n}$, we denote $\mathrm{d}^c A(\tau) \equiv [\mathrm{d}^c A_{ij}(\tau)/(\mathrm{d}\tau)^c]$ as the c-th order element-wise derivative of $A(\tau)$. We use $B > 0$ to denote that B is positive definite. Let $P_A \equiv A(A'A)^+ A'$ and $M_A \equiv \mathbb{I}_m - P_A$, where $\mathbb{I}_m$ denotes an $m \times m$ identity matrix, and B^+ denotes the generalized inverse of a square matrix B. For a positive integer N, we let $[N] \equiv \{1, 2, ..., N\}$. For a positive number a, $\lfloor a \rfloor$ denotes the integer part of a. The operators "$\xrightarrow{P}$", "$\xrightarrow{d}$", and "plim" denote convergence in probability, convergence in distribution, and probability limit, respectively. We use $(N, T) \to \infty$ to denote that N and T pass to infinity jointly. Let $C < \infty$ denote a positive constant that may vary from case to case.

2 Time-Varying Factor-Augmented VAR Model

2.1 The model

Let $\{X_{it}, i \in [N]; t = -p+1, ..., T\}$ be an N-dimensional time series with $T + p$ observations, where p is the lag order of the FAVAR model. For notational simplicity, we assume that the total number of time series observations is $T + p$. We follow the lead of Su and Wang (2017) to assume that X_{it} admits a TV factor structure with R common factors $F_t = (F_{1t}, \cdots, F_{Rt})'$:

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

where $\{e_{it}\}$ can be weakly dependent across both the cross-sectional units i and time periods t. Let $\{Y_t, \, t = -p+1, ..., T\}$ be a K-dimensional random vector with $T + p$ observations. We assume that $(Y_t', F_t')'$ is generated via the following TV VAR(p) process:

$$\begin{pmatrix} Y_t \\ F_t \end{pmatrix} = \sum_{j=1}^{p} \begin{pmatrix} \phi_{jt}^{(1,1)} & \phi_{jt}^{(1,2)} \\ \phi_{jt}^{(2,1)} & \phi_{jt}^{(2,2)} \end{pmatrix} \begin{pmatrix} Y_{t-j} \\ F_{t-j} \end{pmatrix} + \begin{pmatrix} \varepsilon_{Y,t} \\ \varepsilon_{F,t} \end{pmatrix} \equiv \sum_{j=1}^{p} \phi_{jt} \begin{pmatrix} Y_{t-j} \\ F_{t-j} \end{pmatrix} + \varepsilon_t, \tag{2.2}$$

where $\phi_{jt} \equiv \left(\phi_{jt}^{(1,1)}, \phi_{jt}^{(1,2)}; \phi_{jt}^{(2,1)}, \phi_{jt}^{(2,2)}\right)$ is a $(K+R) \times (K+R)$ matrix and $\varepsilon_t = \left(\varepsilon_{Y,t}', \varepsilon_{F,t}'\right)'$ is a $(K+R) \times 1$ error vector. In particular, (2.2) implies that:

$$Y_t = \sum_{j=1}^{p} \left(\phi_{jt}^{(1,1)} Y_{t-j} + \phi_{jt}^{(1,2)} F_{t-j}\right) + \varepsilon_{Y,t}. \tag{2.3}$$

It can be regarded as an extension of Stock and Watson's (2002) diffusion index model by allowing for structural changes in the regression coefficient matrices.

To consistently estimate the proposed TV-FAVAR model, we follow Robinson (1989, 1991) to specify the TV factor loadings and the TV VAR coefficients as deterministic functions of the rescaled time index t/T:

$$\lambda_{it} \equiv \lambda_i(t/T) \text{ and } \phi_{jt} \equiv \phi_j(t/T), \tag{2.4}$$

where $\lambda_i(\cdot)$ and $\phi_j(\cdot)$ are unknown smooth functions on $(0, 1]$ for each i and j. Such a specification is widely adopted by the nonparametric TV models; see Cai (2007), Chen and Hong (2012), and Su and Wang (2017), among many others.

Model (2.1) is the factor model with TV factor loadings considered by Su and Wang (2017, 2020b). As is well known, λ_{it} and F_t are not separately identifiable. Let $\Lambda_t = (\lambda_{1t}, \cdots, \lambda_{Nt})'$ and $F = (F_{-p+1}, \cdots, F_T)'$. We follow Bai and Ng (2002, 2006) and Bai (2003) to impose the identification conditions that $(T+p)^{-1} F'F = \mathbb{I}_R$ and $N^{-1}\Lambda_t'\Lambda_t$ is diagonal with descending diagonal elements. In addition, Su and Wang (2017) have studied consistent determination of R based on local PCA estimates. Therefore, we assume that R is known in this paper.

Remark 1. There is a growing literature on TV factor loadings that specifies the TV factor loadings as a random walk process or a VAR process; 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. (2019). Similarly, Mumtaz and Surico (2012) consider TV coefficients in the factor process. Most of these papers estimate the unknown parameters using the Bayesian approach. The interpretations and implications behind the stochastic and deterministic specifications for the TV parameters are inherently distinct. Cogley and Sargent (2001) point out that the stochastic fluctuations in the parameters of a reduced-form economic system may result from evolving beliefs of the policymaker. In contrast, the deterministic time-varying coefficients arise when structural changes exist. In this paper, we aim to incorporate structural changes into the FAVAR framework. Hence, we specify the TV factor loadings and VAR coefficients as deterministic functions of t/T and apply the nonparametric kernel method to estimate the TV parameters. It is worth mentioning that the kernel method can also be used to estimate parameters associated with stochastic time variation; see Giraitis et al. (2014) and Giraitis et al. (2021) for the kernel estimation of TV coefficients under stochastic specifications without and with endogeneity, respectively. We note that most of the existing studies specify the TV coefficients as either a stochastic or deterministic form without any formal justification. We conjecture that it is possible to follow the lead of Fu et al. (2022) and propose some formal tests to distinguish these two forms in the factor literature.

Remark 2. The factor model in (2.1) is a TV version of the static factor model studied by Bai and Ng (2002) and Bai (2003), where the impact of a factor occurs only contemporaneously in all series. Alternatively, one can follow the lead of Barigozzi et al. (2021) and consider a TV generalized dynamic factor model (GDFM). The latter paper extends the GDFM of Forni et al. (2000, 2005) and Forni et al. (2017) and proves the consistency of their estimators of TV impulse response functions. Nevertheless, the estimation procedures for these two models are different, and it is quite challenging to derive a consistent estimator for the common shocks (primitive factors) in a TV GDFM. Hence, it is difficult to extend our estimation and testing procedure to the TV GDFM, and we leave it for future research.

2.2 Further Representations

Since F_t is unobservable, we replace F_t with a generic estimator $\check{F}_t$, which is consistent up to a TV rotation matrix $H_t = H\left(t/T\right)$. Then, (2.2) implies that

$$Y_t = \sum_{j=1}^{p} \phi_{jt}^{(1,1)} Y_{t-j} + \sum_{j=1}^{p} \phi_{jt}^{(1,2)} H_{t-j}'^{-1} \check{F}_{t-j} - \sum_{j=1}^{p} \phi_{jt}^{(1,2)} H_{t-j}'^{-1} \left(\check{F}_{t-j} - H_{t-j}' F_{t-j}\right) + \varepsilon_{Y,t},$$

$$\check{F}_t = \sum_{j=1}^{p} H_t' \phi_{jt}^{(2,1)} Y_{t-j} + \sum_{j=1}^{p} H_t' \phi_{jt}^{(2,2)} H_{t-j}'^{-1} \check{F}_{t-j} + \left(\check{F}_t - H_t' F_t\right) - \sum_{j=1}^{p} H_t' \phi_{jt}^{(2,2)} H_{t-j}'^{-1} \left(\check{F}_{t-j} - H_{t-j}' F_{t-j}\right)$$
$$+ H_t' \varepsilon_{F,t}.$$

Denote $\psi_{jt}^{(1,1)} = \phi_{jt}^{(1,1)}$, $\psi_{jt}^{(1,2)} = \phi_{jt}^{(1,2)} H_{t-j}'^{-1}$, $\psi_{jt}^{(2,1)} = H_t' \phi_{jt}^{(2,1)}$, and $\psi_{jt}^{(2,2)} = H_t' \phi_{jt}^{(2,2)} H_{t-j}'^{-1}$. Then, (2.2) can be written as:

$$\begin{pmatrix} Y_t \\ \check{F}_t \end{pmatrix} = \sum_{j=1}^{p} \begin{pmatrix} \psi_{jt}^{(1,1)} & \psi_{jt}^{(1,2)} \\ \psi_{jt}^{(2,1)} & \psi_{jt}^{(2,2)} \end{pmatrix} \begin{pmatrix} Y_{t-j} \\ \check{F}_{t-j} \end{pmatrix} + \begin{pmatrix} v_{Y,t} \\ v_{F,t} \end{pmatrix} \equiv \sum_{j=1}^{p} \psi_{jt} \begin{pmatrix} Y_{t-j} \\ \check{F}_{t-j} \end{pmatrix} + v_t, \qquad (2.5)$$

where $\psi_{jt} \equiv \left(\psi_{jt}^{(1,1)}, \psi_{jt}^{(1,2)}; \psi_{jt}^{(2,1)}, \psi_{jt}^{(2,2)}\right)$, $v_{Y,t} = -\sum_{j=1}^{p} \psi_{jt}^{(1,2)} (\check{F}_{t-j} - H_{t-j}' F_{t-j}) + \varepsilon_{Y,t}$ and $v_{F,t} = (\check{F}_t - H_t' F_t) - \sum_{j=1}^{p} \psi_{jt}^{(2,2)} \left(\check{F}_{t-j} - H_{t-j}' F_{t-j}\right) + H_t' \varepsilon_{F,t}$.

Comparing (2.5) with (2.2), we note that $\phi_{jt}^{(a,b)} \neq \psi_{jt}^{(a,b)}$ for $j \in [p]$, except for $(a,b) = (1,1)$. Specifically, if the factor loadings in (2.1) are TV, $\{\check{F}_t\}$ are only consistent estimators for the latent factors pre-multiplied by H_t. Since H_t is incorporated in the representation of $\psi_{jt}^{(a,b)}$ for $(a,b) = (1,2), (2,1), (2,2)$, their estimates always exhibit TV behavior no matter whether $\phi_{jt}^{(a,b)}$ are TV or not. Hence, it is inappropriate to restrict the regression coefficients $\psi_{jt}^{(a,b)}$ for $(a,b) = (1,2), (2,1), (2,2)$ to be constant if one adopts a TV factor model.

Note that the smoothness of $\{\lambda_i(\cdot)\}_{i \in [N]}$ implies that of $H(\cdot)$. With (2.4), we can write $\psi_{jt} = \psi_j(t/T)$, where the functional form of $\psi_j(\cdot)$ depends on both $\phi_j(\cdot)$ and $\{\lambda_i(\cdot)\}_{i \in [N]}$. As a result, $\psi_j(\cdot)$ is second-order continuously differentiable provided both $\phi_j(\cdot)$ and $\{\lambda_i(\cdot)\}_{i \in [N]}$ are. Therefore, for each fixed $r \in \{-p+1, ..., T\}$, we further have

$$Y_t = \sum_{j=1}^{p} \psi_{jr}^{(1,1)} Y_{t-j} + \sum_{j=1}^{p} \psi_{jr}^{(1,2)} \check{F}_{t-j} + \Delta_Y(t,r) + \varepsilon_{Y,t}, \text{ and}$$
$$\check{F}_t = \sum_{j=1}^{p} \psi_{jr}^{(2,1)} Y_{t-j} + \sum_{j=1}^{p} \psi_{jr}^{(2,2)} \check{F}_{t-j} + \Delta_F(t,r) + H_t' \varepsilon_{F,t},$$

where

$$\Delta_Y(t,r) = \sum_{j=1}^{p} \left[\psi_{jt}^{(1,2)} (H_{t-j}' F_{t-j} - \check{F}_{t-j}) + (\psi_{jt}^{(1,2)} - \psi_{jr}^{(1,2)}) \check{F}_{t-j} + (\psi_{jt}^{(1,1)} - \psi_{jr}^{(1,1)}) Y_{t-j}\right], \text{ and}$$

$$\Delta_F(t,r) \;=\; (\check{F}_t - H_t'F_t) + \sum_{j=1}^{p}\left[(\psi_{jt}^{(2,1)} - \psi_{jr}^{(2,1)})Y_{t-j} + (\psi_{jt}^{(2,2)} - \psi_{jr}^{(2,2)})\check{F}_{t-j} + \psi_{jt}^{(2,2)}(H_{t-j}'F_{t-j} - \check{F}_{t-j})\right].$$

Then, we have

$$\begin{pmatrix} Y_t \\ \check{F}_t \end{pmatrix} = \sum_{j=1}^{p}\psi_{jr}\begin{pmatrix} Y_{t-j} \\ \check{F}_{t-j} \end{pmatrix} + U_t^{(r)}, \tag{2.6}$$

where $U_t^{(r)} \equiv (u_{Y,t}^{(r)\prime}, u_{F,t}^{(r)\prime})'$, $u_{Y,t}^{(r)} \equiv \Delta_Y(t,r) + \varepsilon_{Y,t}$ and $u_{F,t}^{(r)} \equiv \Delta_F(t,r) + H_t'\varepsilon_{F,t}$.

3 Estimation

In this section, we propose a two-stage method to estimate the TV-FAVAR model and then establish the asymptotic distributions.

3.1 Two-Stage Estimation Procedure

To consistently estimate the parameters in the TV-FAVAR model described by (2.1) and (2.2), we propose a two-stage estimation procedure. In the first stage, we estimate the TV factor model in (2.1) by the local PCA approach of Su and Wang (2017), while in the second stage, we replace the latent common factor F_t in (2.2) with the estimator $\hat{F}_t$ obtained in the first stage and estimate the TV VAR coefficients via a local smoothing procedure.

3.1.1 Stage 1: Estimating the TV Factor Model

Let $k_{h_1,tr} \equiv h_1^{-1}K(\frac{t-r}{Th_1})$, where $K:\mathbb{R}\to\mathbb{R}^+$ denotes a kernel function and $h_1 \equiv h_{1NT}$ is a bandwidth. For each fixed $r \in \{-p+1, ..., T\}$, we construct a $(T+p)\times N$ matrix $X^{(r)} = (X_1^{(r)}, \cdots, X_N^{(r)})$ with $X_i^{(r)} \equiv (k_{h_1,(-p+1)r}^{1/2}X_{i,-p+1}, \cdots, k_{h_1,Tr}^{1/2}X_{i,T})'$ and a $(T+p)\times R$ matrix of kernel weighted common factors $F^{(r)} = (k_{h_1,(-p+1)r}^{1/2}F_{-p+1}, \cdots, k_{h_1,Tr}^{1/2}F_T)'$. According to Su and Wang (2017), we can consistently estimate the TV factor loadings and the kernel weighted common factors by solving the following constrained minimization problem:

$$\min_{F^{(r)},\Lambda^{(r)}} \operatorname{tr}\left[\left(X^{(r)} - F^{(r)}\Lambda_r'\right)\left(X^{(r)} - F^{(r)}\Lambda_r'\right)'\right]$$

s.t. $(T+p)^{-1}F^{(r)\prime}F^{(r)} \;=\; \mathbb{I}_R$ and $N^{-1}\Lambda_r'\Lambda_r$ is diagonal with descending diagonal elements.

The estimated factor matrix, denoted by $\hat{F}^{(r)} = (\hat{F}^{(r)}_{-p+1}, ..., \hat{F}^{(r)}_T)'$, is $\sqrt{T+p}$ times eigenvectors corresponding to the R largest eigenvalues of $X^{(r)}X^{(r)\prime}$, arranged in descending order, and $\hat{\Lambda}_r' = (\hat{F}^{(r)}\hat{F}^{(r)\prime})^{-1}\hat{F}^{(r)\prime}X^{(r)} = (T+p)^{-1}\hat{F}^{(r)\prime}X^{(r)}$ is the estimator of the corresponding TV factor loading matrix. We note that $\hat{F}^{(r)}$ is only consistent for a rotational version of $F^{(r)}$. To obtain a consistent estimator for the original common factor, we regress $\{X_{it}, i = 1, \cdots, N\}$ on $\{\hat{\lambda}_{it}, i = 1, \cdots, N\}$ for each t to get the updated estimator $\hat{F}_t = \left(\sum_{i=1}^N \hat{\lambda}_{it}\hat{\lambda}_{it}'\right)^{-1}\left(\sum_{i=1}^N \hat{\lambda}_{it}X_{it}\right)$ for $t = -p+1, ..., T$, where $\hat{\lambda}_{it}$ is the ith column of $\hat{\Lambda}_t'$, for $i \in [N]$. As shown by Su and Wang (2017, 2020b), the estimated common factor $\hat{F}_t$ is a consistent estimator for the latent factor F_t up to a TV rotation matrix $H_t \equiv H(t/T)$.

3.1.2 Stage 2: Estimating the TV VAR Coefficients

Now, we replace the generic estimator $\check{F}_t$ in (2.5) and (2.6) with the estimator $\hat{F}_t$ obtained in Stage 1. Obviously, (2.6) is a standard VAR(p) model for each fixed $r \in \{-p+1, ..., T\}$. Let $W_t \equiv (Y_t', F_t'H_t)'$ and $\hat{W}_t = (Y_t', \hat{F}_t')'$, both of which are $(K+R) \times 1$. Let $Z_t \equiv (W_{t-1}', ..., W_{t-p}')'$ and $\hat{Z}_t \equiv (\hat{W}_{t-1}', ..., \hat{W}_{t-p}')'$ be $(K+R)p \times 1$ vectors and let $\Psi_t \equiv (\psi_{1t}, ..., \psi_{pt})'$ be a $(K+R)p \times (K+R)$ matrix. For $r \in [T]$, we consider the local least squares (LLS) problem $\min_{\Psi_r} \sum_{t=1}^T \left[\hat{W}_t - \Psi_r'\hat{Z}_t\right]'\left[\hat{W}_t - \Psi_r'\hat{Z}_t\right] k_{h_2,tr}$, where $k_{h_2,tr} \equiv h_2^{-1}K(\frac{t-r}{Th_2})$, and the bandwidth $h_2 \equiv h_{2T}$. The solution to this problem is: $\hat{\Psi}_r \equiv \hat{\Psi}\left(\frac{r}{T}\right) = \left(\sum_{s=1}^T k_{h_2,sr}\hat{Z}_s\hat{Z}_s'\right)^{-1}\left(\sum_{s=1}^T k_{h_2,sr}\hat{Z}_s\hat{W}_s'\right)$.

3.2 Limiting Distribution

In this subsection, we establish the asymptotic distribution of $\hat{\Psi}_r$. Note that (2.1) is a purely TV factor model, and the asymptotic distributions of the estimated common factors and factor loadings have been explicitly analyzed by Su and Wang (2017, 2020b).

Let $e_t \equiv (e_{1t}, ..., e_{Nt})'$, $e \equiv (e_{-p+1}, ..., e_T)'$, $e_i^{(r)} \equiv (k^{1/2}_{h_1,(-p+1)r}e_{i(-p+1)}, \cdots, k^{1/2}_{h_1,Tr}e_{iT})'$, $e^{(r)} \equiv (e_1^{(r)}, \cdots, e_N^{(r)})$, $\gamma_N(s,t) \equiv N^{-1}E(e_s'e_t)$, $\gamma_{N,F}(s,t) \equiv N^{-1}E(F_se_s'e_t)$, $\gamma_{N,FF}(s,t) \equiv N^{-1}E(F_se_s'e_tF_t')$, and $\zeta_{st} \equiv N^{-1}[e_s'e_t - E(e_s'e_t)]$. Define $\varpi_{NT,1}(r) = \frac{h_1^{1/2}}{\sqrt{NT}}F^{(r)\prime}e^{(r)}\Lambda_r$ and $\varpi_{NT,2}(r,t) = \frac{h_1^{1/2}}{\sqrt{NT}}[F^{(r)\prime}e^{(r)}e_t - E(F^{(r)\prime}e^{(r)}e_t)]$. We further let $W_t^0 \equiv (Y_t', F_t')'$, $Z_t^0 \equiv \left(W_{t-1}', ..., W_{t-p}'\right)'$, $\Sigma_{ZZ,t} \equiv E(Z_t^0Z_t^{0\prime})$,

$\Sigma_{ZY,jt} \equiv E(Z_t^0 Y_{t-j}')$, and $\Sigma_{ZF,jt} \equiv E(Z_t^0 F_{t-j}')$. We make the following assumptions.

Assumption A.1 (i) $E(e_{it}) = 0$, $\max_{i,t} E|e_{it}|^{8+\sigma} < \infty$ for some $\sigma > 0$, and $\|e\|_{\text{sp}} = O_P\left(N^{1/2} + T^{1/2}\right)$; (ii) $\max_t E\|F_t\|^{8+\sigma} \leq C$, $E\left(F_t F_t'\right) = \Sigma_F > 0$ for some $R \times R$ positive definite matrix Σ_F; (iii) $\lambda_{it} = \lambda_i(\cdot)$ is nonrandom and third-order continuously differentiable with respect to $\tau \equiv t/T \in [0,1]$ such that $\max_{i\in[N]} \sup_{\tau\in[0,1]} \|\mathrm{d}^c \lambda_i(\tau)\| \leq C$ for $c = 0,1,2,3$. $N^{-1}\Lambda_r'\Lambda_r = \Sigma_{\Lambda_r} + O\left(N^{-1/2}\right)$ for some $R \times R$ positive definite matrix Σ_{Λ_r} and for all r; (iv) $\max_t \sum_{s=1}^{T} |\mathrm{Cov}\left(F_{mt}F_{nt}, F_{ms}F_{ns}\right)| \leq C$ for $m,\ n = 1,...,R$; (v) $\max_t \sum_{s=1}^{T} \|\gamma(s,t)\| \leq C$ and $\max_s \sum_{t=1}^{T} \|\gamma(s,t)\| \leq C$ for $\gamma = \gamma_N,\ \gamma_{N,F}$, and $\gamma_{N,FF}$; (vi) $\max_{s,t} E\left|N^{1/2}\zeta_{st}\right|^4 \leq C$ and $\max_{r,t} E\left\|N^{-1/2}\Lambda_r' e_t\right\|^4 \leq C$; (vii) $\varpi_{NT,1}(r) = O_P(1)$ and $\max_t E\left\|\varpi_{NT,2}(r,t)\right\|^2 \leq C$ for each r; (viii) For each $r \in \{-p+1,...,T\}$, the eigenvalues of $\Sigma_{\Lambda_r}^{1/2}\Sigma_F\Sigma_{\Lambda_r}^{1/2}$ are distinct.

Assumption A.2 (i) The kernel function $K : \mathbb{R} \to \mathbb{R}^+$ is a symmetric and continuously differentiable PDF with compact support $[-1,1]$; (ii) As $(N,T) \to \infty$, $h_1 \to 0$, $Th_1^2 \to \infty$, $Nh_1^2 \to \infty$, $Th_1/N \to 0$, and $Th_1/N^{1/2} \to \infty$; (iii) As $(N,T) \to \infty$, $h_2 \to 0$, $Th_2^2 \to \infty$, $T(\ln T)\, h_2/N \to 0, Th_2h_1^4 \to 0$, and $Th_2^7 \to 0$.

Assumption A.3 (i) For each $j \in [p]$, $\phi_j(\cdot)$ is nonrandom and third-order continuously differentiable with $\max_{j,t} \|\mathrm{d}^c\phi_j(t/T)\| \leq C$ for $c = 0,1,2,3$; (ii) $\max_t E\|\varepsilon_t\|^{8+\sigma} \leq C$ and $\max_t E\|Z_t^0\|^{8+\sigma} \leq C$; (iii) $\min_t \mu_{\min}(\Sigma_{ZZ,t}) \geq c_Z$ for some fixed constant $c_Z > 0$; (iv) $E(\varepsilon_t | W_t^0, \varepsilon_{t-1}, W_{t-1}^0, \varepsilon_{t-2}, ...) = 0$.

Assumption A.1 slightly strengthens Assumption A.1 in Su and Wang (2017). We require the existence of eight-plus moments for e_{it} and F_t to facilitate the analysis of the specification tests in the next section. The condition on $\|e\|_{\text{sp}}$ in Assumption A.1(i) has been widely assumed in the literature; see, e.g., Moon and Weidner (2015, 2017), Li et al. (2016), and Lu and Su (2016). Assumption A.1(i) allows e_{it} to be both unconditionally heteroskedastic over (i,t) and conditionally heteroskedastic given F_t. Assumption A.1(ii) imposes some moment conditions on the common factors, which appears more restrictive than those in the existing literature (e.g., Eichler et al., 2011; Motta et al., 2011). The condition that $\max_t E\left\|F_t\right\|^{8+\sigma} \leq C$ can be relaxed to $\max_t E\left\|F_t\right\|^{4+\sigma} \leq C$ if one only cares about the estimation but not the testing under our weak dependence conditions below. If we follow some existing literature such as Motta et al. (2011) to impose serial independence on $\{F_t\}$, our moment condition on $\{F_t\}$ can also be weakened to $\max_t E\left\|F_t\right\|^4 \le C$. Assumption A.1(iii) imposes smoothness conditions on λ_{it} explicitly. Assumption A.2 specifies conditions on the kernel function and bandwidths, the first two parts of which are the same as Assumption A.3(i) and (ii) in Su and Wang (2017).

Assumption A.3 relates to the VAR part of the TV-FAVAR model. Assumption A.3(i) requires that the TV VAR coefficients be element-wise third-order continuously differentiable. This condition can be weakened to the second-order continuous differentiability by strengthening the requirements on the bandwidth parameters from $Th_2^7 \to 0$ to $Th_2^5 \to 0$ in Assumption A.2(iii). Assumption A.3(ii) provides moment conditions on ε_t and Z_t^0, and Assumption A.3(iii) assumes that the matrix $\Sigma_{ZZ,t}$ is positive definite for each t. A.3(iv) requires that ε_t satisfies a martingale-difference-type condition in the FAVAR model, which helps us to establish the asymptotic normality of the estimator $\hat{\Psi}_t$. Assumptions A.3 (iii) and (iv) are quite similar to the first part of Assumption E in Bai and Ng (2006).

To derive the limiting distribution of $\hat{\Psi}_t$, we add some notations. Let $\kappa_2 \equiv \int_{-1}^{1} \tau^2 K(\tau) d\tau$, $D_{Q_t^{-1}} \equiv \text{diag}(\mathbb{I}_K, Q_t^{-1})$, and $\mathbb{Q}_t \equiv \text{diag}(\mathbb{I}_K, Q_{t-1}^{\prime -1}, \cdots, \mathbb{I}_K, Q_{t-p}^{\prime -1})$, where $Q_t = V_t^{1/2} \Upsilon_t^{-1} \Sigma_{\Lambda_t}^{-1/2}$ with V_t being the diagonal matrix consisting of the eigenvalues of $\Sigma_{\Lambda_t}^{1/2} \Sigma_F \Sigma_{\Lambda_t}^{1/2}$ in descending order and Υ_t being the corresponding (normalized) eigenvector matrix.

Theorem 3.1. *[Asymptotic distribution of $\hat{\Psi}_t$] Suppose that Assumptions A.1-A.3 hold. Then, for any fixed $t \in \{\lfloor Th_2 \rfloor, ..., T - \lfloor Th_2 \rfloor\}$,*

$$\sqrt{Th_2}\text{vec}\left(\hat{\Psi}_t - \Psi_t - h_2^2 \mathbf{B}_t\right) \xrightarrow{d} N\left(0, \left[\mathbb{I}_{K+R} \otimes \left(\mathbb{Q}_t \Sigma_{ZZ,t} \mathbb{Q}_t'\right)^{-1}\right] \Sigma_{Z\otimes\varepsilon}^{(t)} \left[\mathbb{I}_{K+R} \otimes \left(\mathbb{Q}_t \Sigma_{ZZ,t} \mathbb{Q}_t'\right)^{-1}\right]\right),$$

where $\mathbf{B}_t = \left(\mathbf{B}_{Y,t}', \mathbf{B}_{F,t}'\right)'$, $\mathbf{B}_{Y,t} = \kappa_2 \left(\mathbb{Q}_t \Sigma_{ZZ,t} \mathbb{Q}_t'\right)^{-1} \left(\sum_{j=1}^{p} B_{t,j}^Y\right)$, $\mathbf{B}_{F,t} = \kappa_2 \left(\mathbb{Q}_t \Sigma_{ZZ,t} \mathbb{Q}_t'\right)^{-1} \left(\sum_{j=1}^{p} B_{t,j}^F\right)$, and $\Sigma_{Z\otimes\varepsilon}^{(t)} = \lim_{T\to\infty} \frac{1}{Th_2} \sum_{s=1}^{T} K^2\left(\frac{s-t}{Th_2}\right) E\left\{\left[D_{Q_s^{-1}}' \varepsilon_s \varepsilon_s' D_{Q_s^{-1}}\right] \otimes \left[\mathbb{Q}_s Z_s^0 Z_s^{0\prime} \mathbb{Q}_s'\right]\right\}$ with $B_{t,j}^Y$ and $B_{t,j}^F$ being respectively defined in (S1.1) and (S1.2) in the online supplement.

From the definitions of $B_{t,j}^Y$ and $B_{t,j}^F$ in (S1.1) and (S1.2), we note that the asymptotic bias of $\hat{\Psi}_t$ is fairly complicated. It mainly arises from the error in the first stage estimation of F_t. Replacing F_t by $\hat{F}_t$ in the FAVAR will bring in the estimation error $\hat{F}_t - H_t' F_t$, which is $O_P\left((N/\ln T)^{-1/2} + h_1^2\right)$ uniformly in t. Such an error term does not contribute to the asymptotic distribution of $\hat{\Psi}_t$ under the condition that $\sqrt{Th_2}((N/\ln T)^{-1/2} + h_1^2) = o(1)$, which is ensured by Assumption A.2(iii). Nevertheless, H_t is random and depends on the whole sample $\{X_{it}\}$ and we have to replace it by its probability limit Q_t^{-1} in the asymptotic analysis. When λ_{it} is TV, so is $Q_t \equiv Q(t/T)$. This explains why several terms associated with the first- or second-order derivatives of functions of Q_t enter the asymptotic bias terms. In the special case where Q_t is time-invariant such that it can be written as Q, $B_{t,j}^Y$ and $B_{t,j}^F$ can be simplified to:

$$B_{t,j}^Y = \mathbb{Q}\left[(\mathrm{d}\Sigma_{ZF,jt})\left(\mathrm{d}\phi_{jt}^{(1,2)}\right)' + \frac{1}{2}\Sigma_{ZF,jt}\left(\mathrm{d}^2\phi_{jt}^{(1,2)}\right)' + (\mathrm{d}\Sigma_{ZY,jt})\,\mathrm{d}\phi_{jt}^{(1,1)} + \frac{1}{2}\Sigma_{ZY,jt}\mathrm{d}^2\phi_{jt}^{(1,1)}\right], \text{ and}$$
$$B_{t,j}^F = \mathbb{Q}\left[(\mathrm{d}\Sigma_{ZY,jt})\left(\mathrm{d}\phi_{jt}^{(2,1)}\right)' + \frac{1}{2}\Sigma_{ZY,jt}\left(\mathrm{d}^2\phi_{jt}^{(2,1)}\right)'(\mathrm{d}\Sigma_{ZF,jt})\left(\mathrm{d}\phi_{jt}^{(2,2)}\right)' + \frac{1}{2}\Sigma_{ZF,jt}\left(\mathrm{d}^2\phi_{jt}^{(2,2)}\right)'\right]Q^{-1},$$

where $\mathbb{Q} \equiv \mathbb{I}_p \otimes \mathrm{diag}(\mathbb{I}_K, Q^{-1})$. It is slightly more complicated than the usual second-order bias of local constant functional coefficient estimation because the second moment matrices $\Sigma_{ZY,jt}$ and $\Sigma_{ZF,jt}$ typically depend on time t when ϕ_{jt} is TV.

$B_{t,j}^Y$ and $B_{t,j}^F$ contain terms related to the derivatives of unknown TV coefficients, which are hard to estimate consistently. Thus, we advocate using the under-smoothing bandwidth to eliminate the asymptotic bias effect and hence focus on the estimation of the asymptotic variance-covariance (VC) matrix. Let $\Omega_t \equiv [\mathbb{I}_{K+R} \otimes (\mathbb{Q}_t\Sigma_{ZZ,t}\mathbb{Q}_t')^{-1}]\Sigma_{Z\otimes\varepsilon}^{(t)}[\mathbb{I}_{K+R} \otimes (\mathbb{Q}_t\Sigma_{ZZ,t}\mathbb{Q}_t')^{-1}]$. To estimate Ω_t, it suffices to consider consistent estimate of $\mathbb{Q}_t\Sigma_{ZZ,t}\mathbb{Q}_t'$ and $\Sigma_{Z\otimes\varepsilon}^{(t)}$. By the proof of Theorem 3.1 in the online supplement, we have $\hat{S}_{ZZ,t} \equiv T^{-1}\sum_{s=1}^T k_{h_2,st}\hat{Z}_s\hat{Z}_s' = \mathbb{Q}_t\Sigma_{ZZ,t}\mathbb{Q}_t' + o_P(1)$. Define $\hat{\Sigma}_{Z\otimes\varepsilon}^{(t)} = \frac{1}{T}\sum_{s=1}^T k_{h_2,st}\left(\hat{\varepsilon}_s^{(t)}\hat{\varepsilon}_s^{(t)\prime}\right) \otimes \left(\hat{Z}_s\hat{Z}_s'\right)$, where $\hat{\varepsilon}_s^{(t)} = k_{h_2,st}^{1/2}(\hat{W}_s - \hat{\Psi}_t'\hat{Z}_s)$. It is straightforward to show that $\hat{\Sigma}_{Z\otimes\varepsilon}^{(t)} = \Sigma_{Z\otimes\varepsilon}^{(t)} + o_P(1)$. Then we can consistently estimate Ω_t using $\hat{\Omega}_t = \left(\mathbb{I}_{K+R} \otimes \hat{S}_{ZZ,t}^{-1}\right)\hat{\Sigma}_{Z\otimes\varepsilon}^{(t)}\left(\mathbb{I}_{K+R} \otimes \hat{S}_{ZZ,t}^{-1}\right)$. Obviously, the estimate $\hat{\Omega}_t$ is the local version of White-Eicker estimate of the asymptotic variance and is robust to conditional heteroskedasticity. If we further assume conditional homoskedasticity so that $E(\varepsilon_t\varepsilon_t'|Z_t) = \Sigma_\varepsilon$ for any t, the estimate $\hat{\Omega}_t$ can be simplified to $\hat{\Omega}_t = \left(\mathbb{I}_{K+R} \otimes \hat{S}_{ZZ,t}^{-1}\right)\hat{\Sigma}_\varepsilon^{(t)}$, where $\hat{\Sigma}_\varepsilon^{(t)} = \frac{1}{T}\sum_{s=1}^T \hat{\varepsilon}_s^{(t)}\hat{\varepsilon}_s^{(t)\prime}$.

Remark 3 (Lag Order Selection). In the above estimation procedure, we have assumed that the lag order p is known. In practice, p is usually unknown and should be determined *a priori*. We consider the BIC-type information criterion (IC) for the FAVAR model:

$$IC(p) = \ln V\left(p, \{\hat{\Psi}_t\}_{t\in[T]}\right) + \rho_T(R+K)^2 p, \tag{3.1}$$

where $V\left(p, \{\hat{\Psi}_t\}_{t\in[T]}\right) \equiv T^{-1}\sum_{t=1}^{T}\left(\hat{W}_t - \hat{\Psi}_t'\hat{Z}_t\right)'\left(\hat{W}_t - \hat{\Psi}_t'\hat{Z}_t\right)$, $\rho_T \to 0$, and $\rho_T T \to \infty$. When $\rho_T = \log(T)/T$, the above IC reduces to the commonly used BIC one. Let $\hat{p} = \arg\min_p IC(p)$. It is standard to show that $\hat{p}$ consistently estimates p.

4 Specification Testing

In this section, we propose tests to detect TV features in the FAVAR model. As mentioned above, it is inappropriate to test constancy of ϕ_{jt}, $j \in [p]$, if the factor loadings λ_{it}'s are TV.

4.1 Hypotheses of Interest

We first consider the following null hypothesis:

$$\mathbb{H}_0^{(1)} : \lambda_{it} = \lambda_{i0} \text{ for all } (i,t) \in [N] \times [T] \text{ and } \phi_{jt} = \phi_{j0} \text{ for all } (j,t) \in [p] \times [T],$$

for some $\lambda_{i0} \in \mathbb{R}^R$ and $\phi_{j0} \in \mathbb{R}^{(R+K)\times(R+K)}$. The alternative hypothesis $\mathbb{H}_A^{(1)}$ is the negation of $\mathbb{H}_0^{(1)}$, viz., $\mathbb{H}_A^{(1)} : \lambda_{it} \neq \lambda_{i0}$ for some $(i,t) \in [N] \times [T]$ or $\phi_{jt} \neq \phi_{j0}$ for some $(j,t) \in [p] \times [T]$ and for all $\lambda_{i0} \in \mathbb{R}^R$ and $\phi_{j0} \in \mathbb{R}^{(R+K)\times(R+K)}$. Under $\mathbb{H}_0^{(1)}$, both the VAR coefficients and the factor loadings are time-invariant. Then the model degenerates to the time-invariant FAVAR model considered by Bai and Ng (2006). Under $\mathbb{H}_A^{(1)}$, either the VAR coefficients or the factor loadings can vary over time, so the conventional PCA and the least squares estimators of the FAVAR model parameters typically fail to be consistent.

When we reject $\mathbb{H}_0^{(1)}$, it is of further interest to explore the source of rejection. We then test

$$\mathbb{H}_0^{(2)} : \lambda_{it} = \lambda_{i0} \text{ for all } (i,t) \in [N] \times [T] \text{ and } \mathbb{H}_0^{(3)} : \phi_{jt} = \phi_{j0} \text{ for all } (j,t) \in [p] \times [T],$$

where $\mathbb{H}_0^{(2)}$ and $\mathbb{H}_0^{(3)}$ check for the time-invariance of the factor loadings and the VAR coefficients, respectively. The alternative hypotheses $\mathbb{H}_A^{(2)}$ and $\mathbb{H}_A^{(3)}$ are the negations of $\mathbb{H}_0^{(2)}$ and $\mathbb{H}_0^{(3)}$, respectively. Superficially, we can allow the VAR coefficients to be TV under $\mathbb{H}_0^{(2)}$ and the factor loadings to be TV under $\mathbb{H}_0^{(3)}$. Indeed, there is no problem of allowing for the TV coefficients in the VAR model when testing $\mathbb{H}_0^{(2)}$ since such a test can be done prior to the estimation of the VAR model. When $\mathbb{H}_0^{(2)}$ is rejected, the local PCA estimator $\hat{F}_t$ is only consistent for the true factor F_t up to a TV rotation matrix H_t. It suggests that the coefficients $\psi_{jt}^{(1,2)}$, $\psi_{jt}^{(2,1)}$, and $\psi_{jt}^{(2,2)}$ in (2.6) must be TV even if the regression parameters, ϕ_{jt}'s, in (2.2) are time-invariant. It implies that testing $\mathbb{H}_0^{(3)}$ is meaningful only when one fails to reject $\mathbb{H}_0^{(2)}$. In this case, we can estimate the factor model by the conventional PCA and then test $\mathbb{H}_0^{(3)}$. In the following analysis of the test statistic for $\mathbb{H}_0^{(3)}$, we do not need to assume the factor loadings in (2.1) to be exactly time-invariant. Instead, it suffices to assume that they satisfy certain restrictions under the local alternatives.

4.2 Test Statistics

Under $\mathbb{H}_0^{(1)}$, we can follow Stock and Watson (2002) and Bai and Ng (2006) to estimate a time-invariant FAVAR model. Let $\Lambda_0 \equiv (\lambda_{10}, ..., \lambda_{N0})'$. We estimate the following factor model:

$$X_{it} = \lambda_{i0}' F_t + e_{it} \tag{4.1}$$

by solving the following minimization problem: $\min_{F,\Lambda} \mathrm{tr}(X - F\Lambda_0')(X - F\Lambda_0')$ subject to $(T+p)^{-1}F'F = \mathbb{I}_R$ and $\Lambda_0'\Lambda_0$ is diagonal with descending diagonal elements. Let $\tilde{\lambda}_{i0}$ and $\tilde{F}_t$ be the PCA estimators of λ_{i0} and F_t, respectively. Let $\tilde{F} \equiv (\tilde{F}_{-p+1}, ..., \tilde{F}_T)'$ and $\tilde{\Lambda}_0 \equiv (\tilde{\lambda}_{10}, ..., \tilde{\lambda}_{N0})'$. As is well known, $\tilde{F}$ equals to $\sqrt{T+p}$ times eigenvectors corresponding to the R largest eigenvalues of XX', and $\tilde{\Lambda}_0 = (\tilde{F}'\tilde{F})^{-1}\tilde{F}'X = (T+p)^{-1}\tilde{F}'X$. Let $\tilde{W}_t \equiv (Y_t', \tilde{F}_t')'$, and $\tilde{Z}_t \equiv (\tilde{W}_{t-1}', ..., \tilde{W}_{t-p}')'$. In the second stage, we regress $\tilde{W}_t$ on $\tilde{Z}_t$ to obtain the estimator $\tilde{\Psi}_0 = (\tilde{Z}'\tilde{Z})^{-1}\tilde{Z}'\tilde{W} = (\sum_{t=1}^T \tilde{Z}_t\tilde{Z}_t')^{-1}\sum_{t=1}^T \tilde{Z}_t\tilde{W}_t'$.

After obtaining the restricted estimators $\tilde{\lambda}_{i0}$, $\tilde{F}_t$, and $\tilde{\Psi}_0$, we consider the TV regressions:

$$X_{it} = \lambda_i(t/T)' \tilde{F}_t + e_{it}^{\dagger} \text{ for each } i \in [N], \text{ and } \tilde{W}_t = \Psi^{\dagger}(t/T)' \tilde{Z}_t + U_t^{\dagger}, \tag{4.2}$$

where $e_{it}^{\dagger}$ and $U_t^{\dagger}$ are the respective error terms in the above regressions that account for the estimation errors introduced by replacing F_t with $\tilde{F}_t$. Clearly, (4.2) is a nonparametric time series regression that regresses X_{it} on the estimated common factor $\tilde{F}_t$ for each i. To motivate our test statistics, we consider three potential cases:

1. When $\mathbb{H}_0^{(1)}$ holds, any nonparametric consistent estimators for $\lambda_i(\cdot)$ and $\Psi^{\dagger}(\cdot)$ in (4.2) should converge to the same probability limits as the restricted estimators $\tilde{\lambda}_{i0}$ and $\tilde{\Psi}_0$, respectively. In contrast, when $\mathbb{H}_0^{(1)}$ fails so that the factor loadings or/and the VAR coefficients are TV, the nonparametric consistent estimators for $\lambda_i(\cdot)$ or $\Psi^{\dagger}(\cdot)$ should deviate significantly from those for $\tilde{\lambda}_{i0}$ or $\tilde{\Psi}_0$, respectively. It implies that we can construct a test statistic to test $\mathbb{H}_0^{(1)}$ based on the distance between the nonparametric estimators of $\lambda_i(\cdot)$ and $\Psi^{\dagger}(\cdot)$ and the restricted estimators $\tilde{\lambda}_{i0}$ and $\tilde{\Psi}_0$.

2. When $\mathbb{H}_0^{(2)}$ holds, the nonparametric consistent estimator for $\lambda_i(\cdot)$ in (4.2) should be close to the restricted estimator $\tilde{\lambda}_{i0}$. However, if $\mathbb{H}_0^{(2)}$ is false, the probability limit of the nonparametric consistent estimator of $\lambda_i(\cdot)$ should deviate from that of $\tilde{\lambda}_{i0}$. It suggests that we can construct a test statistic to test $\mathbb{H}_0^{(2)}$ based on the distance between the nonparametric consistent estimator for $\lambda_i(\cdot)$ and the restricted estimator $\tilde{\lambda}_{i0}$.

3. The VAR part of the FAVAR model requires consistent estimation for the latent common factors. When the factor model admits TV factor loadings, the estimated factors will contain the TV features of the factor loadings. As a result, the autoregressive coefficients in the VAR representation will become TV since one cannot consistently estimate ϕ_{jt} but rather ψ_{jt}. Thus, it is meaningless to test constancy of ψ_{jt} if the factor loadings are TV. In fact, if the factor model in (2.1) suffers from structural changes, it is better to adopt the TV-FAVAR model no matter whether ϕ_{jt} is TV or not. We hence recommend testing $\mathbb{H}_0^{(3)}$ only when we fail to reject $\mathbb{H}_0^{(2)}$.

To construct our test statistics, we use the local constant (Nadaraya-Watson) estimators for $\lambda_i(t/T)$ and $\Psi^{\dagger}(t/T)$. To avoid the boundary problem, we follow Hong and Li (2005), Li and Racine (2007), and Su and Wang (2020a) to adopt the following boundary kernel:

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

Note that $k_{h,tr}^{\dagger}$ coincides with $k_{h,tr}$ in the interior region but not in the boundary regions. The local constant estimators of $\lambda_i(t/T)$ and $\Psi^{\dagger}(t/T)$ are respectively given by $\breve{\lambda}_{it} = \breve{\lambda}_i(\frac{t}{T}) = \left(\sum_{s=1}^{T} k_{h_1,st}^{\dagger}\tilde{F}_s\tilde{F}_s'\right)^{-1}\sum_{s=1}^{T} k_{h_1,st}^{\dagger}\tilde{F}_s X_{is}$ and $\breve{\Psi}_t = \breve{\Psi}(\frac{t}{T}) = \left(\sum_{s=1}^{T} k_{h_2,st}^{\dagger}\tilde{Z}_s\tilde{Z}_s'\right)^{-1}\sum_{s=1}^{T} k_{h_2,st}^{\dagger}\tilde{Z}_s\tilde{W}_s'$, where h_1 and h_2 are the bandwidths.

We then test $\mathbb{H}_0^{(1)}$ by measuring the quadratic distance between $(\breve{\lambda}_{it}, \breve{\Psi}_t)$ and $(\tilde{\lambda}_{i0}, \tilde{\Psi}_0)$, and test $\mathbb{H}_0^{(2)}$ (resp. $\mathbb{H}_0^{(3)}$) by measuring the quadratic distance between $\breve{\lambda}_{it}$ and $\tilde{\lambda}_{i0}$ (resp. $\breve{\Psi}_t$ and $\tilde{\Psi}_0$). That is, we define:

$$\hat{M}_1 = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left\|\breve{\lambda}_{it} - \tilde{\lambda}_{i0}\right\|^2 + \frac{1}{T}\sum_{t=1}^{T}\left\|\breve{\Psi}_t - \tilde{\Psi}_0\right\|^2, \tag{4.3}$$

$$\hat{M}_2 = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left\|\breve{\lambda}_{it} - \tilde{\lambda}_{i0}\right\|^2, \text{ and } \hat{M}_3 = \frac{1}{T}\sum_{t=1}^{T}\left\|\breve{\Psi}_t - \tilde{\Psi}_0\right\|^2. \tag{4.4}$$

Given that $\hat{M}_2$ and $\hat{M}_3$ shrink to zero at different rates under the respective null hypotheses, a simple summation of them does not generate a good statistic for testing $\mathbb{H}_0^{(1)}$. As a result, we consider the following standardized test statistics:

$$\begin{aligned}
\widehat{SM}_1 &= 2^{-1/2}\left[\hat{\mathbb{V}}_{2NT}^{-1/2}\left(TN^{1/2}h_1^{1/2}\hat{M}_2 - \hat{\mathbb{B}}_{2NT}\right) + \hat{\mathbb{V}}_{3T}^{-1/2}\left(Th_2^{1/2}\hat{M}_3 - \hat{\mathbb{B}}_{3T}\right)\right], \\
\widehat{SM}_2 &= \hat{\mathbb{V}}_{2NT}^{-1/2}\left(TN^{1/2}h_1^{1/2}\hat{M}_2 - \hat{\mathbb{B}}_{2NT}\right), \text{ and } \widehat{SM}_3 = \hat{\mathbb{V}}_{3T}^{-1/2}\left(Th_2^{1/2}\hat{M}_3 - \hat{\mathbb{B}}_{3T}\right).
\end{aligned}$$

Here, the recentering and scaling factors are defined as follows:

$$\begin{aligned}
\hat{\mathbb{B}}_{2NT} &= \frac{h_1^{1/2}}{N^{1/2}T^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\tilde{F}_s'\left(k_{h_1,st}^{\dagger}\tilde{S}_{FF,t}^{-1} - \mathbb{I}_R\right)\left(k_{h_1,st}^{\dagger}\tilde{S}_{FF,t}^{-1} - \mathbb{I}_R\right)\tilde{F}_s\tilde{e}_{is}^2, \\
\hat{\mathbb{B}}_{3T} &= \frac{h_2^{1/2}}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\tilde{Z}_s'\left(k_{h_2,st}^{\dagger}\tilde{S}_{ZZ,t}^{-1} - \tilde{S}_{ZZ}^{-1}\right)\left(k_{h_2,st}^{\dagger}\tilde{S}_{ZZ,t}^{-1} - \tilde{S}_{ZZ}^{-1}\right)\tilde{Z}_s\tilde{U}_s'\tilde{U}_s, \\
\hat{\mathbb{V}}_{2NT} &= \frac{4}{NT^2h_1}\sum_{s=2}^{T}\sum_{r=1}^{s-1}\bar{K}\left(\frac{s-r}{Th_1}\right)^2\left(\tilde{F}_s'\tilde{F}_r'\tilde{e}_r'\tilde{e}_s\right)^2, \\
\hat{\mathbb{V}}_{3T} &= \frac{4}{T^2h_2}\sum_{s=2}^{T}\sum_{r=1}^{s-1}\bar{K}\left(\frac{s-r}{Th_2}\right)^2\left(\tilde{Z}_s'\hat{\mathbb{S}}_T\tilde{Z}_r\tilde{U}_r'\tilde{U}_s\right)^2,
\end{aligned}$$

where $\tilde{S}_{FF,t} \equiv T^{-1}\sum_{s=1}^{T} k_{h_1,st}^{\dagger}\tilde{F}_s\tilde{F}_s'$, $\tilde{S}_{ZZ,t} \equiv T^{-1}\sum_{s=1}^{T} k_{h_2,st}^{\dagger}\tilde{Z}_s\tilde{Z}_s'$, $\tilde{S}_{ZZ} \equiv T^{-1}\sum_{s=1}^{T}\tilde{Z}_s\tilde{Z}_s'$, $\tilde{e}_{is} \equiv X_{is} - \tilde{\lambda}_{i0}'\tilde{F}_s$, $\tilde{e}_s \equiv (\tilde{e}_{1s}, ..., \tilde{e}_{Ns})'$, $\tilde{U}_s \equiv \tilde{W}_s - \tilde{\Psi}_0'\tilde{Z}_s$, and $\hat{\mathbb{S}}_T \equiv T^{-1}\sum_{t=1}^{T}\tilde{S}_{ZZ,t}^{-1}\tilde{S}_{ZZ,t}^{-1}$. In addition,

$\bar{K}(u) = \int_{-1}^{1} K(v)K(u-v)\mathrm{d}v$ is the two-fold convolution kernel of $K(\cdot)$.

Remark 4. The above test statistics avoid the local PCA estimation. Alternatively, one can also construct the test statistics based on the local PCA estimates. First, we consider the test of $\mathbb{H}_0^{(2)}$. If we estimate the model by the two-stage estimation procedure introduced in Section 3, the local PCA estimator $\hat{F}_t$ is only a consistent estimator for the latent factor F_t up to H_t in the presence of TV factor loadings. As a result, one cannot test $\mathbb{H}_0^{(2)}$ by direct comparison of the unrestricted estimate $\hat{F}_t$ and the restricted estimate $\tilde{F}_t$ but can test $\mathbb{H}_0^{(2)}$ by direct comparison of the restricted and unrestricted of the common component $(\lambda_{it}' F_t)$ as in Su and Wang (2017). This makes the derivation of the asymptotic distribution of the resulting test statistic more complicated than that of $\widehat{SM}_2$. Similar remarks hold for testing $\mathbb{H}_0^{(1)}$. Second, we consider the test of $\mathbb{H}_0^{(3)}$. The VAR part of the FAVAR model requires consistent estimation for the latent common factors. Recall that the local PCA estimator $\hat{F}_t$ is only consistent for $H_t' F_t$ when the factor loadings are TV. This makes it impossible to test the constancy of VAR coefficients even if we use the local PCA estimator. In addition, the derivation of the asymptotic properties of the local PCA estimators is quite involved, not to mention that for the test statistics based on the local PCA estimators. For these reasons, we manage to avoid the use of local PCA estimates by considering the auxiliary regressions in (4.2).

4.3 Asymptotic Null Distributions

In this subsection, we study the asymptotic null distributions of $\widehat{SM}_1$, $\widehat{SM}_2$, and $\widehat{SM}_3$ under their respective null hypotheses. To proceed, we add the following assumption.

Assumption A.4. (i) For each $i \in [N]$, the process $\{(e_{it}, F_t),\ t = -p+1, -p+2, ...\}$ is strong mixing with mixing coefficients $\alpha_i(\cdot)$, and $\alpha(\cdot) \equiv \max_i \alpha_i(\cdot)$ satisfies $\sum_{s=1}^{\infty} \alpha(s)^{\delta/(2+\delta)} \leq C$ for some $\delta > 0$. Moreover, there exists a positive integer $T_0 \in [1, T)$ such that $T^{-2} \max(T_0^4, T_0^3 h_1^{-1}, T_0^2 h_1^{-2}) \to 0$ and $N^2 T h_1^2 \alpha(T_0)^{\delta/(2+\delta)} \to 0$ as $(N,T) \to \infty$; (ii) The process $\{Y_t, t = -p+1, -p+2, ...\}$ is strong mixing with mixing coefficient $\tilde{\alpha}(\cdot)$ that satisfies $\sum_{s=1}^{\infty} \tilde{\alpha}(s)^{\delta/(2+\delta)} \leq C$ for some $\delta > 0$. In addition, there exists a positive integer $\tilde{T}_0 \in [1, T)$ such that $T^{-2} \max(\tilde{T}_0^4, \tilde{T}_0^3 h_2^{-1}, \tilde{T}_0^2 h_2^{-2}) \to 0$ and $T h_2^2 \tilde{\alpha}(\tilde{T}_0)^{\delta/(2+\delta)} \to 0$ as $T \to \infty$; (iii) For each $i \in [N]$, the process $\{e_{it}, t = -p+1, -p+2, ...\}$ is a martingale difference sequence (m.d.s.) with respect to $\mathcal{F}_{NT,t-1}$, such that $E(e_t | \mathcal{F}_{NT,t-1}) =$

0, where $\mathcal{F}_{NT,t-1}$ is the minimal σ-field generated by $(F_t, F_{t-1}, ..., e_{t-1}, e_{t-2}, ...)$; (iv) For each t, W_t^0 and ε_t are independent of the idiosyncratic errors e_{is} for all i and s.

Assumptions A.4(i) and A.4(ii) impose some weak dependence conditions on the process $\{e_{it}, F_t\}$ and $\{Y_t\}$, respectively. As Su and Wang (2020a) remark, with more complicated notation, one can allow different individual time series to have distinct mixing rates. Assumption A.4(iii) assumes that the process $\{e_{it}\}$ 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 proving the asymptotic distribution of our test statistic under the null and local alternative hypotheses. It is possible to allow for both serial and cross-sectional dependence in $\{e_{it}\}$. However, it will substantially complicate the asymptotic analysis and we are not sure how to estimate the asymptotic variance of our test statistics in this case. Assumption A.4(iv) imposes independence between the idiosyncratic errors e_{is} and the regressors and error terms in the FAVAR model. This assumption is also adopted by Bai and Ng (2006). It is essential for the asymptotic independence between $\widehat{SM}_2$ and $\widehat{SM}_3$ and greatly facilitates the derivation of the asymptotic variance of $\widehat{SM}_1$.

As mentioned above, the requirements on the bandwidth parameters h_1 and h_2 for the tests are different from those in Assumptions A.2(ii) and A.2(iii). Instead, we make the following assumption on h_1 and h_2.

Assumption A.5 (i) As $(N,T) \to \infty$, $h_1 \to 0$, $Th_1^2 \to \infty$, $Th_1^2/N^3 \to 0$, $Nh_1^2/T \to 0$, $Th_1(\ln T)^{-2} \to \infty$, $Nh_1^2(\ln T)^{-4} \to \infty$, and $T^2N^{-1}h_1^3(\ln T)^{-6} \to \infty$; (ii) As $(N,T) \to \infty$, $h_2 \to 0$, $Th_2^{3/2} \to \infty$, $Nh_2 \to \infty$, $T^{2/(8+\sigma)}h_2^{-1/2}/N \to 0$, and $Th_2^{-1/2}/N^2 \to 0$.

Assumption A.5(i) is similar to Assumption A.4(ii) in Su and Wang (2020a) and it is necessary to study the asymptotic properties of $\widehat{SM}_1$ and $\widehat{SM}_2$. Assumption A.5(ii) is needed to study $\widehat{SM}_1$ and $\widehat{SM}_3$. It is weaker than that in Assumption A.2(iii) because we do not need to consider the nonparametric estimation of the TV-FAVAR model under the global alternative.

The following theorem provides the asymptotic null distributions of the test statistics.

Theorem 4.1. *[Asymptotic null distributions] Suppose that Assumptions A.1, A.2(i), and A.3 to A.5 hold except for the smoothness conditions on $\lambda_i(\cdot)$ by Assumption A.1(iii) or/and that on $\phi_j(\cdot)$ by Assumption A.3(i). Then, $\widehat{SM}_l \xrightarrow{d} N(0,1)$ under $\mathbb{H}_0^{(l)}$ for $l = 1, 2, 3$.*

Under $\mathbb{H}_0^{(1)}$, the smoothness conditions in Assumption A.1(iii) and A.3(i) are automatically satisfied. The test statistics are based on the sample quadratic forms, which measure the squared distance between the local smoothing estimators and the global least squares estimators. All three tests are asymptotically pivotal and have a convenient asymptotic standard normal distribution under the corresponding null hypotheses. Since a large value of any statistic is in favor of the alternative, our tests are all one-sided.

4.4 Asymptotic Local Power

To study the asymptotic local power properties of $\widehat{SM}_l$, for $l = 1, 2$, and 3, we consider the following classes of local alternatives:

$$\begin{aligned}
\mathbb{H}_A^{(1)}(a_{1NT}, a_{2T}) &: \lambda_{it} = \lambda_{i0} + a_{1NT} g_{1i}(t/T) \text{ and } \Phi_t = \Phi_0 + a_{2T} g_2(t/T) \ \forall\ (i,t) \in [N] \times [T]\,; \\
\mathbb{H}_A^{(2)}(a_{1NT}) &: \lambda_{it} = \lambda_{i0} + a_{1NT} g_{1i}(t/T) \ \forall\ (i,t) \in [N] \times [T]\,; \\
\mathbb{H}_A^{(3)}(a_{2T}) &: \Phi_t = \Phi_0 + a_{2T} g_2(t/T) \ \forall\ t \in [T]\,;
\end{aligned}$$

where $a_{1NT} \to 0$ as $(N,T) \to \infty$, $a_{2T} \to 0$ as $T \to \infty$, and $g_{1i}(t/T)$ and $g_2(t/T)$ are piecewise smooth functions with a finite number of discontinuity points. Obviously, a_{1NT} and a_{2T} control the speed at which the local alternatives converge to the null hypotheses. Since $\lambda_{i0} + a_{1NT} g_{1i}(t/T) = (\lambda_{i0} + c_{1i,NT}) + a_{1NT}[g_{1i}(t/T) - c_{1i,NT}/a_{1NT}]$ for any $c_{1i,NT} \in \mathbb{R}^R$, and $\Phi_0 + a_{2T} g_2(t/T) = (\Phi_0 + c_{2,T}) + a_{2T}[g_2(t/T) - c_{2,T}/a_{2T}]$ for any $c_{2,T} \in \mathbb{R}^{(R+K)p\times(R+K)}$, we impose the normalization restrictions: $\int_0^1 g_{1i}(\tau)\mathrm{d}\tau = 0$ for all $i \in [N]$, and $\int_0^1 g_2(\tau)\mathrm{d}\tau = 0$, which facilitates the analysis of the local power properties.

Assumption A.6 (i) For each $i \in [N]$, $g_{1i}(\cdot)$ is piecewise continuous with a finite number of discontinuity points on $(0, 1]$, and satisfies that $\max_{i\in[N]} \sup_\tau |g_{1i}(\tau)| \le C$; (ii) $g_2(\cdot)$ is piecewise continuous with a finite number of discontinuity points on $(0, 1]$ and satisfies that $\sup_\tau |g_2(\tau)| \le C$; (iii) $\max_{r\in[T]} \left\| N^{-1}T^{-1}\sum_{s=1}^T \sum_{i=1}^N k_{h_1,sr} F_s e_{is} g'_{1ir} \right\| = O_P((NTh_1/\ln(NT))^{-1/2})$ and $\max_{r\in[T]} \left\| T^{-1}\sum_{s=1}^T k_{h_2,sr} Z_s^0 \varepsilon_s g'_{2r} \right\| = O_P((Th_2/\ln T)^{-1/2})$, where $g_{1ir} = g_{1i}(r/T)$ and $g_{2r} = g_2(r/T)$.

Assumptions A.6(i) and A.6(ii) allow for both sudden breaks and smooth changes in the factor loadings and VAR coefficients under the local alternatives. Assumption A.6(iii) is similar to Assumption A.5(ii) in Su and Wang (2020a).

To state the next theorem, we add some notations. Let $\Sigma_\Lambda \equiv \lim_{N\to\infty} N^{-1}\sum_{i=1}^N \lambda_{i0}\lambda_{i0}'$, $H \equiv (N^{-1}\Lambda_0'\Lambda_0)(T^{-1}F'\hat{F})V_{NT}^{-1}$, and $Q \equiv V^{1/2}\Upsilon'\Sigma_\Lambda^{-1/2}$, where V_{NT} and V are $R\times R$ diagonal matrices containing the R largest eigenvalues of $(1/NT)XX'$ and $\Sigma_\Lambda^{1/2}\Sigma_F\Sigma_\Lambda^{1/2}$ in descending order, respectively, and Υ is the eigenvector matrix of $\Sigma_\Lambda^{1/2}\Sigma_F\Sigma_\Lambda^{1/2}$ such that $\Upsilon'\Upsilon = \mathbb{I}_R$. Let $D_{Q^{-1}} \equiv \operatorname{diag}(\mathbb{I}_K, Q^{-1})$, $D_H \equiv \operatorname{diag}(\mathbb{I}_K, H)$, $\mathbb{D} \equiv \mathbb{I}_p \otimes D_H'$, $\mathbb{Q} \equiv \mathbb{I}_p \otimes D_{Q^{-1}}'$, $Z_t^\dagger \equiv \mathbb{D}Z_t^0$, $\varepsilon_t^\dagger \equiv D_H'\varepsilon_t$, $\bar{S}_{ZZ,t} \equiv T^{-1}\sum_{s=1}^T k_{h_2,st}^\dagger \Sigma_{ZZ,s}$, and $\bar{\mathbb{S}}_T \equiv \mathbb{Q}\left(T^{-1}\sum_{t=1}^T \bar{S}_{ZZ,t}^{-1}\bar{S}_{ZZ,t}^{-1}\right)\mathbb{Q}'$. We further define

$$\mathbb{B}_{2NT} = \frac{h_1^{1/2}}{T^2N^{1/2}}\sum_{i=1}^N\sum_{t=1}^T\sum_{s=1}^T F_s'H\left(k_{h_1,st}^\dagger\tilde{S}_{FF,t}^{-1} - \mathbb{I}_R\right)\left(k_{h_1,st}^\dagger\tilde{S}_{FF,t}^{-1} - \mathbb{I}_R\right)H'F_s e_{is}^2,$$

$$\mathbb{B}_{3T} = \frac{h_2^{1/2}}{T^2}\sum_{t=1}^T\sum_{s=1}^T Z_s^{\dagger\prime}\left(k_{h_2,st}^\dagger\tilde{S}_{ZZ,t}^{-1} - \tilde{S}_{ZZ}^{-1}\right)\left(k_{h_2,st}^\dagger\tilde{S}_{ZZ,t}^{-1} - \tilde{S}_{ZZ}^{-1}\right)Z_s^\dagger\varepsilon_s^{\dagger\prime}\varepsilon_s^\dagger,$$

$$\Pi_{2NT} = \frac{1}{N}\sum_{i=1}^N \operatorname{tr}\left(Q\left[\int_0^1 g_{1i}(\tau)\,g_{1i}(\tau)'\,\mathrm{d}\tau\right]Q'\right),$$

$$\Pi_{3T} = \left[\operatorname{vec}\left(D_{Q^{-1}}D_{Q^{-1}}'\right)\right]'\left[\int_0^1 g_2(\tau)'\otimes g_2(\tau)'\mathrm{d}\tau\right]\operatorname{vec}\left(\mathbb{I}_p\otimes D_{Q^{-1}}'^{-1}D_{Q^{-1}}^{-1}\right),$$

$$\mathbb{V}_{2NT} = \frac{4}{NT^2h_1}\sum_{s=2}^T\sum_{r=1}^{s-1}\bar{K}\left(\frac{s-r}{Th_1}\right)^2 E\left[\left(F_r'\Sigma_F^{-1}Q'Q\Sigma_F^{-1}F_s e_r' e_s\right)^2\right]$$

$$\mathbb{V}_{3T} = \frac{4}{T^2h_2}\sum_{s=2}^T\sum_{r=1}^{s-1}\bar{K}\left(\frac{s-r}{Th_2}\right)^2 E\left\{\left[Z_s^{0\prime}\bar{\mathbb{S}}_T Z_r^0\varepsilon_r' D_{Q^{-1}}D_{Q^{-1}}'\varepsilon_s\right]^2\right\}$$

The following theorem provides the asymptotic local power of $\widehat{SM}_l$ for $l = 1, 2, 3$.

Theorem 4.2. *[Asymptotic local power] Suppose that Assumptions A.1, A.2(i), and A.3 to A.6 hold except for the smoothness conditions on $\lambda_i(\cdot)$ by Assumption A.1(iii) or/and that on $\phi_j(\cdot)$ by Assumption A.3(i). Let $a_{1NT} = T^{-1/2}N^{-1/4}h_1^{-1/4}$ and $a_{2T} = T^{-1/2}h_2^{-1/4}$. Let $\pi^{(2)} = \lim_{(N,T)\to\infty}\Pi_{2NT}/\mathbb{V}_{2NT}^{1/2}$, $\pi^{(3)} = \lim_{T\to\infty}\Pi_{3T}/\mathbb{V}_{3T}^{1/2}$, and $\pi^{(1)} = 2^{-1/2}(\pi^{(2)} + \pi^{(3)})$. Then*

(i) $\hat{\mathbb{B}}_{2NT} = \mathbb{B}_{2NT} + o_P(1)$, $\hat{\mathbb{B}}_{3T} = \mathbb{B}_{3T} + o_P(1)$, and $\widehat{SM}_1 \xrightarrow{d} N(\pi^{(1)}, 1)$ under $\mathbb{H}_A^{(1)}(a_{1NT}, a_{2T})$;

(ii) $\hat{\mathbb{B}}_{2NT} = \mathbb{B}_{2NT} + o_P(1)$, and $\widehat{SM}_2 \xrightarrow{d} N(\pi^{(2)}, 1)$ under $\mathbb{H}_A^{(2)}(a_{1NT})$;

(iii) $\hat{\mathbb{B}}_{3T} = \mathbb{B}_{3T} + o_P(1)$, and $\widehat{SM}_3 \xrightarrow{d} N(\pi^{(3)}, 1)$ under $\mathbb{H}_A^{(3)}(a_{2T})$.

Theorem 4.2 shows that $\widehat{SM}_1$, $\widehat{SM}_2$, and $\widehat{SM}_3$ have nontrivial asymptotic power against $\mathbb{H}_A^{(1)}(a_{1NT}, a_{2T})$, $\mathbb{H}_A^{(2)}(a_{1NT})$, and $\mathbb{H}_A^{(3)}(a_{2T})$, respectively, with $a_{1NT} = T^{-1/2}N^{-1/4}h_1^{-1/4}$ and

$a_{2T} = T^{-1/2}h_2^{-1/4}$. We note that Assumption A.6 allows for a finite number of unknown discontinuity points in the factor loadings and VAR coefficients. As a result, our tests have power for smooth structural changes and abrupt structural breaks, with possibly unknown break dates and an unknown number of breaks. Moreover, our tests do not require trimming the boundary regions of the sample. Hence, our tests can detect structural changes near the beginning or the ending of the sample. In the proof of Theorem 4.2(iii), we maintain that $\mathbb{H}_A^{(2)}(a_{1NT})$ holds for the reason that the VAR part of the FAVAR model requires consistent estimation for the common factor. Another reason for doing so is that the asymptotic properties of the factor estimators in the TV factor model remain unknown under the global alternative $\mathbb{H}_A^{(2)}$.

Remark 5. If the factor loadings are TV, it is inappropriate to set the VAR coefficient to be time-invariant due to the presence of TV rotation matrix. However, by the established relationship between ϕ_{jt} and ψ_{jt}, we have that $\psi_{jt}^{(1,1)} = \phi_{jt}^{(1,1)}$ no matter whether $\{\lambda_{it}\}$ are TV or not. Hence, we can test the null hypothesis $\mathbb{H}_0^{(3S)}: \phi_{jt}^{(1,1)} = \phi_{j0}^{(1,1)}$ for all $(j,t) \in [p] \times [T]$, even when $\mathbb{H}_0^{(2)}$ does not hold. The alternative hypothesis $\mathbb{H}_A^{(3S)}$ is the negation of $\mathbb{H}_0^{(3S)}$. Let $L_1 \equiv (\mathbb{I}_K, 0_{K\times R})$ be a $K \times (K+R)$ selection matrix such that $L_1\phi_{jt}L_1' = \phi_{jt}^{(1,1)}$, and $L_1\psi_{jt}L_1' = \psi_{jt}^{(1,1)}$. Then $\Psi_t^{(1,1)\prime} \equiv (\psi_{1t}^{(1,1)}, ..., \psi_{pt}^{(1,1)}) = L_1\Psi_t'\mathbb{L}_1'$ with $\mathbb{L}_1 \equiv \mathbb{I}_p \otimes L_1$. Let $\breve{\Psi}_t^{(1,1)\prime} = L_1\breve{\Psi}_t'\mathbb{L}_1'$ and $\tilde{\Psi}_0^{(1,1)\prime} = L_1\tilde{\Psi}_0'\mathbb{L}_1'$. We test $\mathbb{H}_0^{(3S)}$ by considering the following test statistic: $\hat{M}_{3S} = \frac{1}{T}\sum_{t=1}^T \left\| \breve{\Psi}_t^{(1,1)} - \tilde{\Psi}_0^{(1,1)} \right\|^2$. The standardized version is $\widehat{SM}_{3S} = \hat{\mathbb{V}}_{3T,S}^{-1/2}\left(Th_2^{1/2}\hat{M}_3 - \hat{\mathbb{B}}_{3T,S}\right)$, where $\hat{\mathbb{B}}_{3T,S} = \frac{h_2^{1/2}}{T^2}\sum_{t=1}^T\sum_{s=1}^T \tilde{Z}_s'\left(k_{h_2,st}^{\dagger}\tilde{S}_{ZZ,t}^{-1} - \tilde{S}_{ZZ}^{-1}\right)\mathbb{L}\left(k_{h_2,st}^{\dagger}\tilde{S}_{ZZ,t}^{-1} - \tilde{S}_{ZZ}^{-1}\right)\tilde{Z}_s\tilde{U}_s'L\tilde{U}_s$, and $\hat{\mathbb{V}}_{3T,S} = \frac{4}{T^2h_2}\sum_{s=2}^T\sum_{r=1}^{s-1}\bar{K}\left(\frac{s-r}{Th_2}\right)^2\left(\tilde{Z}_s'\hat{\mathbb{S}}_{T,S}\tilde{Z}_r\tilde{U}_r'L\tilde{U}_s\right)^2$, with $\hat{\mathbb{S}}_{T,S} \equiv T^{-1}\sum_{t=1}^T \tilde{S}_{ZZ,t}^{-1}\mathbb{L}\tilde{S}_{ZZ,t}^{-1}$, and $\mathbb{L} = \mathbb{I}_p \otimes L$ with $L \equiv (\mathbb{I}_K, 0_{K\times R}; 0_{R\times K}, 0_{R\times R})$. Let $\Phi_t^{(1,1)\prime} \equiv (\phi_{1t}^{(1,1)}, ..., \phi_{pt}^{(1,1)}) = L_1\Phi_t'\mathbb{L}_1'$. Following the proof of Theorem 4.2, we can show that under $\mathbb{H}_A^{(3S)}(a_{2T}): \Phi_t^{(1,1)} = \Phi_0^{(1,1)} + a_{2T}\mathbb{L}_1 g_2(t/T) L_1'$ with $a_{2T} = T^{-1/2}h_2^{-1/4}$, $\widehat{SM}_{3S} \xrightarrow{d} N(\mu_{3S}, 1)$, where μ_{3S} are defined in (S3.1) in the online supplement.

4.5 Bootstrap Versions of the Tests

As is well known, nonparametric kernel-based tests can have severe size distortions in finite samples and are also sensitive to the choice of bandwidth. To overcome these problems, we propose a resampling procedure to improve the finite sample performance of our tests.

Since we allow for weak cross-sectional dependence (CD) among the error terms in the factor model, standard wild bootstraps do not work well in the presence of CD. Here, we propose a bootstrap procedure that is robust to the presence of CD in $\{e_{it}\}$. Let $\tilde{\Sigma}_e^0 \equiv T^{-1}\sum_{t=1}^{T}\tilde{e}_t\tilde{e}_t'$, and denote its (i,j)th element as $\tilde{\sigma}_{e,ij}^0$. To generate the bootstrap errors $\{e_t^*\}$ that share the variance-covariance structure as $\{e_t\}$ asymptotically, we follow Fan et al. (2013) to obtain a consistent estimate of Σ_e in terms of spectral norm. Let $\hat{\theta}_{ij} = T^{-1}\sum_{t=1}^{T}[\tilde{e}_{it}\tilde{e}_{jt} - \tilde{\sigma}_{e,ij}^0]^2$. Define $\tilde{\Sigma}_e = \{\tilde{\sigma}_{ij}\}$ with $\tilde{\sigma}_{ij} = \tilde{\sigma}_{e,ij}^0\mathbf{1}(i=j) + s_{ij}(\tilde{\sigma}_{e,ij}^0)\mathbf{1}(i \neq j)$, where $s_{ij}(z) = \text{sgn}(z)(|z| - \tau_{ij})_+$ is the soft thresholding function with $a_+ \equiv \max(a,0)$, $\tau_{ij} = C_0 \cdot [(N \wedge T)^{-1}\log T]^{1/2}\,\hat{\theta}_{ij}^{1/2}$, and C_0 is a positive constant. In this paper, we let $C_0 = 1$ initially. If $C_0 = 1$ can not deliver a positive definite matrix $\tilde{\Sigma}_e$, we choose C_0 to be the smallest value such that $\tilde{\Sigma}_e$ is positive definite with a grid search approach. In most situations, $\tilde{\Sigma}_e$ is positive definite when $C_0 = 1$. In addition, let $\tilde{\Sigma}_U \equiv T^{-1}\sum_{t=1}^{T}\tilde{U}_t\tilde{U}_t'$, and denote its (i,j)th element as $\tilde{\sigma}_{U,ij}$. By constructions, $\tilde{\Sigma}_e$ and $\tilde{\Sigma}_U$ are symmetric and positive semi-definite. The detailed bootstrap procedure is given as follows:

1. Estimate the restricted model $X_{it} = \lambda_{i0}'F_t + e_{it}$ by the conventional PCA to obtain the restricted estimates $\{\tilde{F}_t\}$ and $\{\tilde{\lambda}_{i0}\}$ in the first stage, and estimate $\tilde{W}_t = \Psi_0'\tilde{Z}_t + U_t^{\dagger}$ to obtain the least squares estimate $\tilde{\Psi}_0$ in the second stage. Obtain the nonparametric kernel estimates $\{\check{\lambda}_{it}\}$ and $\{\check{\Psi}_t\}$, and then compute $\widehat{SM}_1$, $\widehat{SM}_2$, and $\widehat{SM}_3$, respectively.

2. For each $t \in [T]$, generate an $N \times 1$ vector $\vartheta_t = (\vartheta_{1t}, ..., \vartheta_{Nt})'$ and a $(R+K) \times 1$ vector $\upsilon_t = \left(\upsilon_{1t}, ..., \upsilon_{(K+R)t}\right)'$, where $\vartheta_{it} \sim i.i.d.\ N(0,1)\ \forall\ i \in [N]$, $\upsilon_{jt} \sim i.i.d.\ N(0,1)\ \forall\ j \in [K+R]$, and ϑ_t is independent of υ_t. We construct the bootstrap errors $e_t^* = \tilde{\Sigma}_e^{1/2}\vartheta_t$ and $U_t^* = \tilde{\Sigma}_U^{1/2}\upsilon_t$. Generate $X_{it}^* = \tilde{\lambda}_{i0}'\tilde{F}_t + e_{it}^*$ and $\tilde{W}_t^* = \tilde{\Psi}_0'\tilde{Z}_t + U_t^*$ for $(i,t) \in [N] \times [T]$.

3. Based on the bootstrap sample $\{X_{it}^*, i \in [N],\ t \in [T]\}$, conduct the conventional time-invariant PCA to obtain the bootstrap versions $\{\tilde{F}_t^*, \tilde{\lambda}_{i0}^*\}$ of $\{\tilde{F}_t, \tilde{\lambda}_{i0}\}$; run the restricted model $\tilde{W}_t^* = \Psi_0'\tilde{Z}_t + U_t^{\dagger}$ to obtain the bootstrap version estimates $\tilde{\Psi}_0^*$; run X_{it}^* on $\tilde{F}_t^*$ to obtain the local constant estimate $\check{\lambda}_{it}^*$ with the same kernel and bandwidth as used to obtain $\check{\lambda}_{it}$; and run $\tilde{W}_t^*$ on $\tilde{Z}_t$ to obtain the local constant estimate $\check{\Psi}_t^*$ with the same kernel and bandwidth as used to obtain $\check{\Psi}_t$. Calculate the bootstrap test statistics $\widehat{SM}_l^*$, the bootstrap versions of $\widehat{SM}_l$ for $l \in [3]$.

4. Repeat Steps 2 and 3 for $\mathfrak{B}$ times and index the bootstrap test statistics as $\{\widehat{SM}_{l,b}^{*}\}_{b=1}^{\mathfrak{B}}$ for $l \in [3]$. The bootstrapped p-values are calculated by $p_l^* = \mathfrak{B}^{-1}\sum_{b=1}^{\mathfrak{B}} \mathbf{1}(\widehat{SM}_{l,b}^{*} > \widehat{SM}_l)$ for $l \in [3]$.

The key in the above bootstrap procedure is Step 2, in which we generate the bootstrap sample $\{X_{it}^*\}$ as in Su and Wang (2017) and apply the fixed-regressor wild bootstrap to generate the bootstrap sample $\{\tilde{W}_t^*\}$. The latter is inspired by the fixed-regressor bootstrap procedure of Hansen (2000) who showed that there is no need to mimic the dynamic feature of a time series process in the bootstrap world.

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

Theorem 4.3. *[Asymptotic validity of the bootstrap procedure] Let Assumptions A.1, A.2(i), and A.3 to A.6 hold. Suppose that (i) there exists some* $\gamma_0 \in [0,1)$ *such that* $\max_{1\leq i\leq N}\sum_{i=1}^{N}|\sigma_{e,ij}|^{\gamma_0} \leq C$, *(ii)* $T^{-1}\sum_{t=1}^{T}\left\|\tilde{F}_t\right\|^8 = O_P(1)$, *and (iii)* $N^{-1}\sum_{i=1}^{N}\left\|\tilde{\lambda}_{i0}\right\|^8 = O_P(1)$. *Then,* $\widehat{SM}_l^{*} \xrightarrow{d^*} N(0,1)$ *in probability for* $l = 1,2,3$, *where* $\xrightarrow{d^*}$ *denotes weak convergence under the bootstrap probability measure conditional on the observed sample* $\{X_{it}, Y_t\}_{i\in[N],\, t\in[T]}$.

The first side condition in Theorem 4.3, viz., $\max_{1\leq i\leq N}\sum_{i=1}^{N}|\sigma_{e,ij}|^{\gamma_0} \leq C$, is the key condition to ensure $\|\tilde{\Sigma}_e - \Sigma_e\|_{sp} = o_p(1)$ by following the analysis of Fan et al. (2013). The other side conditions are also imposed in Su and Wang (2020) to facilitate the proof. Theorem 4.3 shows that the proposed bootstrap procedure provides an asymptotic valid approximation to the asymptotic null distributions of $\widehat{SM}_l$ for $l = 1,2,3$, no matter whether the respective null hypotheses are satisfied or not. This implies the above bootstrap tests have the correct asymptotic size under their respective null hypotheses. While under the alternative hypotheses, we have that the $\widehat{SM}_l$'s diverge to infinity in probability. As a result, the bootstrap tests are consistent against TV parameters in the FAVAR model.

5 Monte Carlo Study

Here we study the finite sample performance of our estimators and tests via simulations.

5.1 Data Generating Process

We consider the following FAVAR(1) model: $X_{it} = \lambda_{it,1}F_{1t} + \lambda_{it,2}F_{2t} + e_{it}$ and

$$\begin{pmatrix} Y_t \\ F_{1t} \\ F_{2t} \end{pmatrix} = \phi_t \begin{pmatrix} Y_{t-1} \\ F_{1(t-1)} \\ F_{2(t-1)} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{pmatrix},$$

where we set the dimensions of Y_t and F_t to be $K = 1$ and $R = 2$, respectively. Denote $\Xi(\cdot;\kappa,\gamma) = \{1+\exp[-\kappa\prod_{l=1}^{L}(\cdot-\gamma_l)/s]\}^{-1}$ as the logistic function with tuning parameter κ and location parameter $\gamma = (\gamma_1,\cdots,\gamma_L)'$. We consider the following setups for the factor loadings $\lambda_{it} = (\lambda_{it,1},\lambda_{it,2})'$ and the VAR coefficients ϕ_t.

DGP 1: (time-invariant factor loadings and VAR coefficients)

$\lambda_{it} = \lambda_{i0} \sim i.i.d.\ N(0,\mathbb{I}_2)$, and $\phi_t = \phi_0 = (0.5, 0.4, 0.3; 0, 0.6, 0; 0, 0, 0.3)$;

DGP 2: (time-invariant factor loadings and VAR coefficients with smooth changes)

$\lambda_{it} = \lambda_{i0} \sim i.i.d.\ N(0,\mathbb{I}_2)$, and $\phi_t = (\phi_t^{(1,1)},\phi_t^{(1,2)},\phi_t^{(1,3)}; 0, 0.6, 0; 0, 0, 0.3)$ with $\phi_t^{(1,1)} = -0.2 + \Xi(10t/T,1,5)$, $\phi_t^{(1,2)} = 0.9 - \Xi(10t/T,1,6)$, and $\phi_t^{(1,3)} = -0.2 + \Xi(10t/T;1,5)$.

DGP 3: (time-invariant factor loadings and VAR coefficients with an abrupt break)

$\lambda_{it} = \lambda_{i0} \sim i.i.d.\ N(0,\mathbb{I}_2)$, and $\phi_t = (\phi_t^{(1,1)},\phi_t^{(1,2)},\phi_t^{(1,3)}; 0, 0.6, 0; 0, 0, 0.3)$,

where $\phi_t^{(1,j)} = \begin{cases} -0.5 + 0.5\phi_0^{(1,j)}, & \text{for } t \le T/2 \\ 0.3 + 0.5\phi_0^{(1,j)}, & \text{for } t \ge T/2+1 \end{cases}$ with $\phi_0^{(1,j)} \sim i.i.d.U(0,1)$ for $j = 1,2,3$.

DGP 4: (smooth changes in both factor loadings and VAR coefficients)

$\lambda_{it,1} = \mu_i + \Xi(10t/T; 0.1, (2,4,6,8)')$ with $\mu_i \sim i.i.d.U(0,1)$, and $\lambda_{it,2} = \lambda_{i0,2} \sim i.i.d.\ N(0,1)$; and ϕ_t is the same as in DGP 2.

DGP 5: (smooth changes in factor loadings and an abrupt break in VAR coefficients)

λ_{it} is the same as in DGP 4, and ϕ_t is the same as that in DGP 3.

DGP 6: (smooth changes in both factor loadings and VAR coefficients)

λ_{it} is the same as in DGP 4, and $\phi_t = (\phi_t^{(1,1)},\phi_t^{(1,2)},\phi_t^{(1,3)}; 0, 0.6, 0; 0, 0, 0.3)$ with $\phi_t^{(1,1)} = -0.2+\Xi(10t/T; 1/3, (2,6))$, $\phi_t^{(1,2)} = -0.2+\Xi(10t/T; 0.3, (4,8)')$, and $\phi_t^{(1,3)} = 1-\Xi(10t/T; 0.1, (2,4,8)')$;

DGP 7: (smooth changes in factor loadings and multiple breaks in VAR coefficients)

λ_{it} is the same as in DGP 4, and $\phi_t = (\phi_t^{(1,1)},\phi_t^{(1,2)},\phi_t^{(1,3)}; 0, 0.6, 0; 0, 0, 0.3)$ with

$$\phi_t^{(1,j)} = \begin{cases} 0.6 + \phi_0^{(1,j)}, & \text{for } 0.1T \le t \le 0.3T \\ 0.3 + 0.5\phi_0^{(1,j)}, & \text{for } 0.4T \le t \le 0.6T \\ \phi_0^{(1,j)}, & \text{otherwise} \end{cases} \quad \text{and } \phi_0^{(1,j)} \sim i.i.d.U(0, 0.3).$$

DGP 8: (Multiple abrupt breaks in both factor loadings and VAR coefficients)

$$\lambda_{it,j} = \begin{cases} \lambda_{i0,j}, & \text{for } 0.1T \le t \le 0.2T \\ 1 + \lambda_{i0,j}, & \text{for } 0.4T \le t \le 0.5T \\ -1 + \lambda_{i0,j}, & \text{otherwise} \end{cases}$$ with $\lambda_{i0,j} \sim i.i.d.N(1, 1)$, and ϕ_t is the same as in DGP 7.

These DGPs describe various TV patterns in the factor loadings and VAR coefficients. DGPs 1 to 3 depict FAVAR models with time-invariant factor loadings and various types of VAR coefficients, i.e., the time-invariant VAR coefficients, VAR coefficients with smooth changes, and VAR coefficients with an abrupt structural break, respectively. DGPs 4 to 7 are FAVAR models with smooth-changing factor loadings and various types of TV VAR coefficients. Among them, DGPs 4 and 6 examine monotonic and non-monotonic smooth-changing VAR coefficients, and DGPs 5 and 7 consider VAR coefficients with a single abrupt structural break and multiple abrupt structural breaks, respectively. DGP 8 captures a FAVAR model with multiple structural breaks in both factor loadings and VAR coefficients.

For each DGP, we consider five types of error terms e_{it} and $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t}, \varepsilon_{3t})'$: (i) the i.i.d. case: $e_{it} \sim i.i.d.\ N(0, 1)$, and $\varepsilon_t \sim i.i.d.\ N(0, \mathbb{I}_3)$; (ii) the heteroskedastic case: $e_{it} = \sigma_i v_{it}$, where $\sigma_i \sim i.i.d.\ U(0.5, 1.5)$ and $v_{it} \sim i.i.d.\ N(0, 1)$, and $\varepsilon_{jt} = \mu_j v_{jt}, v_{jt} \sim i.i.d.\ N(0, 1),\ \mu_j \sim i.i.d.\ U(0.5, 1.5)$, for $j = 1, 2, 3$; (iii) the cross-sectionally dependent case: $e_t = (e_{1t}, ..., e_{Nt})' \sim i.i.d.\ N(0, \Sigma_e)$, and $\varepsilon_t \sim i.i.d.\ N(0, \Sigma_\varepsilon)$, where $\Sigma_e = (c_{ij})_{i,j=1,...,N}$, and $\Sigma_\varepsilon = (d_{ij})_{i,j=1,2,3}$, with $c_{ij} = 0.5^{|i-j|}$, and $d_{ij} = 0.4^{|i-j|}$; (iv) i.i.d. error terms in the factor model and stochastic volatility (SV) error terms in the VAR model: $e_{it} \sim i.i.d.N(0, 1)$, and $\varepsilon_{jt} = \sigma_{jt}\epsilon_{jt}$, where $\epsilon_{jt} \sim i.i.d.N(0, 1)$, and $\log \sigma^2_{j(t+1)} = 0.2 + 0.5 \log \sigma^2_{jt} + \eta_{jt}$ with $\eta_{jt} \sim i.i.d.N(0, 1)$ for $j = 1, 2, 3$; (v) SV error terms in both the factor and the VAR models: $e_{it} = \varpi_{it}u_{it}$, where $u_{it} \sim i.i.d.N(0, 1)$, and $\log \varpi^2_{i(t+1)} = 0.2 + 0.5 \log \varpi^2_{it} + v_{it}$ with $v_{it} \sim i.i.d.N(0, 1)$, and $\varepsilon_{jt} = \sigma_{jt}\epsilon_{jt}$, where $\epsilon_{jt} \sim i.i.d.N(0, 1)$, and $\log \sigma^2_{j(t+1)} = 0.2 + 0.5 \log \sigma^2_{jt} + \eta_{jt}$, with $\eta_{jt} \sim i.i.d.N(0, 1)$, for $j = 1, 2, 3$.

For each DGP, we generate the datasets with sample sizes $N = 100, 200$ and $T = 100, 200$. We use the Epanechnikov kernel and Silverman's rule of thumb (RoT) to set the bandwidths

$h_1 = (2.35/\sqrt{12})T^{-1/5}N^{-1/10}$ and $h_2 = (2.35/\sqrt{12})T^{-1/4}$. We also try the Uniform kernel and the Quartic kernel. We find that the choice of kernel functions have little impact on the simulation results.

5.2 Estimation Results

In this subsection, we evaluate the performance of the proposed estimators for the VAR coefficients. We compare our estimators with those proposed by Bai and Ng (2006). As mentioned above, the estimated VAR coefficients $\hat{\psi}_{jt}^{(a,b)}$ are not comparable with the true values $\phi_{jt}^{(a,b)}$ except for $(a, b) = (1, 1)$. Hence, we evaluate the accuracy of our estimators for the $(1, 1)$st element of the VAR coefficients ϕ_{jt} using the root mean squared error (RMSE): $RMSE = \sqrt{\frac{1}{MTp}\sum_{l=1}^{M}\sum_{t=1}^{T}\sum_{j=1}^{p}[\hat{\psi}_{jt,l}^{(1,1)} - \phi_{jt,l}^{(1,1)}]^2}$, where M is the number of replications, and p is the lag order. In addition, we also evaluate the estimators using the sum of squared residuals (SSR) and the RMSE of the variance estimate defined as follows: $SSR = \frac{1}{M(T-p)}\sum_{l=1}^{M}\sum_{t=p+1}^{T}\left(\hat{\varepsilon}_{1t}^{(l)}\right)^2$ and $RMSE_\sigma = \sqrt{\frac{1}{M}\sum_{l=1}^{M}\left[\frac{1}{T-p}\sum_{t=p+1}^{T}\left(\hat{\varepsilon}_{1t}^{(l)}\right)^2 - 1\right]^2}$, where $\{\hat{\varepsilon}_{1t}^{(l)}\}_{t\in[T],l\in[M]}$ is the residual corresponding to ε_{1t} at the l-th replication. We define the SSR as the average of the squared residuals, which can be regarded as an estimator for the variance of the error term. Recall that $\varepsilon_{1t} \sim i.i.d.N(0, 1)$. Hence, the closer of the SSR is to 1, the better the estimation result is. $RMSE_\sigma$ measures the RMSE of the estimated variance of the error term.

We set the lag order p to be the true value. We also assess the performance of our BIC-type information criterion given by (3.1) with $\rho_T = \log(T)/T$. The results show that our IC works fairly well for all the DGPs under investigation. To save space, we relegate the detailed results to the online supplement.

Table 1 reports the results for the VAR estimators of Bai and Ng (2006) and ours with i.i.d. error terms based on 1000 replications. To save space, we only report the results of $RMSE$ and $RMSE_\sigma$ here and relegate the results of SSR to the online supplement. As shown in the table, the RMSEs of our estimators generally decline as T increases. We note that the convergence rate of the VAR coefficients relies on T rather than N. Hence, it is reasonable that the RMSEs may not decline as N increases. Bai and Ng's (2006) estimator outperforms our estimator under DGP 1, which describes a time-invariant FAVAR model. However, it is not as good as our estimator under the other DGPs. In particular, the RMSEs of Bai and Ng's (2006) estimators generally do not decrease as T grows, indicating that their estimators are inconsistent due to the ignorance of the TV features in the factor loadings and/or VAR coefficients. We also note that the factor loadings and/or VAR coefficients exhibit abrupt structural breaks under DGPs 2, 5, 7, and 8. Although the theoretical result given by Theorem 3.1 is only applicable to the smoothing changes in factor loadings and/or VAR coefficients, the simulation results show that our estimator still outperforms Bai and Ng's (2006) when an underlying DGP admits abrupt structural breaks.

Table 1: Performance of estimation in terms of $RMSE$ and $RMSE_\sigma$

	Bai and Ng (2006)				This paper			
DGP\\(N,T)	(100,100)	(100,200)	(200,100)	(200,200)	(100,100)	(100,200)	(200,100)	(200,200)
	$RMSE$							
1	**0.0803**	**0.0536**	**0.0787**	**0.0547**	0.1627	0.1307	0.1620	0.1312
2	0.4425	0.4459	0.4431	0.4454	**0.2813**	**0.2527**	**0.2734**	**0.2522**
3	0.6636	0.6047	0.4610	0.5187	**0.4558**	**0.3450**	**0.2869**	**0.2585**
4	0.4456	0.4491	0.4439	0.4506	**0.2978**	**0.2586**	**0.2962**	**0.2619**
5	0.4679	0.5762	0.5475	0.5707	**0.3169**	**0.3064**	**0.3452**	**0.3089**
6	0.3556	0.3647	0.3545	0.3643	**0.2225**	**0.1888**	**0.2176**	**0.1890**
7	0.3957	0.5546	0.4185	0.3947	**0.3791**	**0.3154**	**0.3294**	**0.2680**
8	0.4980	0.5582	0.5484	0.6504	**0.3542**	**0.3184**	**0.3768**	**0.2927**
	$RMSE_\sigma$							
1	**0.0160**	**0.0088**	**0.0232**	**0.0096**	0.0391	0.0146	0.0466	0.0142
2	0.4104	0.4598	0.3927	0.4519	**0.0279**	**0.0123**	**0.0329**	**0.0137**
3	0.4403	0.4793	0.4419	0.5006	**0.0261**	**0.0127**	**0.0267**	**0.0127**
4	0.4298	0.4721	0.4667	0.4737	**0.0323**	**0.0123**	**0.0294**	**0.0129**
5	0.4873	0.4976	0.4829	0.4646	**0.0313**	**0.0150**	**0.0317**	**0.0119**
6	0.1654	0.2028	0.1851	0.1960	**0.0279**	**0.0149**	**0.0319**	**0.0132**
7	1.0145	1.0321	0.9700	1.0070	**0.1016**	**0.0806**	**0.1105**	**0.0842**
8	0.8194	0.7798	0.7071	0.7699	**0.0832**	**0.0513**	**0.0664**	**0.0522**

Note: The main entries report the values of $RMSE$ and $RMSE_\sigma$ based on 1000 replications. The bold entries highlight the better performance in each case.

5.3 Performance in Predictions

In this subsection, we examine the performance of the proposed TV diffusion index model in forecasting. Specifically, we report the out-of-sample mean squared forecasting errors (MSFEs)

Table 2: Performance of the prediction in terms of MSFE

DGP\\(N,T)	Bai and Ng (2006) (100,100)	(100,200)	(200,100)	(200,200)	This paper (100,100)	(100,200)	(200,100)	(200,200)
1	**0.9829**	**0.9996**	**0.9964**	**0.9877**	1.2243	1.1422	1.2570	1.1496
2	2.4027	2.3923	2.4105	2.4103	**1.3415**	**1.2305**	**1.4009**	**1.2312**
3	1.6903	1.7505	1.8171	1.7818	**1.3536**	**1.3093**	**1.4850**	**1.3545**
4	2.8344	2.8436	2.8883	2.9342	**1.5284**	**1.3484**	**1.5370**	**1.3687**
5	1.8237	1.8051	1.8548	1.8037	**1.5358**	**1.3575**	**1.5250**	**1.3817**
6	1.8720	1.7917	1.7855	1.7567	**1.5730**	**1.3651**	**1.5274**	**1.3718**
7	**1.2897**	1.2815	**1.2625**	1.3131	1.4070	**1.2496**	1.3749	**1.2775**
8	**1.1894**	1.2257	**1.2438**	1.2757	1.2662	**1.1947**	1.3276	**1.2391**

Note: The main entries report the MSFE based on 500 replications. The bold entries highlight the better performance in each case.

under the framework of Stock and Watson's (2002) time-invariant diffusion index model and our TV diffusion index model given by (2.3). This TV diffusion index model has also been studied by Wei and Zhang (2020), but they do not establish the asymptotic distribution of the estimators and their bandwidth does not satisfy our Assumption A.2.

To deal with the boundary problem for the out-of-sample prediction, we can either follow Yousuf and Ng (2021) to use the one-sided kernel or adopt the reflection method of Hall and Wehrly (1991) to avoid the use of the future information. Here we adopts the latter one. Suppose the sample size of a given in-sample is T_0. For the data in the boundary regions with $t = T_0 - \lfloor T_0 h \rfloor + 1, \cdots, T_0$, we follow Hall and Wehrly (1991) and Chen and Hong (2012) to reflect the data by obtaining the pseudodata $(Y_s, X_s') = (Y_{2T_0-s}, X_{2T_0-s}')$ for $s = T_0+1, \cdots, T_0 + \lfloor T_0 h \rfloor$. We use the synthesized data to recursively estimate the models and do one-step-ahead out-of-sample forecasting from $T_0 + 1$ onwards, where $T_0 = 0.8T$. We also consider different starting points $T_0 = 0.6T$ and $T_0 = 0.7T$. The conclusions are quite similar.

Table 2 reports the MSFEs under DGPs 1–8 with 500 replications. We note that the MSFEs using our approach decline as T increases under all DGPs. In contrast, the MSFEs via Bai and Ng's (2006) approach generally do not decline as T increases under these DGPs. As expected, Bai and Ng's (2006) approach outperforms ours under DGP 1. It also performs slightly better than ours under DGPs 7 and 8 when the sample size is small ($T = 100$). However, as T increases, our predictions outperform those made by Bai and Ng (2006). Besides, our MSFEs are lower than those of Bai and Ng's (2006) prediction under other cases, especially under DGPs 2, 4, and 6, which admit smooth structural changes.

5.4 Performance of the Tests

In this subsection, we show the finite sample performance of our tests for the TV-FAVAR model. For each DGP, we simulate 500 datasets. The number of bootstraps is $\mathfrak{B} = 200$. We note that DGP 1 satisfies $\mathbb{H}_0^{(1)}$ and is used to evaluate the size performance of all proposed tests. DGPs 2 and 3 satisfy $\mathbb{H}_0^{(2)}$ but violate $\mathbb{H}_0^{(1)}$, $\mathbb{H}_0^{(3)}$, and $\mathbb{H}_0^{(3S)}$. We use them to evaluate the size performance of $\widehat{SM}_2$, and the power performance of $\widehat{SM}_1$, $\widehat{SM}_3$, and $\widehat{SM}_{3S}$. Recall that if the factor loadings are TV, we should use the TV-FAVAR model no matter whether the VAR coefficients are TV or not. Hence, we use DGPs 4 to 8 to show the power performance.

Table 3 reports the empirical rejection rates of our tests at both 5% and 10% significance levels with i.i.d. error terms. The results under DGP 1 show that our tests have reasonable size using bootstrapped critical values. In addition, DGPs 2 and 3 only satisfy $\mathbb{H}_0^{(2)}$. Hence, the results under DGPs 2 and 3 show the size performance of $\widehat{SM}_2$ and the power performance of $\widehat{SM}_1$, $\widehat{SM}_3$, and $\widehat{SM}_{3S}$. We note that $\widehat{SM}_2$ exhibits reasonable size under these two DGPs, while the others are powerful in detecting both smooth and abrupt structural changes in the VAR coefficients. DGPs 4 to 8 specify TV-FAVAR models with either smooth or abrupt structural changes in the factor loadings and VAR coefficients. The results show that our tests are powerful in detecting various TV factor loadings and VAR coefficients. Notice that the empirical rejection rates of $\widehat{SM}_1$ and $\widehat{SM}_2$ rely on both N and T, while those of $\widehat{SM}_3$ and $\widehat{SM}_{3S}$ solely depend on T, which is consistent with our theoretical results on the local power properties. We also examine the finite sample performance of our tests for the heteroskedastic and spatially correlated error terms. The results are quite similar to those in Table 3. To save space, we relegate them to the online supplement.

6 Empirical Application

In this section, we first estimate and test the TV-FAVAR model using the U.S. macroeconomic data. We then evaluate the forecasting performance of the proposed TV-FAVAR approach on

Table 3: Empirical rejection rates of the proposed tests (i.i.d. errors)

			$\widehat{SM}_1$		$\widehat{SM}_2$		$\widehat{SM}_3$		$\widehat{SM}_{3S}$	
DGP	N	T	5%	10%	5%	10%	5%	10%	5%	10%
1	100	100	0.058	0.112	0.040	0.094	0.066	0.116	0.032	0.068
	100	200	0.106	0.158	0.050	0.112	0.116	0.158	0.034	0.052
	200	100	0.030	0.082	0.040	0.090	0.030	0.074	0.026	0.054
	200	200	0.056	0.100	0.048	0.124	0.048	0.086	0.020	0.062
2	100	100	0.252	0.376	0.062	0.112	0.254	0.408	0.428	0.540
	100	200	0.774	0.870	0.056	0.104	0.788	0.872	0.872	0.942
	200	100	0.218	0.350	0.046	0.120	0.244	0.356	0.402	0.530
	200	200	0.768	0.846	0.048	0.092	0.790	0.886	0.890	0.938
3	100	100	0.208	0.328	0.066	0.124	0.222	0.326	0.320	0.436
	100	200	0.748	0.878	0.050	0.112	0.774	0.888	0.774	0.860
	200	100	0.214	0.348	0.044	0.080	0.224	0.368	0.326	0.438
	200	200	0.770	0.852	0.042	0.096	0.782	0.874	0.792	0.888
4	100	100	0.636	0.782	0.888	0.930	0.292	0.430	0.446	0.596
	100	200	1.000	1.000	0.998	0.998	0.846	0.910	0.896	0.942
	200	100	0.752	0.838	0.954	0.974	0.264	0.410	0.394	0.544
	200	200	1.000	1.000	0.998	1.000	0.856	0.916	0.902	0.956
5	100	100	0.646	0.796	0.888	0.918	0.326	0.462	0.328	0.464
	100	200	0.994	1.000	1.000	1.000	0.886	0.942	0.798	0.866
	200	100	0.738	0.840	0.940	0.964	0.348	0.454	0.318	0.406
	200	200	0.998	0.998	0.994	0.998	0.878	0.938	0.770	0.864
6	100	100	0.558	0.716	0.896	0.930	0.156	0.242	0.432	0.542
	100	200	0.986	0.998	0.998	1.000	0.612	0.764	0.916	0.942
	200	100	0.656	0.772	0.958	0.980	0.174	0.292	0.434	0.532
	200	200	1.000	1.000	1.000	1.000	0.632	0.756	0.916	0.946
7	100	100	0.462	0.586	0.858	0.912	0.096	0.148	0.048	0.088
	100	200	0.980	0.994	0.998	1.000	0.352	0.480	0.270	0.414
	200	100	0.604	0.718	0.946	0.970	0.080	0.156	0.052	0.114
	200	200	0.986	0.994	0.998	1.000	0.348	0.504	0.294	0.418
8	100	100	0.176	0.320	0.678	0.810	0.048	0.094	0.044	0.094
	100	200	0.712	0.836	0.984	0.994	0.126	0.224	0.278	0.392
	200	100	0.164	0.294	0.698	0.800	0.046	0.086	0.048	0.122
	200	200	0.696	0.818	0.968	0.982	0.150	0.246	0.280	0.442

Note: The main entries report the empirical rejection rates based on 500 iterations.

certain key variables.

We extend Stock and Watson's (2009) dataset on the U.S. macroeconomic variables. By excluding some discontinuous series, we get $N = 101$ quarterly time series spanning 1960:I to 2019:IV. Note that the first two quarters are discarded when calculating the first and second-order differencing. We get a total of $T = 238$ quarterly observations. All the series have been standardized to have zero mean and unit variance. We use them to extract the common factors. For more details on data description, one can refer to Stock and Watson (2009) and Su and Wang (2017). For the VAR part, we focus on the following seven key economic variables: the real GDP index (RGDP), personal consumption expenditures (PCEC), industrial production index (IP), GDP implicit price deflator (GDPDEF), total unit labor cost for manufacturing (LCM), unemployment rate (UR), and Federal funds effective rate (FedR). These variables cover various aspects of the macroeconomic fundamentals, including economic condition (RGDP, RCEC, IP), price level (GDPDEF), labor market (LCM, UR), and monetary policy (FedR). All these variables are transformed as suggested by Stock and Watson (2006).

We use each of these seven variables combined with the estimated common factors to construct the FAVAR model. We first determine the number of common factors and the lag order of the FAVAR model. The maximum number of common factors is set to be 9, while the maximum lag order is set to be 4. Other settings, including the kernel function and bandwidths, are the same as in the simulation studies. Su and Wang's (2007) local information criterion IC_{h2} chooses 3 common factors. We also adopt Bai and Ng's (2002) information criteria PC_{p1}, PC_{p2}, IC_{p1}, and IC_{p2}. The estimated number of factors by PC_{p1} is 7, while the other three information criteria choose 6 common factors. It is consistent with the fact that when the factor loadings have structural changes, Bai and Ng's (2002) information criteria tend to overestimate the number of common factors. Since Su and Wang's (2017) local information criterion is valid when the factor loadings suffer from structural changes, we pick the number of common factors as suggested by Su and Wang (2017). According to the IC given by (3.1), the optimal lag order is 1 for all these targeted variables. See the online supplement for details.

Using the constructed dataset, we first test structural change in the FAVAR(1) model. Table 4 reports the p-values based on 1000 bootstrap resamples for the seven key economic variables. We note that the joint test $\widehat{SM}_1$ rejects the null hypothesis at the 5% significance level for all

Table 4: p-values for the tests of structural changes

	$\widehat{SM}_1$	$\widehat{SM}_2$	$\widehat{SM}_3$	$\widehat{SM}_{3S}$		$\widehat{SM}_1$	$\widehat{SM}_2$	$\widehat{SM}_3$	$\widehat{SM}_{3S}$
RGDP	**0.004**	**0.000**	0.768	0.739	LCM	**0.005**	**0.000**	0.570	0.170
PCEC	**0.040**	**0.000**	0.627	**0.029**	UR	**0.019**	**0.000**	0.791	0.347
IP	**0.001**	**0.000**	0.229	0.472	FedR	**0.031**	**0.000**	0.583	0.069
GDPDEF	0.053	**0.000**	0.764	**0.034**					

Note: The main entries report the p-values based on 1000 bootstrap iterations. The bold entries indicate rejection rate at the 5% significance level.

variables except the GDP deflator. However, it becomes significant at the 10% significance level. We further explore the sources of rejection. The test $\widehat{SM}_2$ significantly rejects the constant parameter factor models for all seven variables under investigation, indicating that the source of rejection for $\widehat{SM}_1$ is the TV factor structure. As mentioned above, if the factor loadings exhibit structural changes, it is inappropriate to assume the VAR coefficients to be constant. Thus, it implies that we should adopt the TV-FAVAR model regardless of the test $\widehat{SM}_3$, which is meaningless under TV factor loadings. However, the test $\widehat{SW}_{3S}$ is still informative. It shows that the top left part of the VAR coefficients is TV for PCEC, GDPDEF, and FedR at the 10% significance level.

Table 5: The out-of-sample MSFEs

	BN06	TV-FAVAR	Ratio	$z-$test	p-value
RGDP	0.3482	**0.2387**	0.6856	2.6273	**0.0086**
PCEC	0.3778	**0.3612**	0.9559	0.4423	0.6583
IP	0.3369	**0.1270**	0.3771	2.3723	**0.0177**
GDPDEF	0.2332	**0.2269**	0.9728	0.2388	0.8113
LCM	**0.5475**	0.5586	1.0204	-0.3016	0.7629
UR	0.0190	**0.0187**	0.9816	0.1322	0.8948
FedR	**0.0026**	0.0035	1.3239	-0.6946	0.4873

Notes: (i) The entries under 'BN06' and 'TV-FAVAR' report the out-of-sample MSFEs for Bai and Ng's (2006, BN06) and the proposed TV-FAVAR models, respectively. The bold entries highlight the better performance in each case. (ii) The entries under 'Ratio' report the ratios of the MSFEs of the local smoothing estimation to those of Bai and Ng's (2006) estimation. (iii) 'z-test' and 'p-value' denote Diebold and Mariano's (1995) z-statistics and the corresponding p-values.

For each of the seven target series, we compare the one-step-ahead out-of-sample forecasting performance of our TV-FAVAR model with Bai and Ng's (2006) diffusion index model. We use the sample before 2010:I to estimate the model and conduct one-step-ahead out-of-sample forecasting recursively. In addition, we use the test for predictive ability proposed by Diebold and Mariano (1995, DM) to check the statistical significance of differences in MSFE. Please see the online supplement S5.3 for a detailed discussion of the predictive ability tests. We admit that a better way is to account for the estimation error as in West (1996), but one needs to extend the latter test to the nonparametric framework first. We leave this for future research. Table 5 reports the MSFE for the seven target series of interest. We observe that our TV-FAVAR model outperforms Bai and Ng's (2006) time-invariant FAVAR model for five out of the seven target series, and the DM's z-statistics are significant for two of the five series.

7 Conclusion

FAVAR models have been widely used in macroeconomic analysis and have drawn great attention in the literature. The conventional FAVAR model assumes that both the factor loadings and the VAR coefficients are time-invariant over a long time span, which is quite restrictive and unrealistic. In this paper, we introduce a TV-FAVAR where both the factor loadings and the VAR coefficients are allowed to change smoothly over time. We propose a two-stage procedure to consistently estimate the TV factor loadings and VAR coefficients and establish the estimators' limiting distributions under the standard large N and large T framework. In addition, we propose three test statistics to gauge the possible sources of TV behavior in the FAVAR model. These test statistics are constructed by measuring the squared L^2-distances between two sets of estimators. Monte Carlo studies demonstrate that our estimators and tests perform well. In an application to the U.S. macroeconomic dataset, we find overwhelming evidence of structural changes in the FAVAR model and show that our TV-FAVAR model outperforms the time-invariant FAVAR models in predicting several macroeconomic time series of interest.

References

Bai, J. (2003), "Inferential Theory for Factor Models of Large Dimensions," Econometrica 71, 135-172.

Bai, J., and Ng, S. (2002), "Determining the Number of Factors in Approximate Factor Models," Econometrica 90, 191-221.

Bai, J., and Ng, S. (2006), “Confidence Intervals for Diffusion Index Forecasts and Inference for Factor-augmented Regressions,” Econometrica, 74(4), 1133-1150.

Bai, J., Li, K., and Lu, L. (2016), “Estimation and Inference of FAVAR Models,” Journal of Business & Economic Statistics, 34(4), 620-641.

Bai, J., and Perron, P. (1998), “Estimating and Testing Linear Models with Multiple Structural Changes,” Econometrica, 66(1), 47-78.

Baltagi, B. H., Kao, C., and Wang, F. (2021), “Estimating and Testing High Dimensional Factor Models with Multiple Structural Changes,” Journal of Econometrics, 220(2), 349-365.

Banerjee, A., Marcellino, M., and Masten, I. (2008), “Forecasting Macroeconomic Variables Using Diffusion Indexes in Short Samples with Structural Change,” Frontiers of Economics and Globalization, 3, 149-194.

Bates B.J., Plagborg-Møller M., Stock, J.H., and Watson, M.W. (2013), “Consistent Factor Estimation in Dynamic Factor Models with Structural Instability,” Journal of Econometrics 177, 289-304.

Barigozzi M., Hallin M., Soccorsi S., and von Sachs, R. (2021), “Time-varying General Dynamic Factor Models and the Measurement of Financial Connectedness,” Journal of Econometrics, 222(1), 324-343.

Bernanke, B. S., Boivin, J., and Eliasz, P. (2005), “Measuring the Effects of Monetary Policy: a Factor-augmented Vector Autoregressive (FAVAR) Approach,” Quarterly Journal of Economics, 120(1), 387-422.

Breitung, J., and Eickmeier, S. (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(1), 163-188.

Cai, Z., and Liu, X. (2021), “Solving the Price Puzzle Via A Functional Coefficient Factor-Augmented VAR Model,” Working Paper, Department of Economics, University of Kansas.

Chen, L., Dolado, J., and Gonzalo, J. (2014), “Detecting Big Structural Breaks in Large Factor Models,” Journal of Econometrics, 180, 30-48.

Chen, B., and Hong, Y. (2012), “Testing for Smooth Structural Changes in Time Series Models via Nonparametric Regression,” Econometrica, 80(3), 1157-1183.

Christiano, L. J., Eichenbaum, M., and Evans, C. L. (1999), “Monetary Policy Shocks: What Have We Learned and to What End?” Handbook of Macroeconomics, 1, 65-148.

Cogley, T., and Sargent, T. J. (2001), “Evolving Post-world War II US Inflation Dynamics,” NBER Macroeconomics Annual, 16, 331-373.

Corradi, V., and Swanson, N.R. (2014), “Testing for Structural Stability of Factor Augmented Forecasting Models,” Journal of Econometrics, 182, 100-118.

Del Negro, M., and Otrok C. (2009), “Dynamic Factor Models with Time-varying Parameters: Measuring Changes in International Business Cycles,” Staff Report No. 326, Federal Reserve Bank of New York .

Diebold F. X., and Mariano R. S. (1995), “Comparing Predictive Accuracy,” Journal of Business & Economic Statistics, 13(3): 253-263.

Eichler M., Motta G., and Von Sachs R. (2011), “Fitting Dynamic Factor Models to Non-stationary Time Series,” Journal of Econometrics, 163(1), 51-70.

Eickmeier, S., Lemke, W., and Marcellino, M. (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(3), 493-533.

Fan J., Liao Y., and Mincheva M. (2013), “Large Covariance Estimation by Thresholding Principal Orthogonal Complements,” Journal of the Royal Statistical Society: Series B, 2013, 75(4), 603-680.

Forni M., Hallin M., Lippi M., and Reichlin L. (2000), “The Generalized Dynamic Factor Model: Identification and Estimation,” Review of Economics and Statistics, 82(4), 540-554.

Forni, M., Hallin, M., Lippi, M., and Reichlin, L. (2005), “The Generalized Dynamic Factor Model: One-sided Estimation and Forecasting,” Journal of the American statistical association, 100(471), 830-840.

Forni, M., Hallin, M., Lippi, M., and Zaffaroni, P. (2017), “Dynamic Factor Models with Infinite-dimensional Factor Spaces: Asymptotic Analysis,” Journal of Econometrics, 199 (1), 74–92.

Fu Z., Hong Y., Su L., and Wang X. (2022), “Specification Tests for Time-varying Coefficient Models,” Journal of Econometrics, forthcoming.

Giraitis, L., Kapetanios, G., and Yates, T. (2014), “Inference on Stochastic Time-varying Coefficient Models,” Journal of Econometrics, 179(1), 46-65.

Giraitis L., Kapetanios G., and Marcellino, M. (2021), “Time-varying Instrumental Variable Estimation,” Journal of Econometrics, 224(2): 394-415.

Hall P., and Wehrly T.E. (1991), “A Geometrical Method for Removing Edge Effects From Kernel-Type Nonparametric Regression Estimators,” Journal of American Statistical Association, 86, 665-672.

Han, X., and Inoue, A. (2015), “Tests for Parameter Instability in Dynamic Factor Models,” Econometric Theory, 31, 1117-1152.

Hansen, B. E. (2000), “Testing for Structural Change in Conditional Models,” Journal of Econometrics, 97(1), 93-115.

Hong, Y., and Li, H. (2005), “Nonparametric Specification Testing for Continuous-time Models with Applications to Term Structure of Interest Rate,” Review of Financial Studies, 18, 37-84.

Li, D., Tosasukul, J., and Zhang, W. (2020), “Nonlinear Factor-Augmented Predictive Regression Models with Functional Coefficients,” Journal of Time Series Analysis, 41(3), 367-386.

Li, D., Qian, J., and Su, L. (2016), “Panel Data Models with Interactive Fixed Effects and Multiple Structural Breaks,” Journal of the American Statistical Association, 111(516), 1804-1819.

Li, Q., and Racine, J.S., (2007), “Nonparametric Econometrics: Theory and Practice,” Princeton University Press.

Lu, X., and Su, L. (2016), “Shrinkage Estimation of Dynamic Panel Data Models with Interactive Fixed Effects,” Journal of Econometrics, 190(1), 148-175.

Mikkelsen, J. G., Hillebrand, E., and Urga, G. (2019), "Consistent Estimation of Time-varying Loadings in High-Dimensional Factor Models," Journal of Econometrics, 208(2), 535-562.

Moon, H., R., and Weidner, M. (2015), "Linear Regression for Panel with Unknown Number of Factors as Interactive Fixed Effects," Econometrica 83, 1543–1579.

Moon, H. R., and Weidner, M. (2017), "Dynamic Linear Panel Regression Models with Interactive Fixed Effects," Econometric Theory, 33(1), 158-195.

Motta G., Hafner C. M., and Von Sachs R. (2011), "Locally Stationary Factor Models: Identification and Nonparametric Estimation," Econometric Theory, 27(6), 1279-1319.

Mumtaz H., and Surico P. (2012), "Evolving International Inflation Dynamics: World and Country-specific Factors," Journal of the European Economic Association, 10(4): 716-734.

Qu, Z., and Perron, P. (2007), "Estimating and Testing Structural Changes in Multivariate Regressions," Econometrica, 75(2), 459-502.

Robinson, P. M. (1989), "Nonparametric Estimation of Time-varying Parameters," In Statistical Analysis and Forecasting of Economic Structural Change, pp.253-264. Springer, Berlin, Heidelberg.

Robinson, P. M. (1991), "Time-varying Nonlinear Regression," In Economic Structural Change, pp.179-190. Springer, Berlin, Heidelberg.

Sims, C. A. (1992), "Interpreting the Macroeconomic Time Series Facts: the Effects of Monetary Policy," European Economic Review, 36(5), 975-1000.

Stock, J. H., and Watson, M. W. (2002), "Forecasting Using Principal Components from a Large Number of Predictors," Journal of the American Statistical Association, 97, 1167-1179.

Stock, J. H., and Watson, M. (2009), "Forecasting in Dynamic Factor Models Subject to Structural Instability," The Methodology and Practice of Econometrics. A Festschrift in Honour of David F. Hendry, 173, 205.

Su, L., and Wang, X. (2017), "On Time-varying Factor Models: Estimation and Testing,"Journal of Econometrics, 198, 84-101.

Su, L., and Wang, X. (2020a), "Testing for Structural Changes in Factor Models via a Nonparametric Regression," Econometric Theory, 36(6), 1127-1158.

Su, L., and Wang, X. (2020b), "Corrigendum to 'On Time-varying Factor Models: Estimation and Testing'," [J. Econometrics 198 (2017) 84-101], Working Paper, Singapore Management University.

Wei, J., and Zhang, Y. (2020), "A Time-varying Diffusion Index Forecasting Model," Economics Letters, 193, 109337.

West K. D., 1996. Asymptotic inference about predictive ability. Econometrica: Journal of the Econometric Society, 1067-1084.

Yan, Y., and Cheng, T. (2022), "Factor-Augmented Forecasting Regressions with Threshold Effects," Econometrics Journal, 25(1), 134-154.

Yousuf K., and Ng S. (2021), Boosting High Dimensional Predictive Regressions with Time Varying Parameters," Journal of Econometrics, 224(1): 60-87.

Online Supplement for

"Estimation and Inference on Time-Varying FAVAR Models"

Zhonghao Fu[a,b], Liangjun Su[c] and Xia Wang[d]

[a] School of Economics, Fudan University, China

[b] Shanghai Institute of Finance and Economics, China

[c] School of Economics and Management, Tsinghua University, China

[d] School of Economics, Renmin University of China, China

The online appendix contains five sections. Section S1 provides the proofs of the main results in the paper by calling upon some technical lemmas and propositions. Section S2 contains the proofs of the technical lemmas and propositions in Section S1. Section S3 outlines the derivation of (4.7) in Section 4. Seconds S4 and S5 provide additional simulation and empirical application results, respectively.

S1 Proofs of the Main Results in the Paper

We first introduce some short-hand notations. Let $k_{h,rt} \equiv h^{-1}K(\frac{r-t}{Th})$ and $\kappa_2 \equiv \int_{-1}^{1} \tau^2 K(\tau)\mathrm{d}\tau$. Let $a_T \lesssim b_T$ indicate that a_T/b_T is bounded in probability. For any $m \times n$ nonrandom time-varying matrix A_t such that the (i,j)th entry is $A_{ij,t} \equiv A_{ij}(t/T)$, we let $(\mathrm{d}A_t)$ and $(\mathrm{d}^2 A_t)$ be $m\times n$ matrices such that the (i,j)th entries of each are $[\mathrm{d}A_{ij}(\tau)/\mathrm{d}\tau]_{|\tau=t/T}$ and $[\mathrm{d}^2 A_{ij}(\tau)/(\mathrm{d}\tau)^2]_{|\tau=t/T}$. We further let $C_{NT} = \min(\sqrt{Th_1}, \sqrt{N})$ and $C_{0NT} = \min(\sqrt{T}, \sqrt{N})$.

To prove Theorem 3.1, we introduce the following two lemmas.

Lemma S1.1. *Suppose that Assumptions A.1-A.3 hold. Then*

$$A_{t,2}^Y \equiv \frac{h_2^{1/2}}{T^{1/2}} \sum_{s=1}^{T} k_{h_2,st} \hat{Z}_s \Delta_Y'(s,t) = T^{1/2} h_2^{5/2} \kappa_2 \sum_{j=1}^{p} B_{t,j}^Y + o_P(1),$$

where

$$\begin{aligned} B_{t,j}^Y \equiv\ & \left[\mathrm{d}\left(\mathbb{Q}_t \Sigma_{ZF,jt}\right)\right] \left[\mathrm{d}\phi_{jt}^{(1,2)}\right]' + \frac{1}{2}\mathbb{Q}_t \Sigma_{ZF,jt} \left[\mathrm{d}^2 \phi_{jt}^{(1,2)}\right]' \\ & + \left\{ \left[\mathrm{d}\left(\mathbb{Q}_t \Sigma_{ZF,jt} Q_{t-j}^{-1}\right)\right] \left(\mathrm{d}Q_{t-j}\right) + \frac{1}{2}\mathbb{Q}_t \Sigma_{ZF,jt} Q_{t-j}^{-1} \left(\mathrm{d}^2 Q_{t-j}\right) \right\} \phi_{jt}^{(1,2)\prime} \end{aligned}$$

$$+\left[\mathrm{d}\left(\mathbb{Q}_t\Sigma_{ZY,jt}\right)\right]\left[\mathrm{d}\phi_{jt}^{(1,1)}\right]+\frac{1}{2}\mathbb{Q}_t\Sigma_{ZY,jt}\left[\mathrm{d}^2\phi_{jt}^{(1,1)}\right]. \tag{S1.1}$$

Lemma S1.2. *Suppose that Assumptions A.1-A.3 hold. Then*

$$A_{t,2}^F \equiv \frac{h_2^{1/2}}{T^{1/2}}\sum_{s=1}^{T} k_{h_2,st}\hat{Z}_s\Delta_F'(s,t) = T^{1/2}h_2^{5/2}\kappa_2\sum_{j=1}^{p} B_{t,j}^F + o_P(1),$$

where

$$\begin{aligned}
B_{t,j}^F \;\equiv\; & \left[\mathrm{d}\left(\mathbb{Q}_t\Sigma_{ZY,jt}\right)\right]\left[\mathrm{d}\phi_{jt}^{(2,1)}\right]' Q_t^{-1} + \mathbb{Q}_t\Sigma_{ZY,jt}\left[\mathrm{d}\phi_{jt}^{(2,1)}\right]'\left(\mathrm{d}Q_t^{-1}\right) + \frac{1}{2}\mathbb{Q}_t\Sigma_{ZY,jt}\left[\mathrm{d}^2\phi_{jt}^{(2,1)}\right]' Q_t^{-1} \\
& + \left[\mathrm{d}(\mathbb{Q}_t\Sigma_{ZY,jt})\right]\phi_{jt}^{(2,1)\prime}\left(\mathrm{d}Q_t^{-1}\right) + \frac{1}{2}\mathbb{Q}_t\Sigma_{ZY,jt}\phi_{jt}^{(2,1)\prime}\left(\mathrm{d}^2Q_t^{-1}\right) \\
& + \left[\mathrm{d}\left(\mathbb{Q}_t\Sigma_{ZF,jt}\right)\right]\left[\mathrm{d}\phi_{jt}^{(2,2)}\right]' Q_t^{-1} + \mathbb{Q}_t\Sigma_{ZF,jt}\left[\mathrm{d}\phi_{jt}^{(2,2)}\right]'\left(\mathrm{d}Q_t^{-1}\right) + \frac{1}{2}\mathbb{Q}_t\Sigma_{ZF,jt}\left[\mathrm{d}^2\phi_{jt}^{(2,2)}\right]' Q_t^{-1} \\
& + \left[\mathrm{d}\left(\mathbb{Q}_t\Sigma_{ZF,jt}Q_{t-j}^{-1}\right)\right]\left(\mathrm{d}Q_t\right)\phi_{jt}^{(2,2)\prime}Q_t^{-1} + \mathbb{Q}_t\Sigma_{ZF,jt}Q_{t-j}^{-1}\left(\mathrm{d}Q_t\right)\phi_{jt}^{(2,2)\prime}\left(\mathrm{d}Q_t^{-1}\right) \\
& + \frac{1}{2}\mathbb{Q}_t\Sigma_{ZF,jt}Q_{t-j}^{-1}\left(\mathrm{d}^2Q_t\right)\phi_{jt}^{(2,2)\prime}Q_t^{-1} \\
& + \left[\mathrm{d}\left(\mathbb{Q}_t\Sigma_{ZF,jt}Q_{t-j}^{-1}\right)\right]Q_{t-j}\phi_{jt}^{(2,2)\prime}\left(\mathrm{d}Q_t^{-1}\right) + \frac{1}{2}\mathbb{Q}_t\Sigma_{ZF,jt}\phi_{jt}^{(2,2)\prime}\left(\mathrm{d}^2Q_t^{-1}\right).
\end{aligned} \tag{S1.2}$$

Proof of Theorem 3.1. Recall that $\hat{S}_{ZZ,t} = T^{-1}\sum_{s=1}^{T} k_{h_2,st}\hat{Z}_s\hat{Z}_s'$. By (2.6), we have

$$\begin{aligned}
\hat{\Psi}_t - \Psi_t &= \hat{S}_{ZZ,t}^{-1}\left(\frac{1}{T}\sum_{s=1}^{T} k_{h_2,st}\hat{Z}_s\hat{W}_s'\right) - \Psi_t \\
&= \hat{S}_{ZZ,t}^{-1}\left[\frac{1}{T}\sum_{s=1}^{T} k_{h_2,st}\hat{Z}_s\left(\Psi_t'\hat{Z}_s + U_s^{(t)}\right)'\right] - \Psi_t \\
&= \hat{S}_{ZZ,t}^{-1}\left(\frac{1}{T}\sum_{s=1}^{T} k_{h_2,st}\hat{Z}_s U_s^{(t)\prime}\right),
\end{aligned}$$

where $U_s^{(t)} = \begin{pmatrix} u_{Y,s}^{(t)} \\ u_{F,s}^{(t)} \end{pmatrix} = \begin{pmatrix} \Delta_Y(s,t) + \varepsilon_{Y,s} \\ \Delta_F(s,t) + H_s'\varepsilon_{F,s} \end{pmatrix}$. Partition $\hat{\Psi}_t = \left(\hat{\Psi}_{Y,t}, \hat{\Psi}_{F,t}\right)$ and $\Psi_t = (\Psi_{Y,t}, \Psi_{F,t})$, where $\hat{\Psi}_{Y,t}$ and $\Psi_{Y,t}$ are $(K+R)p\times K$ matrices, and $\hat{\Psi}_{F,t}$ and $\Psi_{F,t}$ are $(K+R)p\times R$ matrices. We first consider $\hat{\Psi}_{Y,t}$.

$$\sqrt{Th_2}\left(\hat{\Psi}_{Y,t} - \Psi_{Y,t}\right) = \sqrt{Th_2}\hat{S}_{ZZ,t}^{-1}\left[\frac{1}{T}\sum_{s=1}^{T} k_{h_2,st}\hat{Z}_s u_{Y,s}^{(t)\prime}\right]$$

$$
\begin{aligned}
&= \hat{S}_{ZZ,t}^{-1}\left[\frac{h_2^{1/2}}{T^{1/2}}\sum_{s=1}^{T}k_{h_2,st}\hat{Z}_s\varepsilon_{Y,s}' + \frac{h_2^{1/2}}{T^{1/2}}\sum_{s=1}^{T}k_{h_2,st}\hat{Z}_s\Delta_Y'(s,t)\right] \\
&\equiv \hat{S}_{ZZ,t}^{-1}\left[A_{t,1}^{Y} + A_{t,2}^{Y}\right].
\end{aligned}
$$

Recall that $Z_t = \left(Y_{t-1}',\ F_{t-1}'H_{t-1}, ...,\ Y_{t-p}',\ F_{t-p}'H_{t-p}\right)'$, $\hat{Z}_t = (Y_{t-1}',\ \hat{F}_{t-1}', ...,\ Y_{t-p}',\ \hat{F}_{t-p}')'$, and $Z_t^0 = \left(Y_{t-1}',\ F_{t-1}', ...,\ Y_{t-p}',\ F_{t-p}'\right)' = \left(W_{t-1}', ..., W_{t-p}'\right)'$. Denote $\mathbb{D}_t = \text{diag}(\mathbb{I}_K, H_{t-1}', \cdots, \mathbb{I}_K, H_{t-p}')$ and recall that $\mathbb{Q}_t = \text{diag}(\mathbb{I}_K, Q_{t-1}'^{-1}, \cdots, \mathbb{I}_K, Q_{t-p}'^{-1})$. Then $Z_t = \mathbb{D}_t Z_t^0$. By Lemma A.1(iii) in Su and Wang (2017), we have that $\mathbb{D}_t = \mathbb{Q}_t + O_P(C_{NT}^{-1})$. Note that

$$
\begin{aligned}
\hat{S}_{ZZ,t} &= \frac{1}{T}\sum_{s=1}^{T}k_{h_2,st}Z_sZ_s' + \frac{1}{T}\sum_{s=1}^{T}k_{h_2,st}(\hat{Z}_s - Z_s)(\hat{Z}_s - Z_s)' \\
&\quad + \frac{1}{T}\sum_{s=1}^{T}k_{h_2,st}(\hat{Z}_s - Z_s)Z_s' + \frac{1}{T}\sum_{s=1}^{T}k_{h_2,st}Z_s(\hat{Z}_s - Z_s)' \\
&= \bar{S}_{ZZ,t} + O_P(C_{NT}^{-1}), \qquad \text{(S1.3)}
\end{aligned}
$$

where

$$
\begin{aligned}
\bar{S}_{ZZ,t} &\equiv \frac{1}{T}\sum_{s=1}^{T}k_{h_2,st}Z_sZ_s' = \frac{1}{T}\sum_{s=1}^{T}k_{h_2,st}\mathbb{D}_sZ_s^0Z_s^{0\prime}\mathbb{D}_s' \\
&= \frac{1}{T}\sum_{s=1}^{T}k_{h_2,st}\mathbb{Q}_sZ_s^0Z_s^{0\prime}\mathbb{Q}_s' + O_P(C_{NT}^{-1}) = \mathbb{Q}_t\Sigma_{ZZ,t}\mathbb{Q}_t' + O_P(C_{NT}^{-1}). \qquad \text{(S1.4)}
\end{aligned}
$$

For $A_{t,1}^{Y}$, we have

$$
\begin{aligned}
A_{t,1}^{Y} &= \frac{h_2^{1/2}}{T^{1/2}}\sum_{s=1}^{T}k_{h_2,st}Z_s\varepsilon_{Y,s}' + \frac{h_2^{1/2}}{T^{1/2}}\sum_{s=1}^{T}k_{h_2,st}\left(\hat{Z}_s - Z_s\right)\varepsilon_{Y,s}' \\
&= \frac{h_2^{1/2}}{T^{1/2}}\sum_{s=1}^{T}k_{h_2,st}Z_s\varepsilon_{Y,s}' + O_P(C_{NT}^{-1}),
\end{aligned}
$$

where the last equation holds by Theorem 3.3 in Su and Wang (2017) and Theorem 2.3 in Su and Wang (2020b). For $A_{t,2}^{Y}$, we have by Lemma S1.1, $A_{t,2}^{Y} = T^{1/2}h_2^{5/2}\kappa_2\sum_{j=1}^{p}B_{t,j}^{Y} + o_P(1)$. Combining the above results yields

$$
\sqrt{Th_2}\left(\hat{\Psi}_{Y,t} - \Psi_{Y,t} - h_2^2\mathbf{B}_{Y,t}\right) = \hat{S}_{ZZ,t}^{-1}\frac{1}{\sqrt{Th_2}}\sum_{s=1}^{T}K\left(\frac{s-t}{Th_2}\right)Z_s\varepsilon_{Y,s}' + o_P(1),
$$

where $\mathbf{B}_{Y,t} = (\mathbb{Q}_t \Sigma_{ZZ,t} \mathbb{Q}_t')^{-1} \kappa_2 \sum_{j=1}^p B_{t,j}^Y$.

We now consider the $(K+R)p \times R$ matrix $\hat{\Psi}_{F,t}$.

$$\begin{aligned}
\sqrt{Th_2}\left(\hat{\Psi}_{F,t} - \Psi_{F,t}\right) &= \sqrt{Th_2}\hat{S}_{ZZ,t}^{-1}\left[\frac{1}{T}\sum_{s=1}^T k_{h_2,st}\hat{Z}_s u_{F,s}^{(t)\prime}\right] \\
&= \hat{S}_{ZZ,t}^{-1}\left[\frac{h_2^{1/2}}{T^{1/2}}\sum_{s=1}^T k_{h_2,st}\hat{Z}_s\varepsilon_{F,s}' H_s + \frac{h_2^{1/2}}{T^{1/2}}\sum_{s=1}^T k_{h_2,st}\hat{Z}_s \Delta_F'(s,t)\right] \\
&\equiv \hat{S}_{ZZ,t}^{-1}\left[A_{t,1}^F + A_{t,2}^F\right].
\end{aligned}$$

For $A_{t,1}^F$, we can apply (S2.3) in the next appendix to show that

$$\begin{aligned}
A_{t,1}^F &= \frac{h_2^{1/2}}{T^{1/2}}\sum_{s=1}^T k_{h_2,st} Z_s \varepsilon_{F,s}' H_s + \frac{h_2^{1/2}}{T^{1/2}}\sum_{s=1}^T k_{h_2,st}\left(\hat{Z}_s - Z_s\right)\varepsilon_{F,s}' H_s \\
&= \frac{h_2^{1/2}}{T^{1/2}}\sum_{s=1}^T k_{h_2,st} Z_s \varepsilon_{F,s}' H_s + o_P(1).
\end{aligned}$$

For $A_{t,2}^F$, we have by Lemma S1.2 that $A_{t,2}^F = T^{1/2} h_2^{5/2} \kappa_2 \sum_{j=1}^p B_{t,j}^F + o_P(1)$. Then

$$\sqrt{Th_2}\left(\hat{\Psi}_{F,t} - \Psi_{F,t} - h_2^2 \mathbf{B}_{F,t}\right) = \hat{S}_{ZZ,t}^{-1}\frac{1}{\sqrt{Th_2}}\sum_{s=1}^T K\left(\frac{s-t}{Th_2}\right) Z_s \varepsilon_{F,s}' H_s + o_P(1),$$

where $\mathbf{B}_{F,t} = (\mathbb{Q}_t \Sigma_{ZZ,t}\mathbb{Q}_t')^{-1}\kappa_2 \sum_{j=1}^p B_{t,j}^F$.

Now, let $\mathbf{B}_t = \left(\mathbf{B}_{Y,t}', \mathbf{B}_{F,t}'\right)'$ and $\tilde{\varepsilon}_s \equiv D_{H_s}'\varepsilon_s = (\varepsilon_{Y,s}', \varepsilon_{F,s}' H_s)'$, where $D_{H_s} \equiv \text{diag}(\mathbb{I}_K, H_s)$. Recall that $D_{Q_s^{-1}} \equiv \text{diag}(\mathbb{I}_K, Q_s^{-1})$. Then by Lemma A.1(iii) of Su and Wang (2017), $D_{H_s} = D_{Q_s^{-1}} + O_P(C_{NT}^{-1})$. Then

$$\begin{aligned}
\sqrt{Th_2}\left(\hat{\Psi}_t - \Psi_t - h_2^2\mathbf{B}_t\right) &= \hat{S}_{ZZ,t}^{-1}\frac{1}{\sqrt{Th_2}}\sum_{s=1}^T K\left(\frac{s-t}{Th_2}\right) Z_s \tilde{\varepsilon}_s' + o_P(1) \\
&= \hat{S}_{ZZ,t}^{-1}\frac{1}{\sqrt{Th_2}}\sum_{s=1}^T K\left(\frac{s-t}{Th_2}\right) \mathbb{D}_s Z_s^0 \varepsilon_s' D_{H_s} + o_P(1) \\
&= \hat{S}_{ZZ,t}^{-1}\frac{1}{\sqrt{Th_2}}\sum_{s=1}^T K\left(\frac{s-t}{Th_2}\right) \mathbb{Q}_s Z_s^0 \varepsilon_s' D_{Q_s^{-1}} + o_P(1),
\end{aligned}$$

where the last equality follows from an application of (S2.3) in Appendix B and Assumption A.2. Note that $\hat{S}_{ZZ,t} = \mathbb{Q}_t \Sigma_{ZZ,t}\mathbb{Q}_t' + O_P(C_{NT}^{-1})$ by (S1.3) and (S1.4). Under Assumption A.3, we can readily apply the martingale central limit theorem to obtain[1]

$$\text{vec}\left(\frac{1}{\sqrt{Th_2}}\sum_{s=1}^{T}K\left(\frac{s-t}{Th_2}\right)\mathbb{Q}_s Z_s^0\varepsilon_s' D_{Q_s^{-1}}\right) = \frac{1}{\sqrt{Th_2}}\sum_{s=1}^{T}K\left(\frac{s-t}{Th_2}\right)\left[D'_{Q_s^{-1}}\varepsilon_s\right]\otimes[\mathbb{Q}_s Z_s^0]$$
$$\xrightarrow{d} N(0,\Sigma_{Z\otimes\varepsilon}^{(t)}),$$

where

$$\begin{aligned}\Sigma_{Z\otimes\varepsilon}^{(t)} &= \lim_{T\to\infty}\frac{1}{Th_2}\sum_{s=1}^{T}K^2\left(\frac{s-t}{Th_2}\right)E\left[\left\{\left(D'_{Q_s^{-1}}\varepsilon_s\right)\otimes\left(\mathbb{Q}_s Z_s^0\right)\right\}\left\{\left(\varepsilon_s' D_{Q_s^{-1}}\right)\otimes\left(Z_s^{0\prime}\mathbb{Q}_s'\right)\right\}\right]\\ &= \lim_{T\to\infty}\frac{1}{Th_2}\sum_{s=1}^{T}K^2\left(\frac{s-t}{Th_2}\right)E\left\{\left[D'_{Q_s^{-1}}\varepsilon_s\varepsilon_s' D_{Q_s^{-1}}\right]\otimes[\mathbb{Q}_s Z_s^0 Z_s^{0\prime}\mathbb{Q}_s']\right\}.\end{aligned}$$

Consequently,

$$\begin{aligned}\sqrt{Th_2}\text{vec}\left(\hat{\Psi}_t-\Psi_t-h_2^2\mathbf{B}_t\right) &= \text{vec}\left(\hat{S}_{ZZ,t}^{-1}\frac{1}{\sqrt{Th_2}}\sum_{s=1}^{T}K\left(\frac{s-t}{Th_2}\right)\mathbb{Q}_s Z_s^0\varepsilon_s' D_{Q_s^{-1}}\right)+o_P(1)\\ &= \left[\mathbb{I}_{K+R}\otimes\hat{S}_{ZZ,t}^{-1}\right]\text{vec}\left(\frac{1}{\sqrt{Th_2}}\sum_{s=1}^{T}K\left(\frac{s-t}{Th_2}\right)\mathbb{Q}_s Z_s^0\varepsilon_s' D_{Q_s^{-1}}\right)+o_P(1)\\ &\xrightarrow{d} N\left(0,\left[\mathbb{I}_{K+R}\otimes(\mathbb{Q}_t\Sigma_{ZZ,t}\mathbb{Q}_t')^{-1}\right]\Sigma_{Z\otimes\varepsilon}^{(t)}\left[\mathbb{I}_{K+R}\otimes(\mathbb{Q}_t\Sigma_{ZZ,t}\mathbb{Q}_t')^{-1}\right]\right).\end{aligned}$$

This completes the proof of Theorem 3.1. ■

Proof of Theorem 4.1. The result in Theorem 4.1 follows as a special case of Theorem 4.2 with $a_{1NT}=a_{2T}=0$. We omit the proof for brevity. ■

Recall that in Section 4.4, we let that $D_H=\text{diag}(\mathbb{I}_K,H)$, $D_{Q^{-1}}=\text{diag}(\mathbb{I}_K,Q^{-1})$, $\mathbb{D}=\mathbb{I}_p\otimes D_H'$, $\mathbb{Q}=\mathbb{I}_p\otimes D'_{Q^{-1}}$, $\varepsilon_t^\dagger=D_H'\varepsilon_t$, and $Z_t^\dagger=\mathbb{D}Z_t^0$, where $Z_t^0=(Y_{t-1}',F_{t-1}',\dots,Y_{t-p}',F_{t-p}')'$. Together with the notations that we introduced in Section 4.2, we define

$$\mathbb{B}_{3T}^{(1)}=\frac{h_2^{1/2}}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}k_{h_2,st}^{\dagger 2}Z_s^{\dagger\prime}\tilde{S}_{ZZ,t}^{-1}\tilde{S}_{ZZ,t}^{-1}Z_s^\dagger\varepsilon_s^{\dagger\prime}\varepsilon_s^\dagger,\ \mathbb{B}_{3T}^{(2)}=\frac{h_2^{1/2}}{T}\sum_{s=1}^{T}Z_s^{\dagger\prime}\tilde{S}_{ZZ}^{-1}\tilde{S}_{ZZ}^{-1}Z_s^\dagger\varepsilon_s^{\dagger\prime}\varepsilon_s^\dagger,\ \text{and}$$
$$\mathbb{B}_{3T}^{(3)}=\frac{h_2^{1/2}}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}k_{h_2,st}^{\dagger}Z_s^{\dagger\prime}\tilde{S}_{ZZ,t}^{-1}\tilde{S}_{ZZ}^{-1}Z_s^\dagger\varepsilon_s^{\dagger\prime}\varepsilon_s^\dagger.$$

[1]Noting that $\text{vec}(ABC)=(C'\otimes A)\text{vec}(B)$ for conformable matrices A, B and C, the left hand side can also be written as $\frac{h_2^{1/2}}{T^{1/2}}\sum_{s=1}^{T}k_{h_2,st}(\text{diag}(\mathbb{I}_K,Q_s'^{-1})\otimes\mathbb{Q}_s)\text{vec}(Z_s^0\varepsilon_s')$.

Let $\mathbb{B}_{3T}$ and $\mathbb{V}_{3T}$ be as defined as in Section 4.4. We have that $\mathbb{B}_{3T} = \mathbb{B}_{3T}^{(1)} + \mathbb{B}_{3T}^{(2)} - 2\mathbb{B}_{3T}^{(3)}$. The following five propositions are used in the proof of Theorem 4.2.

Proposition S1.1. *Suppose that the conditions in Theorem 4.2 Hold. Then* $M_{31} - \mathbb{B}_{3T}^{(1)} - \Pi_{3T} \xrightarrow{d} N(0, \mathbb{V}_{3T})$ *under* $\mathbb{H}_A^{(3)}(a_{2T})$.

Proposition S1.2. *Suppose that the conditions in Theorem 4.2 Hold. Then* $M_{32} - \mathbb{B}_{3T}^{(2)} = o_P(1)$ *under* $\mathbb{H}_A^{(3)}(a_{2T})$.

Proposition S1.3. *Suppose that the conditions in Theorem 4.2 Hold. Then* $M_{33} - \mathbb{B}_{3T}^{(3)} = o_P(1)$ *under* $\mathbb{H}_A^{(3)}(a_{2T})$.

Proposition S1.4. *Suppose that the conditions in Theorem 4.2 Hold. Then* $\hat{\mathbb{B}}_{3T} = \mathbb{B}_{3T} + o_P(1)$ *under* $\mathbb{H}_A^{(2)}(a_{2T})$.

Proposition S1.5. *Suppose that the conditions in Theorem 4.2 Hold. Then* $\hat{\mathbb{V}}_{3T} = \mathbb{V}_{3T} + o_P(1)$ *under* $\mathbb{H}_A^{(3)}(a_{2T})$.

Proof of Theorem 4.2. Since $\widehat{SM}_1 = 2^{-1/2}(\widehat{SM}_2 + \widehat{SM}_3)$, we first show (ii) and (iii), and then show (i). We note that (ii) holds according to Theorem 3.4 in Su and Wang (2020a).

We then consider (iii). Under $\mathbb{H}_A^{(3)}(a_{2T}) : \Psi_t = \Psi_0 + a_{2T} g_2(t/T)$, we decompose

$$\begin{aligned}
Th_2^{1/2}\hat{M}_3 &= h_2^{1/2}\sum_{t=1}^{T}\left\|\breve{\Psi}_t - \tilde{\Psi}_0\right\|^2 \\
&= h_2^{1/2}\sum_{t=1}^{T}\left\|\left(\breve{\Psi}_t - \Psi_0\right) - \left(\tilde{\Psi}_0 - \Psi_0\right)\right\|^2 \\
&= h_2^{1/2}\sum_{t=1}^{T}\left\|\breve{\Psi}_t - \Psi_0\right\|^2 + h_2^{1/2}\sum_{t=1}^{T}\left\|\tilde{\Psi}_0 - \Psi_0\right\|^2 - 2h_2^{1/2}\sum_{t=1}^{T}\mathrm{tr}\left[\left(\breve{\Psi}_t - \Psi_0\right)\left(\tilde{\Psi}_0 - \Psi_0\right)'\right] \\
&\equiv M_{31} + M_{32} - 2M_{33}.
\end{aligned}$$

Note that (iii) holds if $\mathbb{H}_A^{(3)}(a_{2T})$, (iii.1) $\mathbb{V}_{3T}^{-1/2}(M_{31} - \mathbb{B}_{3T}^{(1)} - \Pi_{3T}) \xrightarrow{d} N(0,1)$, (iii.2) $M_{32} - \mathbb{B}_{3T}^{(2)} = o_P(1)$, (iii.3) $M_{33} - \mathbb{B}_{3T}^{(3)} = o_P(1)$, (iii.4) $\hat{\mathbb{B}}_{3T} = \mathbb{B}_{3T} + o_P(1)$, and (iii.5) $\hat{\mathbb{V}}_{3T} = \mathbb{V}_{3T} + o_P(1)$. The claims (iii.1)–(iii.5) are established by Propositions S1.1–S1.5, respectively. Combining these results yields $\widehat{SM}_3 = \hat{\mathbb{V}}_{3T}^{-1/2}(Th_2^{1/2}\hat{M}_3 - \hat{\mathbb{B}}_{3T}) \xrightarrow{d} N(\pi_3, 1)$ where $\pi_3 = \lim_{T\to\infty}\Pi_{3T}/\mathbb{V}_{3T}^{1/2}$.

In the end, we show (i). By the proof of Theorem 3.4 in Su and Wang (2020a) and the result in part (ii), we have

$$\widehat{SM}_2 \;=\; \hat{\mathbb{V}}_{2T}^{-1/2}(TN^{1/2}h_1^{1/2}\hat{M}_2 - \hat{\mathbb{B}}_{2NT})$$

$$
\begin{aligned}
&= \mathbb{V}_{2NT}^{-1/2}\left[\frac{2}{TN^{1/2}h_1^{1/2}}\sum_{1\le r<s\le T}\sum_{i=1}^{N}\bar{K}\left(\frac{s-r}{Th_1}\right)F_s'\Sigma_F^{-1}Q'Q\Sigma_F^{-1}F_r e_{ir}e_{is}\right]+o_P(1)\\
&\equiv \mathbf{V}_1+o_P(1),
\end{aligned}
$$

and

$$
\begin{aligned}
\widehat{SM}_3 &= \hat{\mathbb{V}}_{3T}^{-1/2}(Th_2^{1/2}\hat{M}_3-\hat{\mathbb{B}}_{3T})\\
&= \mathbb{V}_{3T}^{-1/2}\left[\frac{2}{Th_2^{1/2}}\sum_{1\le r<s\le T}\bar{K}\left(\frac{s-r}{Th_2}\right)Z_s^{0\prime}\mathbb{S}_T Z_r^0\varepsilon_r' D_{Q^{-1}}D_{Q^{-1}}'\varepsilon_s\right]+o_P(1)\\
&\equiv \mathbf{V}_2+o_P(1),
\end{aligned}
$$

where $E(\mathbf{V}_\ell)=0$ and $\text{Var}(\mathbf{V}_\ell)=1$ for $\ell=1,2$. It follows that

$$
\begin{aligned}
\widehat{SM}_1 &= \widehat{SM}_2+\widehat{SM}_3=\hat{\mathbb{V}}_{2NT}^{-1/2}(TN^{1/2}h_1^{1/2}\hat{M}_2-\hat{\mathbb{B}}_{2NT})+\hat{\mathbb{V}}_{3T}^{-1/2}(Th_2^{1/2}\hat{M}_3-\hat{\mathbb{B}}_{3T})\\
&= \mathbf{V}_1+\mathbf{V}_2+o_P(1),
\end{aligned}
$$

where $\text{Var}(\mathbf{V}_1+\mathbf{V}_2)=2+2E(\mathbf{V}_1\mathbf{V}_2)$. By Assumption A.3(iii)–(iv), we have

$$
\begin{aligned}
E(V_1V_2) &= \mathbb{V}_{2NT}^{-1/2}\mathbb{V}_{3T}^{-1/2}\frac{4}{T^2N^{1/2}h_1^{1/2}h_2^{1/2}}\sum_{1\le r_1<s_1\le T}^{T}\sum_{1\le r_2<s_2\le T}^{T}\sum_{i=1}^{N}\bar{K}\left(\frac{s_1-r_1}{Th_2}\right)\bar{K}\left(\frac{s_2-r_2}{Th_1}\right)\\
&\quad\times E\left[Z_{s_1}^{0\prime}\mathbb{S}_T Z_{r_1}^0 F_{s_2}'\Sigma_F^{-1}Q'Q\Sigma_F^{-1}F_{r_2}\varepsilon_{r_1}'D_{Q^{-1}}D_{Q^{-1}}'\varepsilon_{s_1}e_{ir_2}e_{is_2}\right]\\
&= \mathbb{V}_{2NT}^{-1/2}\mathbb{V}_{3T}^{-1/2}\frac{4}{T^2N^{1/2}h_1^{1/2}h_2^{1/2}}\sum_{1\le r_1<s_1\le T}^{T}\sum_{1\le r_2<s_2\le T}^{T}\sum_{i=1}^{N}\bar{K}\left(\frac{s_1-r_1}{Th_2}\right)\bar{K}\left(\frac{s_2-r_2}{Th_1}\right)\\
&\quad\times E\left[Z_{s_1}^{0\prime}\mathbb{S}_T Z_{r_1}^0 F_{s_2}'\Sigma_F^{-1}Q'Q\Sigma_F^{-1}F_{r_2}\varepsilon_{r_1}'D_{Q^{-1}}D_{Q^{-1}}'\varepsilon_{s_1}\right]E(e_{ir_2}e_{is_2})=0.
\end{aligned}
$$

Then $\text{Var}(\mathbf{V}_1+\mathbf{V}_2)=2$. Since $\mathbf{V}_1$ and $\mathbf{V}_2$ are both asymptotically standard normally distributed, we can readily apply the Cramér-Wold device to conclude that $\widehat{SM}_1\xrightarrow{d}N(0,1)$. ■

Proof of Theorem 4.3. The result that $\widehat{SM}_2^{*}\xrightarrow{d^*}N(0,1)$ essentially follows similar arguments as used in Su and Wang (2020a) under a slightly different bootstrap scheme. The major difference lies in the generation of the error terms used in the bootstrap world. Following Su and Wang (2020a), we can establish the uniform convergence rates for the estimators $\tilde{\lambda}_{i0}$ and $\tilde{F}_t$ under either the null or local alternatives. This ensures that $\max_{i\in[N]}\frac{1}{T}\sum_{t=1}^{T}(\tilde{e}_{it}-e_{it})^2=O_P(a_{NT})$ for some $a_{NT}=o(1)$ and $\max_{i\in[N]}\max_{t\in[T]}|\tilde{e}_{it}-e_{it}|=o_P(1)$. Then we can readily follow Fan et al. (2013) and show that $\left\|\tilde{\Sigma}_e - \Sigma_e\right\|_{sp} = O_P(a_{NT}^{1-\gamma_0/2}) = o_P(1)$. As a result, $\left\|\tilde{\Sigma}_e\right\|_{sp} \leq \|\Sigma_e\|_{sp} + \left\|\tilde{\Sigma}_e - \Sigma_e\right\|_{sp} = O_P(1)$. With this result, we can follow Su and Wang (2020a) and justify the distributional result for $\widehat{SM}_2^*$. Below we outline the proof that $\widehat{SM}_3^* \xrightarrow{d^*} N(0,1)$. Then by the asymptotic independence between $\widehat{SM}_2^*$ and $\widehat{SM}_3^*$ conditional on the data $\mathcal{W}_{NT} = \{X_{it}, Y_t\}_{i\in[N],-p+1\leq t\leq T}$, we have $\widehat{SM}_1^* \xrightarrow{d^*} N(0,1)$.

Let P^* denote the probability measure induced by the bootstrap conditional on $\mathcal{W}_{NT}$. Let E^* and Var^* denote the expectation and variance under P^*. Let O_{P^*} and o_{P^*} denote the probability order under P^*, e.g., $b_{NT} = o_{P^*}(1)$ if for any $\epsilon > 0$, $P^*(\|b_{NT}\| > \epsilon) = o_P(1)$. In addition, let $\tilde{\lambda}_{i0}^*$, $\tilde{F}_t^*$, $\tilde{\lambda}_{it}^*$, $\tilde{\Psi}_0^*$, and $\breve{\Psi}_t^*$ denote the bootstrap analogue of $\tilde{\lambda}_{i0}$, $\tilde{F}_t$, $\tilde{\lambda}_{it}$, $\tilde{\Psi}_0$, and $\breve{\Psi}_t$, respectively. Let $\hat{M}_3^*$, $\widehat{SM}_3^*$, $\mathbb{B}_{3T}^{(l*)}$, $\mathbb{B}_{3T}^*$, $\mathbb{V}_{3T}^*$, $\hat{\mathbb{B}}_{3T}^*$ and $\hat{\mathbb{V}}_{3T}^*$ denote the bootstrap analogues of $\hat{M}_3$, $\widehat{SM}_3$, $\mathbb{B}_{3T}^{(l)}$, $\mathbb{B}_{3T}$, $\mathbb{V}_{3T}$, $\hat{\mathbb{B}}_{3T}^*$ and $\hat{\mathbb{V}}_{3T}^*$, respectively, where $l = 1,2,3$. Then $\widehat{SM}_3^* = (Th_2^{1/2}\hat{M}_3^* - \hat{\mathbb{B}}_{3T}^*)/\sqrt{\hat{\mathbb{V}}_{3T}^*}$. Let $SM_3^* \equiv (Th_2^{1/2}\hat{M}_3^* - \mathbb{B}_{3T}^*)/\sqrt{\mathbb{V}_{3T}^*}$. As in the proof of Theorem 4.2, we make the following decomposition:

$$
\begin{aligned}
Th_2^{1/2}\hat{M}_3^* &= h_2^{1/2}\sum_{t=1}^T \left\|\breve{\Psi}_t^* - \tilde{\Psi}_0^*\right\|^2 \\
&= h_2^{1/2}\sum_{t=1}^T \left\|\breve{\Psi}_t^* - \tilde{\Psi}_0\right\|^2 + h_2^{1/2}\sum_{t=1}^T \left\|\tilde{\Psi}_0^* - \tilde{\Psi}_0\right\|^2 - 2h_2^{1/2}\sum_{t=1}^T \text{tr}\left[(\breve{\Psi}_t^* - \tilde{\Psi}_0)(\tilde{\Psi}_0^* - \tilde{\Psi}_0)'\right] \\
&\equiv M_{31}^* + M_{32}^* - 2M_{33}^*.
\end{aligned}
$$

Noting that $\tilde{W}_t^* = \tilde{\Psi}_0'\tilde{Z}_t + U_t^*$, we have

$$
\breve{\Psi}_t^* - \tilde{\Psi}_0 = \left(\frac{1}{T}\sum_{s=1}^T k_{h_2,st}^\dagger \tilde{Z}_s\tilde{Z}_s'\right)^{-1}\frac{1}{T}\sum_{s=1}^T k_{h_2,st}^\dagger \tilde{Z}_s\tilde{W}_s^{*\prime} - \tilde{\Psi}_0 = \tilde{S}_{ZZ,t}^{-1}\frac{1}{T}\sum_{s=1}^T k_{h_2,st}^\dagger \tilde{Z}_s U_s^{*\prime} \equiv D_{1t}^*.
$$

Following the proof of Proposition S1.1, we can readily show that $M_{31}^* - \mathbb{B}_{3T}^{(1*)} = 2\sum_{s=2}^T \mathcal{Z}_{T,s}^* + o_{P^*}(1)$, where $\mathcal{Z}_{T,s}^* = T^{-1}h_2^{-1/2}\sum_{r=1}^{s-1}\bar{k}_{sr}\tilde{Z}_s'\tilde{S}_{ZZ,t}^{-1}\tilde{S}_{ZZ,t}^{-1}\tilde{Z}_r U_s^{*\prime}U_r^*$. Let $\mathcal{F}_{NT,t}^* = \sigma(\mathcal{W}_{NT}, v_t, v_{t-1}, \ldots)$, the minimal σ-field formed by $\mathcal{W}_{NT}$, v_t and v_t's history. Then $E^*\left(\mathcal{Z}_{T,s}^*|\mathcal{F}_{NT,s-1}^*\right) = T^{-1}h_2^{-1/2}\sum_{r=1}^{s-1}\bar{k}_{sr}\tilde{Z}_s'\tilde{S}_{ZZ,t}^{-1}\tilde{S}_{ZZ,t}^{-1}\tilde{Z}_r U_r^{*\prime}E^*(U_s^*|\mathcal{F}_{NT,s-1}^*) = 0$ and

$$
\begin{aligned}
\text{Var}^*\left(2\sum_{s=2}^T \mathcal{Z}_{T,s}^*\right) &= 4T^{-2}h_2^{-1}\sum_{s=2}^T\sum_{r=1}^{s-1}\bar{k}_{sr}^2\left(\tilde{Z}_s'\tilde{S}_{ZZ,t}^{-1}\tilde{S}_{ZZ,t}^{-1}\tilde{Z}_r\right)^2 E^*(U_r^{*\prime}U_s^*)^2 \\
&= 4T^{-2}h_2^{-1}\sum_{s=2}^T\sum_{r=1}^{s-1}\bar{k}_{sr}^2\left(\tilde{Z}_s'\tilde{S}_{ZZ,t}^{-1}\tilde{S}_{ZZ,t}^{-1}\tilde{Z}_r\right)^2 \text{tr}(\tilde{\Sigma}_U\tilde{\Sigma}_U) \equiv \mathbb{V}_{3T}^*.
\end{aligned}
$$

One can readily show that $(\mathbb{V}_{3T}^{*})^{-1/2}\left(M_{31}^{*}-\mathbb{B}_{3T}^{(1*)}\right)=2\sum_{s=2}^{T}\mathcal{Z}_{T,s}^{*}+o_{P^*}(1)\overset{d^*}{\to}N(0,1)$ by the martingale CLT. Noting that

$$\tilde{\Psi}_0^{*}-\tilde{\Psi}_0=\left(\frac{1}{T}\sum_{s=1}^{T}\tilde{Z}_s\tilde{Z}_s'\right)^{-1}\frac{1}{T}\sum_{s=1}^{T}\tilde{Z}_s\tilde{W}_s^{*\prime}-\tilde{\Psi}_0=\tilde{S}_{ZZ}^{-1}\frac{1}{T}\sum_{s=1}^{T}\tilde{Z}_sU_s^{*\prime}\equiv D_1^{*},$$

it is trivial to show that

$$\begin{aligned}M_{32}^{*} &= \frac{h_2^{1/2}}{T}\sum_{s=1}^{T}\tilde{Z}_s'\tilde{S}_{ZZ}^{-1}\tilde{S}_{ZZ}^{-1}\tilde{Z}_sU_s^{*\prime}U_s^{*}+o_{P^*}(1)\equiv B_{3T}^{(2*)}+o_{P^*}(1)\,,\text{ and}\\ M_{33}^{*} &= h_2^{1/2}\sum_{t=1}^{T}\mathrm{tr}(D_{1t}^{*}D_1^{*\prime})=\frac{h_2^{1/2}}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}k_{h_2,st}^{\dagger}\tilde{Z}_s'\tilde{S}_{ZZ}^{-1}\tilde{S}_{ZZ,t}^{-1}\tilde{Z}_sU_s^{*\prime}U_s^{*}+o_{P^*}(1)\equiv B_{3T}^{(3*)}+o_{P^*}(1)\,.\end{aligned}$$

In fact both $B_{3T}^{(2*)}$ and $B_{3T}^{(3*)}$ are $O_{P^*}(h_2^{1/2})$ and thus asymptotically negligible. Consequently, we have shown that

$$SM_3^{*}\equiv(\mathbb{V}_{3T}^{*})^{-1/2}\left(Th_2^{1/2}\hat{M}_3^{*}-\mathbb{B}_{3T}^{*}\right)=2\sum_{s=2}^{T}\mathcal{Z}_{T,s}^{*}+o_{P^*}(1)\overset{d^*}{\to}N(0,1)$$

where $\mathbb{B}_{3T}^{*}=B_{3T}^{(1*)}+B_{3T}^{(2*)}-2B_{3T}^{(3*)}=\frac{h_2^{1/2}}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\tilde{Z}_s'(k_{h_2,st}^{\dagger}\tilde{S}_{ZZ,t}^{-1}-\tilde{S}_{ZZ}^{-1})(k_{h_2,st}^{\dagger}\tilde{S}_{ZZ,t}^{-1}-\tilde{S}_{ZZ}^{-1})\tilde{Z}_sU_s^{*\prime}U_s^{*}$.

Next, note that the bootstrap analogoues of $\hat{\mathbb{B}}_{3T}$ and $\hat{\mathbb{V}}_{3T}$ are respectively given by

$$\begin{aligned}\hat{\mathbb{B}}_{3T}^{*} &= \frac{h_2^{1/2}}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\tilde{Z}_s'(k_{h_2,st}^{\dagger}\tilde{S}_{ZZ,t}^{-1}-\tilde{S}_{ZZ}^{-1})(k_{h_2,st}^{\dagger}\tilde{S}_{ZZ,t}^{-1}-\tilde{S}_{ZZ}^{-1})\tilde{Z}_s\tilde{U}_s^{*\prime}\tilde{U}_s^{*}\text{ and}\\ \hat{\mathbb{V}}_{3T}^{*} &= \frac{4}{T^2h_2}\sum_{s=2}^{T}\sum_{r=1}^{s-1}\bar{k}_{sr}^{2}\left(\tilde{Z}_s'\tilde{S}_{ZZ,t}^{-1}\tilde{S}_{ZZ,t}^{-1}\tilde{Z}_r\right)^2(\tilde{U}_r^{*\prime}\tilde{U}_s^{*})^2,\end{aligned}$$

where $\tilde{U}_s^{*}\equiv\tilde{W}_s^{*}-\tilde{\Psi}_0^{*\prime}\tilde{Z}_s$. Following the proofs of Propositions S1.4–S1.5, we can readily show that $\hat{\mathbb{B}}_{3T}^{*}=\mathbb{B}_{3T}^{*}+o_{P^*}(1)$ and $\hat{\mathbb{V}}_{3T}^{*}=\mathbb{V}_{3T}^{*}+o_{P^*}(1)$. Then $\widehat{SM}_3^{*}\overset{d^*}{\to}N(0,1)\,.$ ■

S2 Proofs of the Technical Lemmas and Propositions in Appendix S1

To prove the technical lemmas and propositions in Appendix A, we frequently call upon the following lemma.

Lemma S2.1. *Let* $\{V_t, t \geq 1\}$ *be a strong mixing process with mixing coefficient* $\alpha(\cdot)$. *Let* $G_{t_1,...,t_m}(v_1,...,v_m)$ *denote the cumulative 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*

$$\max_{1\leq j\leq m} \left\{ \int |\vartheta(v_1,\ldots,v_m)|^{1+\tilde{\eta}} \mathrm{d}G_{t_1,\ldots,t_j}(v_1,\ldots,v_j) \mathrm{d}\mathfrak{G}_{t_{j+1},\ldots,t_m}(v_{j+1},\ldots,v_m), \int |\vartheta(v_1,\ldots,v_m)|^{1+\tilde{\eta}} \mathrm{d}G_{t_1,\ldots,t_m}(v_1,\ldots,v_m) \right\} \leq C,$$

for some $\tilde{\eta} > 0$. *Then* $|\int \vartheta(v_1,\ldots,v_m) \mathrm{d}G_{t_1,\ldots,t_m}(v_1,\ldots,v_m) - \int \vartheta(v_1,\ldots,v_m) \ \mathrm{d}G_{t_1,\ldots,t_j}$ $(v_1,\ldots,v_j) \mathrm{d}G_{t_{j+1},\ldots,t_m}(v_{j+1}, \ldots, v_m)| \leq 4M^{1/(1+\tilde{\eta})}\alpha(t_{j+1}-t_j)^{\tilde{\eta}/(1+\tilde{\eta})}$.

For the proof of the Lemma S2.1, see Sun and Chiang (1997).

Proof of Lemma S1.1. Note that

$$\begin{aligned} A_{t,2}^Y &= \frac{h_2^{1/2}}{T^{1/2}} \sum_{s=1}^T k_{h_2,st} \hat{Z}_s \Delta_Y'(s,t) \\ &= \frac{h_2^{1/2}}{T^{1/2}} \sum_{s=1}^T \sum_{j=1}^p k_{h_2,st} \hat{Z}_s \left(H_{s-j}'F_{s-j} - \hat{F}_{s-j}\right)' \psi_j^{(1,2)}\left(\frac{s}{T}\right)' \\ &+ \frac{h_2^{1/2}}{T^{1/2}} \sum_{s=1}^T \sum_{j=1}^p k_{h_2,st} \hat{Z}_s \hat{F}_{s-j}' \left[\psi_j^{(1,2)}\left(\frac{s}{T}\right) - \psi_j^{(1,2)}\left(\frac{t}{T}\right)\right]' \\ &+ \frac{h_2^{1/2}}{T^{1/2}} \sum_{s=1}^T \sum_{j=1}^p k_{h_2,st} \hat{Z}_s Y_{s-j}' \left[\psi_j^{(1,1)}\left(\frac{s}{T}\right) - \psi_j^{(1,1)}\left(\frac{t}{T}\right)\right]' \equiv A_{t,21}^Y + A_{t,22}^Y + A_{t,23}^Y. \end{aligned}$$

Consider $A_{t,21}^Y$.

$$\begin{aligned} \left\|A_{t,21}^Y\right\| &= \frac{h_2^{1/2}}{T^{1/2}} \left\| \sum_{j=1}^p \sum_{s=1}^T k_{h_2,st} \hat{Z}_s \left(\hat{F}_{s-j} - H_{s-j}'F_{s-j}\right)' \psi_j^{(1,2)}\left(\frac{s}{T}\right)' \right\| \\ &\leq T^{1/2} h_2^{1/2} \max_{-p+1\leq s\leq T} \left\|\hat{F}_s - H_s'F_s\right\| \frac{1}{T} \sum_{j=1}^p \sum_{s=1}^T k_{h_2,st} \left\|\hat{Z}_s\right\| \left\|\psi_j^{(1,2)}\left(\frac{s}{T}\right)\right\| \end{aligned}$$

With the proof of Theorem 3.5 in Su and Wang (2017) by taking into the bias terms as in Su and Wang (2020b), we can show that

$$\max_{-p+1\leq s\leq T} \left\|\hat{F}_s - H_s'F_s\right\| = O_P\left((N/\ln T)^{-1/2} + h_1^2\right). \tag{S2.1}$$

In addition, $T^{-1}\sum_{j=1}^{p}\sum_{s=1}^{T}k_{h_2,st}\left\|\hat{Z}_s\right\|\left\|\psi_{js}^{(1,2)}\right\|=O_P\left(1\right).$ It follows that

$$\left\|A_{t,21}^{Y}\right\|=T^{1/2}h_2^{1/2}O_P\left((N/\ln T)^{-1/2}+h_1^2\right)O_P\left(1\right)=o_P(1),$$

where we use the conditions that $Th_2\ln T/N=o\left(1\right)$ and $Th_1^4h_2=o\left(1\right)$ under Assumption A.2(iii).

For $A_{t,22}^{Y}$, we have

$$\begin{aligned}A_{t,22}^{Y} &= \frac{h_2^{1/2}}{T^{1/2}}\sum_{s=1}^{T}\sum_{j=1}^{p}k_{h_2,st}\hat{Z}_sF_{s-j}'H_{s-j}\left[\psi_j^{(1,2)}\left(\frac{s}{T}\right)-\psi_j^{(1,2)}\left(\frac{t}{T}\right)\right]' \\ &\quad+\frac{h_2^{1/2}}{T^{1/2}}\sum_{s=1}^{T}\sum_{j=1}^{p}k_{h_2,st}\hat{Z}_s\left(\hat{F}_{s-j}-H_{s-j}'F_{s-j}\right)'\left[\psi_j^{(1,2)}\left(\frac{s}{T}\right)-\psi_j^{(1,2)}\left(\frac{t}{T}\right)\right]'\equiv A_{t,221}^{Y}+A_{t,222}^{Y}.\end{aligned}$$

It is easy to show that $A_{t,222}^{Y}$ is of smaller order than $A_{t,221}^{Y}$. So we focus on $A_{t,221}^{Y}$. Note that

$$\begin{aligned}A_{t,221}^{Y} &= \frac{h_2^{1/2}}{T^{1/2}}\sum_{s=1}^{T}\sum_{j=1}^{p}k_{h_2,st}\hat{Z}_sF_{s-j}'H_{s-j}\left(\phi_j^{(1,2)}\left(\frac{s}{T}\right)H_{s-j}^{-1\prime}-\phi_j^{(1,2)}\left(\frac{t}{T}\right)H_{t-j}^{-1\prime}\right)' \\ &= \frac{h_2^{1/2}}{T^{1/2}}\sum_{s=1}^{T}\sum_{j=1}^{p}k_{h_2,st}\hat{Z}_sF_{s-j}'\left[\phi_j^{(1,2)}\left(\frac{s}{T}\right)-\phi_j^{(1,2)}\left(\frac{t}{T}\right)\right]' \\ &\quad+\frac{h_2^{1/2}}{T^{1/2}}\sum_{s=1}^{T}\sum_{j=1}^{p}k_{h_2,st}\hat{Z}_sF_{s-j}'H_{s-j}\left(H_{s-j}^{-1}-H_{t-j}^{-1}\right)\phi_j^{(1,2)}\left(\frac{t}{T}\right)' \\ &= \frac{h_2^{1/2}}{T^{1/2}}\sum_{s=1}^{T}\sum_{j=1}^{p}k_{h_2,st}Z_sF_{s-j}'\left[\phi_j^{(1,2)}\left(\frac{s}{T}\right)-\phi_j^{(1,2)}\left(\frac{t}{T}\right)\right]' \\ &\quad+\frac{h_2^{1/2}}{T^{1/2}}\sum_{s=1}^{T}\sum_{j=1}^{p}k_{h_2,st}\left(\hat{Z}_s-Z_s\right)F_{s-j}'\left[\phi_j^{(1,2)}\left(\frac{s}{T}\right)-\phi_j^{(1,2)}\left(\frac{t}{T}\right)\right]' \\ &\quad+\frac{h_2^{1/2}}{T^{1/2}}\sum_{s=1}^{T}\sum_{j=1}^{p}k_{h_2,st}Z_sF_{s-j}'H_{s-j}\left(H_{s-j}^{-1}-H_{t-j}^{-1}\right)\phi_j^{(1,2)}\left(\frac{t}{T}\right)' \\ &\quad+\frac{h_2^{1/2}}{T^{1/2}}\sum_{s=1}^{T}\sum_{j=1}^{p}k_{h_2,st}\left(\hat{Z}_s-Z_s\right)F_{s-j}'H_{s-j}\left(H_{s-j}^{-1}-H_{t-j}^{-1}\right)\phi_j^{(1,2)}\left(\frac{t}{T}\right)' \\ &\equiv A_{t,2211}^{Y}+A_{t,2212}^{Y}+A_{t,2213}^{Y}+A_{t,2214}^{Y}.\end{aligned}$$

For $A^Y_{t,2211}$, we apply the Taylor expansion to each element of $\phi^{(1,2)}_{js}$ obtain

$$\begin{aligned}
A^Y_{t,2211} &= \frac{h_2^{1/2}}{T^{1/2}} \sum_{s=1}^{T} \sum_{j=1}^{p} k_{h_2,st} Z_s F'_{s-j} \left[\phi_j^{(1,2)} \left(\frac{s}{T} \right) - \phi_j^{(1,2)} \left(\frac{t}{T} \right) \right]' \\
&= \frac{h_2^{1/2}}{T^{1/2}} \sum_{s=1}^{T} \sum_{j=1}^{p} k_{h_2,st} \mathbb{D}_s Z_s^0 F'_{s-j} \left[h_2 \left(\mathrm{d}\phi_{jt}^{(1,2)} \right)' \left(\frac{s-t}{Th_2} \right) \right] \\
&\quad + \frac{h_2^{1/2}}{T^{1/2}} \sum_{s=1}^{T} \sum_{j=1}^{p} k_{h_2,st} \mathbb{D}_s Z_s^0 F'_{s-j} \left[\frac{1}{2} h_2^2 \left(\mathrm{d}^2\phi_{jt}^{(1,2)} \right)' \left(\frac{s-t}{Th_2} \right)^2 \right] + o_P(1) \\
&\equiv A^Y_{t,22111} + A^Y_{t,22112} + o_P(1),
\end{aligned}$$

where $Z_s = \mathbb{D}_s Z_s^0$ with $\mathbb{D}_s = \mathrm{diag}\big(\mathbb{I}_K, H'_{s-1}, ..., \mathbb{I}_K, H'_{s-p}\big)$. We now show that both $A^Y_{t,22111}$ and $A^Y_{t,22112}$ contribute to the asymptotic bias term because $\Sigma_{ZF,js} \equiv \Sigma_{ZF,j}(s/T) \equiv E(Z_s^0 F'_{s-j})$ is typically s-dependent in the time-varying FAVAR model. Noting that $H_t = Q_t^{-1} + O_P(C_{NT}^{-1})$ and $\mathbb{D}_t = \mathbb{Q}_t + O_P(C_{NT}^{-1})$, we can show that

$$\begin{aligned}
A^Y_{t,22111} &= \frac{h_2^{3/2}}{T^{1/2}} \sum_{s=1}^{T} \sum_{j=1}^{p} k_{h_2,st} \mathbb{Q}_s \Sigma_{ZF,js} \left(\mathrm{d}\phi_{jt}^{(1,2)} \right)' \left(\frac{s-t}{Th_2} \right) + o_P(1) \\
&= \frac{h_2^{3/2}}{T^{1/2}} \sum_{s=1}^{T} \sum_{j=1}^{p} k_{h_2,st} \mathbb{Q}_t \Sigma_{ZF,jt} \left(\mathrm{d}\phi_{jt}^{(1,2)} \right)' \left(\frac{s-t}{Th_2} \right) \\
&\quad + \frac{h_2^{3/2}}{T^{1/2}} \sum_{s=1}^{T} \sum_{j=1}^{p} k_{h_2,st} \left[\mathbb{Q}_s \Sigma_{ZF,js} - \mathbb{Q}_t \Sigma_{ZF,jt} \right] \left(\mathrm{d}\phi_{jt}^{(1,2)} \right)' \left(\frac{s-t}{Th_2} \right) + o_P(1),
\end{aligned}$$

where the first term is of order $T^{1/2} h_2^{3/2} O\left(\int \tau K(\tau) \mathrm{d}\tau + (Th_2)^{-1} \right) = o(1)$ by the property of Riemann sum approximation of an integral and the fact that $\int \tau K(\tau) \mathrm{d}\tau = 0$, and the second term is equal to

$$\begin{aligned}
& \frac{h_2^{5/2}}{T^{1/2}} \sum_{s=1}^{T} \sum_{j=1}^{p} k_{h_2,st} \mathrm{d}\left(\mathbb{Q}_t \Sigma_{ZF,jt} \right) \left(\mathrm{d}\phi_{jt}^{(1,2)} \right)' \left(\frac{s-t}{Th_2} \right)^2 + o(1) \\
= \;& T^{1/2} h_2^{5/2} \sum_{j=1}^{p} \mathrm{d}\left(\mathbb{Q}_t \Sigma_{ZF,jt} \right) \left(\mathrm{d}\phi_{jt}^{(1,2)} \right)' \kappa_2 + o(1).
\end{aligned}$$

where $\kappa_2 \equiv \int \tau^2 K(\tau) \mathrm{d}\tau$. Then $A^Y_{t,22111} = T^{1/2} h_2^{5/2} \kappa_2 \sum_{j=1}^{p} \mathrm{d}(\mathbb{Q}_t \Sigma_{ZF,jt}) \left(\mathrm{d}\phi_{jt}^{(1,2)} \right)' + o(1)$. Similarly,

$$\begin{aligned} A^Y_{t,22112} &= \frac{h_2^{1/2}}{T^{1/2}}\sum_{s=1}^T\sum_{j=1}^p k_{h_2,st}\mathbb{Q}_s\Sigma_{ZF,js}\left[\frac{1}{2}h_2^2\left(\mathrm{d}^2\phi_{jt}^{(1,2)}\right)'\left(\frac{s-t}{Th_2}\right)^2\right]+o_P(1) \\ &= T^{1/2}h_2^{5/2}\kappa_2\sum_{j=1}^p\left[\frac{1}{2}\mathbb{Q}_t\Sigma_{ZF,jt}\left(\mathrm{d}^2\phi_{jt}^{(1,2)}\right)'\right]+o_P(1). \end{aligned}$$

It follows that

$$A^Y_{t,2211}=T^{1/2}h_2^{5/2}\kappa_2\sum_{j=1}^p\left\{\mathrm{d}\left(\mathbb{Q}_t\Sigma_{ZF,jt}\right)\left(\mathrm{d}\phi_{jt}^{(1,2)}\right)'+\frac{1}{2}\mathbb{Q}_t\Sigma_{ZF,jt}\left(\mathrm{d}^2\phi_{jt}^{(1,2)}\right)'\right\}+o_P(1). \tag{S2.2}$$

For $A^Y_{t,2212}$, we have

$$\begin{aligned} \left\|A^Y_{t,2212}\right\| &= \frac{h_2^{1/2}}{T^{1/2}}\left\|\sum_{s=1}^T\sum_{j=1}^p k_{h_2,st}\left(\hat{Z}_s-Z_s\right)F'_{s-j}\left[\phi_j^{(1,2)}\left(\frac{s}{T}\right)-\phi_j^{(1,2)}\left(\frac{t}{T}\right)\right]'\right\| \\ &\leq \max_{s\in[T]}\left\|\hat{Z}_s-Z_s\right\|T^{1/2}h_2^{1/2}\sum_{j=1}^p\left\|\frac{1}{T}\sum_{s=1}^T k_{h_2,st}F'_{s-j}\left[\phi_j^{(1,2)}\left(\frac{s}{T}\right)-\phi_j^{(1,2)}\left(\frac{t}{T}\right)\right]'\right\| \\ &= O_P\left((N/\ln T)^{-1/2}+h_1^2\right)O_P\left(T^{1/2}h_2^{1/2}h_2^2\right)=o_P(1), \end{aligned}$$

where we use the fact that $\max_{s\in[T]}\left\|\hat{Z}_s-Z_s\right\| = O_P\left((N/\ln T)^{-1/2}+h_1^2\right)$ by (S2.1). For $A^Y_{t,221}(3)$, we have

$$\begin{aligned} A^Y_{t,2213} &= \frac{h_2^{1/2}}{T^{1/2}}\sum_{s=1}^T\sum_{j=1}^p k_{h_2,st}Z_sF'_{s-j}\left(H_{t-j}-H_{s-j}\right)H_{t-j}^{-1}\phi_{jt}^{(1,2)\prime} \\ &= \frac{h_2^{1/2}}{T^{1/2}}\sum_{s=1}^T\sum_{j=1}^p k_{h_2,st}\mathbb{D}_sZ_s^0F'_{s-j}\left(Q_{t-j}^{-1}-Q_{s-j}^{-1}\right)Q_{t-j}\phi_{jt}^{(1,2)\prime} \\ &\quad+\frac{h_2^{1/2}}{T^{1/2}}\sum_{s=1}^T\sum_{j=1}^p k_{h_2,st}\mathbb{D}_sZ_s^0F'_{s-j}\left(Q_{s-j}^{-1}-H_{s-j}\right)Q_{t-j}^{-1}\phi_{jt}^{(1,2)\prime} \\ &\quad+\frac{h_2^{1/2}}{T^{1/2}}\sum_{s=1}^T\sum_{j=1}^p k_{h_2,st}\mathbb{D}_sZ_s^0F'_{s-j}H_{s-j}\left(Q_{t-j}-H_{t-j}^{-1}\right)\phi_{jt}^{(1,2)\prime}\equiv\sum_{\ell=1}^3 A^Y_{t,2213\ell}. \end{aligned}$$

Based on the proof of Lemma A.1(iii) in Su and Wang (2017), we can show that

$$\max_{t\in[T]} \left\| H_t - Q_t^{-1} \right\| = O_P\left((Th_1 \wedge N)^{-1/2} (\ln T)^{1/2} \right). \tag{S2.3}$$

With this result, we can readily show that $A^Y_{t,2213\ell} = O_P\left(T^{1/2} h_2^{1/2} (Th_1 \wedge N)^{-1/2} (\ln T)^{1/2} \right) = o_P(1)$ for $\ell = 2, 3$ under Assumption A.2(iii). For $A^Y_{t,22131}$, we can follow the analysis of $A^Y_{t,2211}$ and show that

$$\begin{aligned}
A^Y_{t,22131} &= \frac{h_2^{1/2}}{T^{1/2}} \sum_{s=1}^T \sum_{j=1}^p k_{h_2,st} \mathbb{D}_s Z_s^0 F'_{s-j} Q_{s-j}^{-1} \left(Q_{s-j} - Q_{t-j} \right) \phi_{jt}^{(1,2)\prime} \\
&= \frac{h_2^{1/2}}{T^{1/2}} \sum_{s=1}^T \sum_{j=1}^p k_{h_2,st} \mathbb{Q}_s \Sigma_{ZF,js} Q_{s-j}^{-1} \left[h_2 \left(\mathrm{d} Q_{t-j} \right) \left(\frac{s-t}{Th_2} \right) + \frac{h_2^2}{2} \left(\mathrm{d}^2 Q_{t-j} \right) \left(\frac{s-t}{Th_2} \right)^2 \right] \phi_{jt}^{(1,2)\prime} \\
&\quad + o_P(1) \\
&= T^{1/2} h_2^{5/2} \sum_{j=1}^p \left\{ \left[\mathrm{d} \left(\mathbb{Q}_t \Sigma_{ZF,jt} Q_{t-j}^{-1} \right) \right] \left(\mathrm{d} Q_{t-j} \right) + \frac{1}{2} \mathbb{Q}_t \Sigma_{ZF,jt} Q_{t-j}^{-1} \left(\mathrm{d}^2 Q_{t-j} \right) \right\} \phi_{jt}^{(1,2)\prime} \kappa_2 + o_P(1).
\end{aligned}$$

Then

$$A^Y_{t,2213} = T^{1/2} h_2^{5/2} \kappa_2 \sum_{j=1}^p \left\{ \left[\mathrm{d} \left(\mathbb{Q}_t \Sigma_{ZF,jt} Q_{t-j}^{-1} \right) \right] \left(\mathrm{d} Q_{t-j} \right) + \frac{1}{2} \mathbb{Q}_t \Sigma_{ZF,jt} Q_{t-j}^{-1} \left(\mathrm{d}^2 Q_{t-j} \right) \right\} \phi_{jt}^{(1,2)\prime} + o_P(1). \tag{S2.4}$$

For $A^Y_{t,221}(4)$, we have

$$\begin{aligned}
\left\| A^Y_{t,2214} \right\| &= \frac{h_2^{1/2}}{T^{1/2}} \left\| \sum_{s=1}^T \sum_{j=1}^p k_{h_2,st} \left(\hat{Z}_s - Z_s \right) F'_{s-j} H_{s-j} \left(H_{s-j}^{-1} - H_{t-j}^{-1} \right) \phi_j^{(1,2)} \left(\frac{t}{T} \right)' \right\| \\
&\lesssim T^{1/2} h_2^{1/2} \max_{1\le s\le T} \left\| \hat{Z}_s - Z_s \right\| \sum_{j=1}^p \frac{1}{Th_2} \sum_{s=1}^T K\left(\frac{s-t}{Th_2} \right) \left\| H'_{s-j} F_{s-j} \right\| \left\| \phi_j^{(1,2)} \left(\frac{t}{T} \right) \right\| \left\| H_{s-j}^{-1} - H_{t-j}^{-1} \right\| \\
&= T^{1/2} h_2^{1/2} O_P\left((N/\ln T)^{-1/2} + h_1^2 \right) O_P(1) = o_P(1).
\end{aligned}$$

In sum, we conclude that

$$\begin{aligned}
A^Y_{t,22} &= T^{1/2} h_2^{5/2} \kappa_2 \sum_{j=1}^p \left\{ \left[\mathrm{d} \left(\mathbb{Q}_t \Sigma_{ZF,jt} \right) \right] \left(\mathrm{d} \phi_{jt}^{(1,2)} \right)' + \frac{1}{2} \mathbb{Q}_t \Sigma_{ZF,jt} \left(\mathrm{d}^2 \phi_{jt}^{(1,2)} \right)' \right. \\
&\quad \left. + \left[\left[\mathrm{d} \left(\mathbb{Q}_t \Sigma_{ZF,jt} Q_{t-j}^{-1} \right) \right] \left(\mathrm{d} Q_{t-j} \right) + \frac{1}{2} \mathbb{Q}_t \Sigma_{ZF,jt} Q_{t-j}^{-1} \left(\mathrm{d}^2 Q_{t-j} \right) \right] \phi_{jt}^{(1,2)\prime} \right\} + o_P(1).
\end{aligned} \tag{S2.5}$$

It remains to analyze $A^Y_{t,23}$. As in the analysis of $A^Y_{t,22}$, we can show that

$$\begin{aligned}
A^Y_{t,23} &= \frac{h_2^{1/2}}{T^{1/2}} \sum_{s=1}^{T} \sum_{j=1}^{p} k_{h_2,st} \hat{Z}_s Y'_{s-j} \left(\psi_j^{(1,1)} \left(\frac{s}{T} \right) - \psi_j^{(1,1)} \left(\frac{t}{T} \right) \right)' \\
&= \frac{h_2^{1/2}}{T^{1/2}} \sum_{s=1}^{T} \sum_{j=1}^{p} k_{h_2,st} Z_s Y'_{s-j} \left(\phi_j^{(1,1)} \left(\frac{s}{T} \right) - \phi_j^{(1,1)} \left(\frac{t}{T} \right) \right)' + o_P(1) \\
&= \frac{h_2^{1/2}}{T^{1/2}} \sum_{s=1}^{T} \sum_{j=1}^{p} k_{h_2,st} \mathbb{Q}_s \Sigma_{ZY,js} \left[h_2 \mathrm{d}\phi_{jt}^{(1,1)} \frac{s-t}{Th_2} + \frac{1}{2} h_2^2 \mathrm{d}^2 \phi_{jt}^{(1,1)} \left(\frac{s-t}{Th_2} \right)^2 \right]' + o_P(1) \\
&= T^{1/2} h_2^{5/2} \kappa_2 \sum_{j=1}^{p} \left\{ \left[\mathrm{d} \left(\mathbb{Q}_t \Sigma_{ZY,jt} \right) \right] \left[\mathrm{d}\phi_{jt}^{(1,1)} \right] + \frac{1}{2} \mathbb{Q}_t \Sigma_{ZY,jt} \left[\mathrm{d}^2 \phi_{jt}^{(1,1)} \right] \right\} + o_P(1). \qquad \text{(S2.6)}
\end{aligned}$$

where $\Sigma_{ZY,js} \equiv \Sigma_{ZY,j}(s/T) \equiv E(Z_s^0 Y'_{s-j})$.

We hereby finish the proof. ■

Proof Lemma S1.2. Note that

$$\begin{aligned}
A^F_{t,2} &= \frac{h_2^{1/2}}{T^{1/2}} \sum_{s=1}^{T} k_{h_2,st} \hat{Z}_s \Delta'_F(s,t) \\
&= \frac{h_2^{1/2}}{T^{1/2}} \sum_{j=1}^{p} \sum_{s=1}^{T} k_{h_2,st} \hat{Z}_s Y'_{s-j} \left[\psi_j^{(2,1)} \left(\frac{s}{T} \right) - \psi_j^{(2,1)} \left(\frac{t}{T} \right) \right]' \\
&\quad + \frac{h_2^{1/2}}{T^{1/2}} \sum_{j=1}^{p} \sum_{s=1}^{T} k_{h_2,st} \hat{Z}_s \hat{F}'_{s-j} \left[\psi_j^{(2,2)} \left(\frac{s}{T} \right) - \psi_j^{(2,2)} \left(\frac{t}{T} \right) \right]' \\
&\quad + \frac{h_2^{1/2}}{T^{1/2}} \sum_{j=1}^{p} \sum_{s=1}^{T} k_{h_2,st} \hat{Z}_s \left(H_{s-j} F_{s-j} - \hat{F}_{s-j} \right)' \psi_j^{(2,2)} \left(\frac{s}{T} \right)' \\
&\quad + \frac{h_2^{1/2}}{T^{1/2}} \sum_{s=1}^{T} k_{h_2,st} \hat{Z}_s (\hat{F}_s - H'_s F_s)' \equiv \sum_{\ell=1}^{4} A^F_{t,2\ell}.
\end{aligned}$$

For $A^F_{t,21}$, we decompose

$$\begin{aligned}
A^F_{t,21} &= \frac{h_2^{1/2}}{T^{1/2}} \sum_{j=1}^{p} \sum_{s=1}^{T} k_{h_2,st} \hat{Z}_s Y'_{s-j} \left[H'_s \phi_j^{(2,1)} \left(\frac{s}{T} \right) - H'_t \phi_j^{(2,1)} \left(\frac{t}{T} \right) \right]' \\
&= \frac{h_2^{1/2}}{T^{1/2}} \sum_{j=1}^{p} \sum_{s=1}^{T} k_{h_2,st} Z_s Y'_{s-j} \left[\phi_j^{(2,1)} \left(\frac{s}{T} \right) - \phi_j^{(2,1)} \left(\frac{t}{T} \right) \right]' H_s
\end{aligned}$$

$$+\frac{h_2^{1/2}}{T^{1/2}}\sum_{j=1}^{p}\sum_{s=1}^{T}k_{h_2,st}\left(\hat{Z}_s-Z_s\right)Y_{s-j}'\left[\phi_j^{(2,1)}\left(\frac{s}{T}\right)-\phi_j^{(2,1)}\left(\frac{t}{T}\right)\right]'H_s$$

$$+\frac{h_2^{1/2}}{T^{1/2}}\sum_{j=1}^{p}\sum_{s=1}^{T}k_{h_2,st}\hat{Z}_sY_{s-j}'\phi_j^{(2,1)}\left(\frac{t}{T}\right)'(H_s-H_t)\equiv\sum_{\ell=1}^{3}A_{t,21\ell}^F.$$

For $A_{t,211}^F$ and $A_{t,212}^F$, we can follow the analyses of $A_{t,2211}^Y$ and $A_{t,2212}^Y$ to obtain

$$\begin{aligned}
A_{t,211}^F &= \frac{h_2^{1/2}}{T^{1/2}}\sum_{j=1}^{p}\sum_{s=1}^{T}k_{h_2,st}Z_sY_{s-j}'\left[\phi_j^{(2,1)}\left(\frac{s}{T}\right)-\phi_j^{(2,1)}\left(\frac{t}{T}\right)\right]'H_s \\
&= \frac{h_2^{1/2}}{T^{1/2}}\sum_{j=1}^{p}\sum_{s=1}^{T}k_{h_2,st}\mathbb{Q}_s\Sigma_{ZY,js}\left[h_2\left[\mathrm{d}\phi_{jt}^{(2,1)}\right]\frac{s-t}{Th_2}+\frac{1}{2}h_2^2\left[\mathrm{d}^2\phi_{jt}^{(2,1)}\right]\left(\frac{s-t}{Th_2}\right)^2\right]'Q_s^{-1}+o_P(1) \\
&= T^{1/2}h_2^{5/2}\kappa_2\sum_{j=1}^{p}\left\{\mathrm{d}\left(\mathbb{Q}_t\Sigma_{ZY,jt}\right)\left[\mathrm{d}\phi_{jt}^{(2,1)}\right]'Q_t^{-1}+\mathbb{Q}_t\Sigma_{ZY,jt}\left[\mathrm{d}\phi_{jt}^{(2,1)}\right]'\left(\mathrm{d}Q_t^{-1}\right)\right. \\
&\quad\left.+\frac{1}{2}\mathbb{Q}_t\Sigma_{ZY,jt}\left[\mathrm{d}^2\phi_{jt}^{(2,1)}\right]'Q_t^{-1}\right\}+o_P(1),
\end{aligned}$$

and

$$\begin{aligned}
\left\|A_{t,212}^F\right\| &= \frac{h_2^{1/2}}{T^{1/2}}\left\|\sum_{j=1}^{p}\sum_{s=1}^{T}k_{h_2st}\left(\hat{Z}_s-Z_s\right)Y_{s-j}'\left[\phi_j^{(2,1)}\left(\frac{s}{T}\right)-\phi_j^{(2,1)}\left(\frac{t}{T}\right)\right]'H_s\right\| \\
&\le T^{1/2}h_2^{1/2}\max_{s\in[T]}\left\|\hat{Z}_s-Z_s\right\|\sum_{j=1}^{p}\frac{1}{T}\sum_{s=1}^{T}k_{h_2st}\left\|Y_{s-j}'\left[\phi_j^{(2,1)}\left(\frac{s}{T}\right)-\phi_j^{(2,1)}\left(\frac{t}{T}\right)\right]'H_s\right\| \\
&= T^{1/2}h_2^{1/2}O_P((N/\ln T)^{-1/2}+h_1^2)O_P(h_2)=o_P(1),
\end{aligned}$$

and

$$\begin{aligned}
A_{t,213}^F &= \frac{1}{\sqrt{Th_2}}\sum_{j=1}^{p}\sum_{s=1}^{T}K\left(\frac{s-t}{Th_2}\right)\hat{Z}_sY_{s-j}'\phi_{jt}^{(2,1)\prime}(H_s-H_t) \\
&= \frac{1}{\sqrt{Th_2}}\sum_{j=1}^{p}\sum_{s=1}^{T}K\left(\frac{s-t}{Th_2}\right)\mathbb{Q}_s\Sigma_{ZY,js}\phi_{jt}^{(2,1)\prime}\left(Q_s^{-1}-Q_t^{-1}\right)+o_P(1) \\
&= T^{1/2}h_2^{5/2}\kappa_2\sum_{j=1}^{p}\left\{\left[\mathrm{d}(\mathbb{Q}_t\Sigma_{ZY,jt})\right]\phi_{jt}^{(2,1)\prime}\left(\mathrm{d}Q_t^{-1}\right)+\frac{1}{2}\mathbb{Q}_t\Sigma_{ZY,jt}\phi_{jt}^{(2,1)\prime}\left(\mathrm{d}^2Q_t^{-1}\right)\right\}+o_P(1)
\end{aligned}$$

In sum, we have

$$
\begin{aligned}
A^F_{t,21} &= T^{1/2}h_2^{5/2}\kappa_2\sum_{j=1}^{p}\Big\{\left[\mathrm{d}\left(\mathbb{Q}_t\Sigma_{ZY,jt}\right)\right]\left(\mathrm{d}\phi_{jt}^{(2,1)}\right)'Q_t^{-1}+\mathbb{Q}_t\Sigma_{ZY,jt}\left[\mathrm{d}\phi_{jt}^{(2,1)}\right]'\left(\mathrm{d}Q_t^{-1}\right)\\
&\quad+\frac{1}{2}\mathbb{Q}_t\Sigma_{ZY,jt}\left[\mathrm{d}^2\phi_{jt}^{(2,1)}\right]'Q_t^{-1}+\left[\mathrm{d}(\mathbb{Q}_t\Sigma_{ZY,jt})\right]\phi_{jt}^{(2,1)\prime}\left(\mathrm{d}Q_t^{-1}\right)+\frac{1}{2}\mathbb{Q}_t\Sigma_{ZY,jt}\phi_{jt}^{(2,1)\prime}\left(\mathrm{d}^2Q_t^{-1}\right)\Big\}\\
&\quad+o_P(1). \qquad\qquad (\text{S2.7})
\end{aligned}
$$

Next, we consider $A^F_{t,22}$. Note that

$$
\begin{aligned}
A^F_{t,22} &= \frac{h_2^{1/2}}{T^{1/2}}\sum_{j=1}^{p}\sum_{s=1}^{T}k_{h_2,st}\hat{Z}_s\hat{F}'_{s-j}\left[H_{s-j}^{-1}\phi_j^{(2,2)}\left(\frac{s}{T}\right)'H_s-H_{t-j}^{-1}\phi_j^{(2,2)}\left(\frac{t}{T}\right)'H_t\right]\\
&= \frac{h_2^{1/2}}{T^{1/2}}\sum_{j=1}^{p}\sum_{s=1}^{T}k_{h_2,st}\hat{Z}_s\hat{F}'_{s-j}H_{s-j}^{-1}\left[\phi_j^{(2,2)}\left(\frac{s}{T}\right)-\phi_j^{(2,2)}\left(\frac{t}{T}\right)\right]'H_s\\
&\quad+\frac{h_2^{1/2}}{T^{1/2}}\sum_{j=1}^{p}\sum_{s=1}^{T}k_{h_2,st}\hat{Z}_s\hat{F}'_{s-j}\left[H_{s-j}^{-1}-H_{t-j}^{-1}\right]\phi_j^{(2,2)}\left(\frac{t}{T}\right)'H_s\\
&\quad+\frac{h_2^{1/2}}{T^{1/2}}\sum_{j=1}^{p}\sum_{s=1}^{T}k_{h_2,st}\hat{Z}_s\hat{F}'_{s-j}H_{t-j}^{-1}\phi_j^{(2,2)}\left(\frac{t}{T}\right)'(H_s-H_t)\;\equiv\;\sum_{\ell=1}^{3}A^F_{t,22\ell}.
\end{aligned}
$$

For $A^F_{t,221}$, we can follow the analysis of $A^Y_{t,2213}$ to obtain

$$
\begin{aligned}
A^F_{t,221} &= \frac{h_2^{1/2}}{T^{1/2}}\sum_{j=1}^{p}\sum_{s=1}^{T}k_{h_2,st}\hat{Z}_sF'_{s-j}\left[\phi_j^{(2,2)}\left(\frac{s}{T}\right)-\phi_j^{(2,2)}\left(\frac{t}{T}\right)\right]'Q_s^{-1}+o_P(1)\\
&= \frac{h_2^{1/2}}{T^{1/2}}\sum_{j=1}^{p}\sum_{s=1}^{T}k_{h_2,st}\mathbb{Q}_s\Sigma_{ZF,js}\left[h_2\left(\mathrm{d}\phi_{jt}^{(2,2)}\right)\left(\frac{s-t}{T}\right)+\frac{h_2^2}{2}\left[\mathrm{d}^2\phi_{jt}^{(2,2)}\right]\left(\frac{s-t}{T}\right)^2\right]'Q_s^{-1}+o_P(1)\\
&= T^{1/2}h_2^{5/2}\kappa_2\sum_{j=1}^{p}\Big\{\left[\mathrm{d}\left(\mathbb{Q}_t\Sigma_{ZF,jt}\right)\right]\left[\mathrm{d}\phi_{jt}^{(2,2)}\right]'Q_t^{-1}+\mathbb{Q}_t\Sigma_{ZF,jt}\left[\mathrm{d}\phi_{jt}^{(2,2)}\right]'\left(\mathrm{d}Q_t^{-1}\right)\\
&\quad+\frac{1}{2}\mathbb{Q}_t\Sigma_{ZF,jt}\left[\mathrm{d}^2\phi_{jt}^{(2,2)}\right]'Q_t^{-1}\Big\}+o_P(1).
\end{aligned}
$$

Similarly,

$$
A^F_{t,222} = \frac{h_2^{1/2}}{T^{1/2}}\sum_{j=1}^{p}\sum_{s=1}^{T}k_{h_2,st}\hat{Z}_s\hat{F}'_{s-j}\left[H_{s-j}^{-1}-H_{t-j}^{-1}\right]\phi_{jt}^{(2,2)\prime}H_s
$$

$$
\begin{aligned}
&= \frac{h_2^{1/2}}{T^{1/2}} \sum_{j=1}^{p} \sum_{s=1}^{T} k_{h_2,st} Z_s F_{s-j}' H_{s-j} \left[H_{s-j}^{-1} - H_{t-j}^{-1} \right] \phi_{jt}^{(2,2)\prime} H_s + o_P(1) \\
&= \frac{h_2^{1/2}}{T^{1/2}} \sum_{j=1}^{p} \sum_{s=1}^{T} k_{h_2,st} \mathbb{Q}_s \Sigma_{ZF,js} Q_{s-j}^{-1} \left[Q_{s-j} - Q_{t-j} \right] \phi_{jt}^{(2,2)\prime} Q_s^{-1} + o_P(1) \\
&= T^{1/2} h_2^{5/2} \kappa_2 \sum_{j=1}^{p} \Big\{ \left[\mathrm{d}\left(\mathbb{Q}_t \Sigma_{ZF,jt} Q_{t-j}^{-1} \right) \right] (\mathrm{d}Q_t) \phi_{jt}^{(2,2)\prime} Q_t^{-1} + \mathbb{Q}_t \Sigma_{ZF,jt} Q_{t-j}^{-1} (\mathrm{d}Q_t) \phi_{jt}^{(2,2)\prime} \left(\mathrm{d}Q_t^{-1} \right) \\
&\quad + \frac{1}{2} \mathbb{Q}_t \Sigma_{ZF,jt} Q_{t-j}^{-1} \left(\mathrm{d}^2 Q_t \right) \phi_{jt}^{(2,2)\prime} Q_t^{-1} \Big\} + o_P(1),
\end{aligned}
$$

and

$$
\begin{aligned}
A_{t,223}^F &= \frac{h_2^{1/2}}{T^{1/2}} \sum_{j=1}^{p} \sum_{s=1}^{T} k_{h_2,st} Z_s F_{s-j}' H_{s-j} H_{t-j}^{-1} \phi_{jt}^{(2,2)\prime} \left(Q_s^{-1} - Q_t^{-1} \right) + o_P(1) \\
&= \frac{h_2^{1/2}}{T^{1/2}} \sum_{j=1}^{p} \sum_{s=1}^{T} k_{h_2,st} \mathbb{Q}_s \Sigma_{ZF,js} Q_{s-j}^{-1} Q_{t-j} \phi_{jt}^{(2,2)\prime} \left(Q_s^{-1} - Q_t^{-1} \right) + o_P(1) \\
&= T^{1/2} h_2^{5/2} \kappa_2 \sum_{j=1}^{p} \left\{ \left[\mathrm{d}\left(\mathbb{Q}_t \Sigma_{ZF,jt} Q_{t-j}^{-1} \right) \right] Q_{t-j} \phi_{jt}^{(2,2)\prime} \left(\mathrm{d}Q_t^{-1} \right) + \frac{1}{2} \mathbb{Q}_t \Sigma_{ZF,jt} \phi_{jt}^{(2,2)\prime} \left(\mathrm{d}^2 Q_t^{-1} \right) \right\} + o_P(1).
\end{aligned}
$$

Then

$$
\begin{aligned}
A_{t,22}^F &= T^{1/2} h_2^{5/2} \kappa_2 \sum_{j=1}^{p} \Big\{ \left[\mathrm{d}\left(\mathbb{Q}_t \Sigma_{ZF,jt} \right) \right] \left[\mathrm{d}\phi_{jt}^{(2,2)} \right]' Q_t^{-1} + \mathbb{Q}_t \Sigma_{ZF,jt} \left[\mathrm{d}\phi_{jt}^{(2,2)} \right]' \left(\mathrm{d}Q_t^{-1} \right) \\
&\quad + \frac{1}{2} \mathbb{Q}_t \Sigma_{ZF,jt} \left[\mathrm{d}^2 \phi_{jt}^{(2,2)} \right]' Q_t^{-1} + \left[\mathrm{d}\left(\mathbb{Q}_t \Sigma_{ZF,jt} Q_{t-j}^{-1} \right) \right] (\mathrm{d}Q_t) \phi_{jt}^{(2,2)\prime} Q_t^{-1} \\
&\quad + \mathbb{Q}_t \Sigma_{ZF,jt} Q_{t-j}^{-1} (\mathrm{d}Q_t) \phi_{jt}^{(2,2)\prime} \left(\mathrm{d}Q_t^{-1} \right) + \frac{1}{2} \mathbb{Q}_t \Sigma_{ZF,jt} Q_{t-j}^{-1} \left(\mathrm{d}^2 Q_t \right) \phi_{jt}^{(2,2)\prime} Q_t^{-1} \\
&\quad + \left[\mathrm{d}\left(\mathbb{Q}_t \Sigma_{ZF,jt} Q_{t-j}^{-1} \right) \right] Q_{t-j} \phi_{jt}^{(2,2)\prime} \left(\mathrm{d}Q_t^{-1} \right) + \frac{1}{2} \mathbb{Q}_t \Sigma_{ZF,jt} \phi_{jt}^{(2,2)\prime} \left(\mathrm{d}^2 Q_t^{-1} \right) \Big\} + o_P(1). \qquad \text{(S2.8)}
\end{aligned}
$$

For $A_{t,23}^F$ and $A_{t,24}^F$, we use (S2.1) to obtain:

$$
\begin{aligned}
\left\| A_{t,23}^F \right\| &= \frac{h_2^{1/2}}{T^{1/2}} \left\| \sum_{j=1}^{p} \sum_{s=1}^{T} k_{h_2,st} \hat{Z}_s \left(H_{s-j}' F_{s-j} - \hat{F}_{s-j} \right)' \psi_j^{(2,2)} \left(\frac{s}{T} \right)' \right\| \\
&\le T^{1/2} h_2^{1/2} \max_s \left\| \hat{F}_s - H_s' F_s \right\| \frac{1}{T} \sum_{j=1}^{p} \sum_{s=1}^{T} k_{h_2,st} \left\| \hat{Z}_s \right\| \left\| \psi_j^{(2,2)} \left(\frac{s}{T} \right)' \right\|
\end{aligned}
$$

$$= T^{1/2}h_2^{1/2}O_P\left((N/\ln T)^{-1/2}+h_1^2\right)O_P(1)=o_P(1),$$

and

$$\begin{aligned} A_{t,24}^F &\leq T^{1/2}h_2^{1/2}\max_s\left\|\hat{F}_s-H_s'F_{s-j}\right\|\frac{1}{T}\sum_{s=1}^T k_{h_2,st}\left\|\hat{Z}_s\right\| \\ &= T^{1/2}h_2^{1/2}O_P\left((N/\ln T)^{-1/2}+h_1^2\right)O_P(1)=o_P(1). \end{aligned}$$

This completes the proof. ■

Proof of Proposition S1.1 Let H denote the rotation matrix as defined in the conventional PCA of Bai and Ng (2002) such that $\tilde{F}_t$ is consistent for $H'F_t$ under $\mathbb{H}_0^{(1)}$. Recall that $W_t^0 \equiv (Y_t', F_t')'$ and $Z_t^0 \equiv \left(W_{t-1}^{0\prime}, ..., W_{t-p}^{0\prime}\right)'$. Let $W_t^\dagger \equiv (Y_t', F_t'H)'$ and $Z_t^\dagger \equiv \left(W_{t-1}^{\dagger\prime}, ..., W_{t-p}^{\dagger\prime}\right)'$. Obviously, $W_t^\dagger$ is equal to W_t with H_t replaced by H. Recall that $D_H = \text{diag}(\mathbb{I}_K, H)$ and $\mathbb{D} = \mathbb{I}_p \otimes D_H'$. Then, $Z_t^\dagger = \mathbb{D}Z_t^0$. Note that

$$Y_t = \sum_{j=1}^p\left[\phi_{jt}^{(1,1)}Y_{t-j}+\phi_{jt}^{(1,2)}H'^{-1}H'F_{t-j}\right]+\varepsilon_{Y,t}=\sum_{j=1}^p\left[\psi_{jt}^{\dagger(1,1)}Y_{t-j}+\psi_{jt}^{\dagger(1,2)}H'F_{t-j}\right]+\varepsilon_{Y,t},$$

and

$$\begin{aligned} H'F_t &= \sum_{j=1}^p\left[H'\phi_{jt}^{(2,1)}Y_{t-j}+H'\phi_{jt}^{(2,2)}H'^{-1}H'F_{t-j}\right]+H'\varepsilon_{F,t} \\ &= \sum_{j=1}^p\left[\psi_{jt}^{\dagger(2,1)}Y_{t-j}+\psi_{jt}^{\dagger(2,2)}H'F_{t-j}\right]+H'\varepsilon_{F,t}, \end{aligned}$$

where $\psi_{jt}^{\dagger(1,1)}=\phi_{jt}^{(1,1)}$, $\psi_{jt}^{\dagger(1,2)}=\phi_{jt}^{(1,2)}H'^{-1}$, $\psi_{jt}^{\dagger(2,1)}=H'\phi_{jt}^{(2,1)}$, and $\psi_{jt}^{\dagger(2,2)}=H'\phi_{jt}^{(2,2)}H'^{-1}$. Let

$$\Psi_t^\dagger=\left(\Psi_{1t}^\dagger,...,\Psi_{pt}^\dagger\right)' \text{ where } \Psi_{j,t}^\dagger=\begin{pmatrix}\psi_{jt}^{\dagger(1,1)} & \psi_{jt}^{\dagger(1,2)} \\ \psi_{jt}^{\dagger(2,1)} & \psi_{jt}^{\dagger(2,2)}\end{pmatrix} \text{ for } j\in[p].$$

Then

$$W_t^\dagger=\Psi_t^{\dagger\prime}Z_t^\dagger+D_H'\varepsilon_t=\Psi_t^{\dagger\prime}Z_t^\dagger+\varepsilon_t^\dagger, \tag{S2.9}$$

where $\varepsilon_t^{\dagger} = D_H' \varepsilon_t$. In addition, notice that

$$\Psi_{j,t}^{\dagger} = D_H' \begin{pmatrix} \phi_{jt}^{(1,1)} & \phi_{jt}^{1,2} \\ \phi_{jt}^{(2,1)} & \phi_{jt}^{(2,2)} \end{pmatrix} D_H'^{-1} = D_H' \phi_{jt} D_H'^{-1} = D_H' \phi_{j0} D_H'^{-1} + a_{2T} D_H' g_{j,2}\left(t/T\right) D_H'^{-1},$$

under $\mathbb{H}_A^{(3)}\left(a_{2T}\right)$, we then have

$$\begin{aligned}
\Psi_t^{\dagger\prime} &= \left(D_H' \phi_{10} D_H'^{-1}, ..., D_H' \phi_{p0} D_H'^{-1}\right) + a_{2T}\left(D_H' g_{1,2}\left(t/T\right) D_H'^{-1}, ..., D_H' g_{p,2}\left(t/T\right) D_H'^{-1}\right) \\
&= D_H' \left(\phi_{10}, ..., \phi_{p0}\right)\left(\mathbb{I}_p \otimes D_H'^{-1}\right) + a_{2T} D_H' g_2\left(t/T\right)'\left(\mathbb{I}_p \otimes D_H'^{-1}\right) \\
&= \Psi_0' + a_{2T}\tilde{g}_2\left(t/T\right)', \qquad \text{(S2.10)}
\end{aligned}$$

where $\Psi_0' = D_H' \Phi_0' \left(\mathbb{I}_p \otimes D_H'^{-1}\right)$, $\Phi_0 = \left(\phi_{10}, ..., \phi_{p0}\right)'$, $\tilde{g}_2\left(t/T\right)' = D_H' g_2\left(t/T\right)'\left(\mathbb{I}_p \otimes D_H'^{-1}\right)$ and $g_2\left(t/T\right)' = \left(g_{1,2}\left(t/T\right), ..., g_{p,2}\left(t/T\right)\right)$.

With the above preparation, we can use (S2.9) and (S2.10) to decompose $\tilde{W}_t$ as follows:

$$\begin{aligned}
\tilde{W}_t &= \Psi_t^{\dagger\prime}\tilde{Z}_t + U_t^{\dagger} = \Psi_0'\tilde{Z}_t + a_{2T}\tilde{g}_2(t/T)'\tilde{Z}_t + U_t^{\dagger} \\
&= \Psi_0' Z_t + a_{2T}\tilde{g}_2(t/T)' Z_t + \Psi_0'(\tilde{Z}_t - Z_t) + a_{2T}\tilde{g}_2(t/T)'(\tilde{Z}_t - Z_t) + U_t^{\dagger}, \qquad \text{(S2.11)}
\end{aligned}$$

where the error term $U_t^{\dagger}$ can be represented as follows:

$$U_t^{\dagger} \equiv (\tilde{W}_t - W_t^{\dagger}) - \Psi_t^{\dagger\prime}(\tilde{Z}_t - Z_t^{\dagger}) + \varepsilon_t^{\dagger}. \qquad \text{(S2.12)}$$

It follows that

$$\begin{aligned}
\breve{\Psi}_t - \Psi_0 &= \left(\frac{1}{T}\sum_{s=1}^{T} k_{h_2,st}^{\dagger}\tilde{Z}_s\tilde{Z}_s'\right)^{-1}\frac{1}{T}\sum_{s=1}^{T} k_{h_2,st}^{\dagger}\tilde{Z}_s\tilde{W}_s' - \Psi_0 \\
&= \left(\frac{1}{T}\sum_{s=1}^{T} k_{h_2,st}^{\dagger}\tilde{Z}_s\tilde{Z}_s'\right)^{-1}\frac{1}{T}\sum_{s=1}^{T} k_{h_2,st}\tilde{Z}_s\left[\Psi_0'\tilde{Z}_s + a_{2T}\tilde{g}_2(s/T)'\tilde{Z}_s + U_s^{\dagger}\right]' - \Psi_0 \\
&= \tilde{S}_{ZZ,t}^{-1}\frac{1}{T}\sum_{s=1}^{T} k_{h_2,st}^{\dagger}\tilde{Z}_s U_s^{\dagger\prime} + a_{2T}\tilde{S}_{ZZ,t}^{-1}\frac{1}{T}\sum_{s=1}^{T} k_{h_2,st}^{\dagger}\tilde{Z}_s\tilde{Z}_s'\tilde{g}_2(s/T) \\
&= \tilde{S}_{ZZ,t}^{-1}\frac{1}{T}\sum_{s=1}^{T} k_{h_2,st}^{\dagger}\tilde{Z}_s U_s^{\dagger\prime} + a_{2T}\tilde{S}_{ZZ,t}^{-1}\frac{1}{T}\sum_{s=1}^{T} k_{h_2,st}^{\dagger}\tilde{Z}_s\tilde{Z}_s'\tilde{g}_2(s/T) \\
&= \tilde{S}_{ZZ,t}^{-1}\frac{1}{T}\sum_{s=1}^{T} k_{h_2,st}^{\dagger}\tilde{Z}_s\varepsilon_s^{\dagger\prime} + a_{2T}\tilde{S}_{ZZ,t}^{-1}\frac{1}{T}\sum_{s=1}^{T} k_{h_2,st}^{\dagger}\tilde{Z}_s\tilde{Z}_s'\tilde{g}_2(s/T) + \tilde{S}_{ZZ,t}^{-1}\frac{1}{T}\sum_{s=1}^{T} k_{h_2,st}^{\dagger}\tilde{Z}_s(\tilde{W}_s - W_s^{\dagger})'
\end{aligned}$$

$$
\begin{aligned}
& -\tilde{S}_{ZZ,t}^{-1}\frac{1}{T}\sum_{s=1}^{T}k_{h_2,st}^{\dagger}\tilde{Z}_s(\tilde{Z}_s - Z_s^{\dagger})'\Psi_0 - a_{2T}\tilde{S}_{ZZ,t}^{-1}\frac{1}{T}\sum_{s=1}^{T}k_{h_2,st}^{\dagger}\tilde{Z}_s(\tilde{Z}_s - Z_s^{\dagger})'g_2(s/T) \\
\equiv\ & D_{1t} + D_{2t} + D_{3t} + D_{4t} + D_{5t},
\end{aligned}
$$

where $\tilde{S}_{ZZ,t} \equiv \frac{1}{T}\sum_{s=1}^{T} k_{h_2,st}^{\dagger}\tilde{Z}_s\tilde{Z}_s'$. Then

$$
\begin{aligned}
Th_2^{1/2}M_{31} &= h_2^{1/2}\sum_{t=1}^{T}\left\|\breve{\Psi}_t - \Psi_0\right\|^2 = h_2^{1/2}\sum_{t=1}^{T}\|D_{1t} + D_{2t} + D_{3t} + D_{4t} + D_{5t}\|^2 \\
&= h_2^{1/2}\sum_{t=1}^{T}[\|D_{1t}\|^2 + \|D_{2t}\|^2 + \|D_{3t}\|^2 + \|D_{4t}\|^2 + \|D_{5t}\|^2 + 2\mathrm{tr}(D_{1t}D_{2t}') + 2\mathrm{tr}(D_{1t}D_{3t}') \\
&+ 2\mathrm{tr}(D_{1t}D_{4t}') + 2\mathrm{tr}(D_{1t}D_{5t}') + 2\mathrm{tr}(D_{2t}D_{3t}') + 2\mathrm{tr}(D_{2t}D_{4t}') + 2\mathrm{tr}(D_{2t}D_{5t}') + 2\mathrm{tr}(D_{3t}D_{4t}') \\
&+ 2\mathrm{tr}(D_{3t}D_{5t}') + 2\mathrm{tr}(D_{4t}D_{5t}')] \\
&\equiv M_{31,1} + M_{31,2} + M_{31,3} + M_{31,4} + M_{31,5} + 2M_{31,6} + 2M_{31,7} + 2M_{31,8} + 2M_{31,9} \\
&\quad + 2M_{31,10} + 2M_{31,11} + 2M_{31,12} + 2M_{31,13} + 2M_{31,14} + 2M_{31,15}.
\end{aligned}
$$

We prove the proposition by showing that (i) $M_{31,1} - \mathbb{B}_{3T}^{(1)} \xrightarrow{d} N(0, \mathbb{V}_{3T})$; (ii) $M_{31,2} = \Pi_3^{(1)} + o_P(1)$; and (iii) $M_{31,j} = o_P(1)$ for $j = 3, ..., 15$.

We first prove (i).

$$
\begin{aligned}
M_{31,1} &= h_2^{1/2}\sum_{t=1}^{T}\left\|\tilde{S}_{ZZ,t}^{-1}\frac{1}{T}\sum_{s=1}^{T}k_{h_2,st}^{\dagger}\tilde{Z}_s\varepsilon_s^{\dagger\prime}\right\|^2 \\
&= \frac{h_2^{1/2}}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{r=1}^{T}k_{h_2,st}^{\dagger}k_{h_2,rt}^{\dagger}\mathrm{tr}\left(\varepsilon_s^{\dagger}\tilde{Z}_s'\tilde{S}_{ZZ,t}^{-1}\tilde{S}_{ZZ,t}^{-1}\tilde{Z}_r\varepsilon_r^{\dagger\prime}\right) \\
&= \frac{h_2^{1/2}}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}k_{h_2,st}^{\dagger 2}\mathrm{tr}\left(\varepsilon_s^{\dagger}Z_s^{\dagger\prime}\tilde{S}_{ZZ,t}^{-1}\tilde{S}_{ZZ,t}^{-1}Z_s^{\dagger}\varepsilon_s^{\dagger\prime}\right) \\
&\quad + \frac{2h_2^{1/2}}{T^2}\sum_{t=1}^{T}\sum_{1\le r<s\le T}k_{h_2,st}^{\dagger}k_{h_2,rt}^{\dagger}\mathrm{tr}\left(\varepsilon_s^{\dagger}Z_s^{\dagger\prime}\tilde{S}_{ZZ,t}^{-1}\tilde{S}_{ZZ,t}^{-1}Z_r^{\dagger}\varepsilon_r^{\dagger\prime}\right) \\
&\quad + \frac{h_2^{1/2}}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{r=1}^{T}k_{h_2,st}^{\dagger}k_{h_2,rt}^{\dagger}\mathrm{tr}\left(\varepsilon_s^{\dagger}\left(\tilde{Z}_s - Z_s^{\dagger}\right)'\tilde{S}_{ZZ,t}^{-1}\tilde{S}_{ZZ,t}^{-1}\left(\tilde{Z}_r - Z_r^{\dagger}\right)\varepsilon_r^{\dagger\prime}\right) \\
&\quad + \frac{2h_2^{1/2}}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{r=1}^{T}k_{h_2,st}^{\dagger}k_{h_2,rt}^{\dagger}\mathrm{tr}\left(\varepsilon_s^{\dagger}Z_s^{\dagger\prime}\tilde{S}_{ZZ,t}^{-1}\tilde{S}_{ZZ,t}^{-1}\left(\tilde{Z}_r - Z_r^{\dagger}\right)\varepsilon_r^{\dagger\prime}\right) \\
&\equiv M_{31,1}^{(1)} + M_{31,1}^{(2)} + M_{31,1}^{(3)} + M_{31,1}^{(4)}.
\end{aligned}
$$

In what follows, we study $M_{31,1}^{(\ell)}$, $\ell = 1, 2, 3, 4$, in turn.

First, we observe that

$$M_{31,1}^{(1)} = \frac{h_2^{1/2}}{T^2} \sum_{t=1}^{T} \sum_{s=1}^{T} k_{h_2,st}^{\dagger 2} Z_s^{\dagger\prime} \tilde{S}_{ZZ,t}^{-1} \tilde{S}_{ZZ,t}^{-1} Z_s^{\dagger} \varepsilon_s^{\dagger\prime} \varepsilon_s^{\dagger} = \mathbb{B}_{3T}^{(1)}.$$

Next, we consider $M_{31,1}^{(2)}$. Recall that $Z_t^{0\prime} = (Y_{t-1}', F_{t-1}', \cdots, Y_{t-p}', F_{t-p}')$, $\tilde{Z}_t = (\tilde{W}_{t-1}', \cdots, \tilde{W}_{t-p}')'$, $\mathbb{Q} \equiv \mathbb{I}_p \otimes D_{Q^{-1}}'$, $\mathbb{D} \equiv \mathbb{I}_p \otimes D_H'$. By that $Z_t^{\dagger} = \mathbb{D} Z_t^0$, we have

$$\begin{aligned} Z_t^{\dagger} Z_t^{\dagger\prime} &= \mathbb{Q} Z_t^0 Z_t^{0\prime} \mathbb{Q}' + [\mathbb{D} - \mathbb{Q}] Z_t^0 Z_t^{0\prime} [\mathbb{D} - \mathbb{Q}]' + [\mathbb{D} - \mathbb{Q}] Z_t^0 Z_t^{0\prime} \mathbb{Q}' + \mathbb{Q} Z_t^0 Z_t^{0\prime} [\mathbb{D} - \mathbb{Q}]' \\ &\equiv d_{1t} + d_{2t} + d_{3t} + d_{4t}. \end{aligned}$$

We further let $\tilde{Z}_t \tilde{Z}_t' - Z_t^{\dagger} Z_t^{\dagger\prime} = (\tilde{Z}_t - Z_t^{\dagger})(\tilde{Z}_t - Z_t^{\dagger})' + (\tilde{Z}_t - Z_t^{\dagger}) Z_t^{\dagger\prime} + Z_t^{\dagger} (\tilde{Z}_t - Z_t^{\dagger})' \equiv d_{5t} + d_{6t} + d_{7t}$, we have

$$\tilde{S}_{ZZ,t} = \frac{1}{T} \sum_{s=1}^{T} k_{h_2,st}^{\dagger} Z_s^{\dagger} Z_s^{\dagger\prime} + \frac{1}{T} \sum_{s=1}^{T} k_{h_2,st}^{\dagger} \left[\tilde{Z}_s \tilde{Z}_s' - Z_s^{\dagger} Z_s^{\dagger\prime} \right] = \sum_{\ell=1}^{7} \frac{1}{T} \sum_{s=1}^{T} k_{h_2,st}^{\dagger} d_{\ell s} \equiv \sum_{\ell=1}^{7} \tilde{S}_{ZZ,t,\ell}.$$

Using the Bernstein inequality for strong mixing processes, we can readily show that

$$\max_t \left\| \tilde{S}_{ZZ,t,1} - E(\tilde{S}_{ZZ,t,1}) \right\| \lesssim \max_t \left\| \frac{1}{T} \sum_{s=1}^{T} k_{h_2,st}^{\dagger} [Z_s^0 Z_s^{0\prime} - E(Z_s^0 Z_s^{0\prime})] \right\| = O_P((Th_2/\ln T)^{-1/2}).$$

In addition, we can readily show that $\max_t \left\| \tilde{S}_{ZZ,t,\ell}^{(2)} \right\| = O_P\left(C_{0NT}^{-1}\right)$ for $\ell = 2, 3, 4$ by using the fact that

$$H - Q^{-1} = O_P\left((N \wedge T)^{-1/2}\right) = O_P\left(C_{0NT}^{-1}\right) \tag{S2.13}$$

by Lemma A.2(vi) in Su and Wang (2020a). Next, by Lemma A.3 in Su and Wang (2020a), we can readily show that

$$\begin{aligned} \max_t \left\| \tilde{S}_{ZZ,t,5} \right\| &= \max_t \left\| \frac{1}{T} \sum_{s=1}^{T} k_{h_2,st}^{\dagger} \left(\tilde{Z}_s - Z_s^{\dagger} \right) \left(\tilde{Z}_s - Z_s^{\dagger} \right)' \right\| \\ &\lesssim \max_t \sum_{j=1}^{p} \left\| \frac{1}{T} \sum_{s=1}^{T} k_{h_2,st}^{\dagger} \left(\tilde{F}_{s-j} - H' F_{s-j} \right) \left(\tilde{F}_{s-j} - H' F_{s-j}' \right)' \right\| \\ &= O_P\left(T^{-1} \ln T + N^{-1}\right), \end{aligned}$$

$$\max_t \left\| \tilde{S}_{ZZ,t,\ell} \right\| = \max_t \left\| \frac{1}{T}\sum_{s=1}^{T} k^{\dagger}_{h_2,st}\left(\tilde{Z}_s - Z^{\dagger}_s\right) Z^{\dagger\prime}_s \right\| = O_P\left(T^{-1}\ln T + N^{-1}\right) + o_P\left(a_{1NT}\right) \text{ for } \ell = 6,7.$$

It follows that

$$\max_t \left\| \tilde{S}_{ZZ,t} - S_{ZZ,t,1} \right\| = O_P\left((Th_2/\ln T)^{-1/2}\right) \tag{S2.14}$$

where $S_{ZZ,t,1} = E(\tilde{S}_{ZZ,t,1})$.

Now we study $M^{(2)}_{31,1}$ by making the following decomposition:

$$\begin{aligned}
M^{(2)}_{31,1} &= \frac{2h_2^{1/2}}{T^2}\sum_{t=1}^{T}\sum_{1\le r<s\le T}^{T} k^{\dagger}_{h_2,st}k^{\dagger}_{h_2,rt}Z^{\dagger\prime}_s \tilde{S}^{-1}_{ZZ,t}\tilde{S}^{-1}_{ZZ,t}Z^{\dagger}_r\varepsilon^{\dagger\prime}_r\varepsilon^{\dagger}_s \\
&= \frac{2h_2^{1/2}}{T^2}\sum_{t=1}^{T}\sum_{1\le r<s\le T}^{T} k^{\dagger}_{h_2,st}k^{\dagger}_{h_2,rt}Z^{\dagger\prime}_s S^{-1}_{ZZ,t,1}S^{-1}_{ZZ,t,1}Z^{\dagger}_r\varepsilon^{\dagger\prime}_r\varepsilon^{\dagger}_s \\
&\quad + \frac{2h_2^{1/2}}{T^2}\sum_{t=1}^{T}\sum_{1\le r<s\le T}^{T} k^{\dagger}_{h_2,st}k^{\dagger}_{h_2,rt}Z^{\dagger\prime}_s \left(\tilde{S}^{-1}_{ZZ,t}\tilde{S}^{-1}_{ZZ,t} - S^{-1}_{ZZ,t,1}S^{-1}_{ZZ,t,1}\right) Z^{\dagger}_r\varepsilon^{\dagger\prime}_r\varepsilon^{\dagger}_s \\
&= \frac{2}{Th_2^{1/2}}\sum_{1\le r<s\le T}^{T} \bar{K}\left(\frac{s-r}{Th_2}\right) Z^{\dagger\prime}_s \frac{1}{T}\sum_{t=1}^{T} S^{-1}_{ZZ,t,1}S^{-1}_{ZZ,t,1}Z^{\dagger}_r\varepsilon^{\dagger\prime}_r\varepsilon^{\dagger}_s \\
&\quad + \frac{2}{Th_2^{1/2}}\sum_{1\le r<s\le T}^{T} Z^{\dagger\prime}_s \frac{h_2}{T}\sum_{t=1}^{T}\left[k^{\dagger}_{h_2,st}k^{\dagger}_{h_2,rt} - \bar{K}\left(\frac{s-r}{Th_2}\right)\right] S^{-1}_{ZZ,t,1}S^{-1}_{ZZ,t,1}Z^{\dagger}_r\varepsilon^{\dagger\prime}_r\varepsilon^{\dagger}_s \\
&\quad + \frac{2h_2^{1/2}}{T^2}\sum_{t=1}^{T}\sum_{1\le r<s\le T}^{T} k^{\dagger}_{h_2,st}k^{\dagger}_{h_2,rt}Z^{\dagger\prime}_s \left(\tilde{S}^{-1}_{ZZ,t}\tilde{S}^{-1}_{ZZ,t} - S^{-1}_{ZZ,t,1}S^{-1}_{ZZ,t,1}\right) Z^{\dagger}_r\varepsilon^{\dagger\prime}_r\varepsilon^{\dagger}_s \equiv \sum_{\ell=1}^{3} M^{(2,\ell)}_{31,1},
\end{aligned}$$

where $\bar{K}(v) = \int_{-1}^{1} K(u)K(u-v)du$. Noting that $\varepsilon^{\dagger\prime}_r\varepsilon^{\dagger}_s = \varepsilon'_r D_H D'_H \varepsilon_s = (\varepsilon'_s \otimes \varepsilon'_r)\text{vec}(D_H D'_H)$ and $Z^{\dagger\prime}_t = Z^{0\prime}_t(\mathbb{I}_p \otimes D_H)$ by the definition of $Z^{\dagger}_t$, we have

$$\begin{aligned}
M^{(2,1)}_{31,1} &= \frac{2}{Th_2^{1/2}}\sum_{1\le r<s\le T}^{T} \bar{K}\left(\frac{s-r}{Th_2}\right) Z^{\dagger\prime}_s \frac{1}{T}\sum_{t=1}^{T} S^{-1}_{ZZ,t,1}S^{-1}_{ZZ,t,1}Z^{\dagger}_r\varepsilon^{\dagger\prime}_r\varepsilon^{\dagger}_s \\
&= \frac{2}{Th_2^{1/2}}\sum_{1\le r<s\le T}^{T} \bar{K}\left(\frac{s-r}{Th_2}\right) Z^{0\prime}_s \tilde{\mathbb{S}}_T Z^0_r\left(\varepsilon'_s \otimes \varepsilon'_r\right)\text{vec}\left(D_H D'_H\right) \\
&= 2\left[\text{vec}\left(\tilde{\mathbb{S}}_T\right)\right]' \frac{1}{Th_2^{1/2}}\sum_{1\le r<s\le T}^{T} \bar{K}\left(\frac{s-r}{Th_2}\right)\left(Z^0_r \otimes Z^0_s\right)\left(\varepsilon'_s \otimes \varepsilon'_r\right)\text{vec}\left(D_H D'_H\right) \\
&= 2\left[\text{vec}\left(\tilde{\mathbb{S}}_T\right)\right]' \frac{1}{Th_2^{1/2}}\sum_{1\le r<s\le T}^{T} \bar{K}\left(\frac{s-r}{Th_2}\right)\left(Z^0_r\varepsilon'_s\right) \otimes \left(Z^0_s\varepsilon'_r\right)\text{vec}\left(D_H D'_H\right)
\end{aligned}$$

$$= \quad 2\left[\text{vec}\left(\tilde{\mathbb{S}}_T\right)\right]'\sum_{s=2}^{T}\mathcal{Z}_s^0\text{vec}\left(D_H D_H'\right),$$

where $\tilde{\mathbb{S}}_T \equiv (\mathbb{I}_p \otimes D_H)\frac{1}{T}\sum_{t=1}^{T} S_{ZZ,t,1}^{-1} S_{ZZ,t,1}^{-1}(\mathbb{I}_p \otimes D_H')$ is a symmetric $p(K+R)\times p(K+R)$ matrix, and $\mathcal{Z}_T^0 \equiv T^{-1}h_2^{-1/2}\sum_{r=1}^{s-1}\bar{K}\left(\frac{s-r}{Th_2}\right)(Z_r^0\varepsilon_s')\otimes(Z_s^0\varepsilon_r')$. By straightforward moment calculations via the repeated use of Lemma S2.1 as in what follows, we can readily show that $\sum_{s=2}^{T}\mathcal{Z}_s^0 = O_P(1)$ by verifying that $E\left\|\sum_{s=2}^{T}\mathcal{Z}_s^0\right\|^2 = O(1)$. This along with (S2.13) implies that

$$\begin{aligned}
M_{31,1}^{(2,1)} &= 2\left[\text{vec}\left(\mathbb{S}_T\right)\right]'\sum_{s=2}^{T}\mathcal{Z}_s^0\text{vec}(D_{Q^{-1}}D_{Q^{-1}}') + o_P(1) \\
&= \frac{2}{Th_2^{1/2}}\sum_{1\le r<s\le T}\bar{k}_{sr}Z_s^{0\prime}\mathbb{S}_T Z_r^0\varepsilon_r' D_{Q^{-1}}D_{Q^{-1}}'\varepsilon_s + o_P(1) = 2\sum_{s=2}^{T}\mathcal{Z}_s + o_P(1) \equiv \bar{M}_{31,1}^{(2,1)} + o_P(1),
\end{aligned}$$

where $\mathbb{S}_T \equiv \mathbb{Q}'\frac{1}{T}\sum_{t=1}^{T} S_{Tt,1}^{-1}S_{Tt,1}^{-1}\mathbb{Q}$, $\bar{k}_{sr} = \bar{K}\left(\frac{s-r}{Th_2}\right)$, and $\mathcal{Z}_s = T^{-1}h_2^{-1/2}\sum_{r=1}^{s-1}\bar{k}_{sr}Z_s^{0\prime}\mathbb{S}_T Z_r^0\varepsilon_r' D_{Q^{-1}}D_{Q^{-1}}'\varepsilon_s$. Let

$$\mathbb{V}_{3T} \equiv 4T^{-2}h_2^{-1}\sum_{s=2}^{T}\sum_{r=1}^{s-1}\bar{k}_{sr}^2 E\left[\left(Z_r^{0\prime}\mathbb{S}_T Z_s^0\varepsilon_r' D_{Q^{-1}}D_{Q^{-1}}'\varepsilon_s\right)^2\right]. \tag{S2.15}$$

Below we will apply the martingale central limit theorem (CLT) to show $(\mathbb{V}_{3T})^{-1/2}\bar{M}_{31,1}^{(2,1)} \xrightarrow{d} N(0,1)$. Let $\mathcal{F}_t = \sigma\left(W_{t+1}^0, \varepsilon_t, W_t^0, \varepsilon_{t-1}...\right)$. Under Assumption A.3(iv),

$$E(\mathcal{Z}_s|\mathcal{F}_{s-1}) = T^{-1}h_2^{-1/2}\sum_{r=1}^{s-1}\bar{k}_{sr}Z_s^{0\prime}\mathbb{S}_T Z_r^0\varepsilon_r' D_{Q^{-1}}D_{Q^{-1}}' E(\varepsilon_s'|\mathcal{F}_{s-1}) = 0.$$

By the martingale central limit theorem (e.g., Pollard, 1984, p.171), it suffices to prove $(\mathbb{V}_{3T})^{-1/2}\bar{M}_{31,1}^{(2,1)} \xrightarrow{d} N(0,1)$ by showing that

$$\mathbb{Z} = \sum_{s=2}^{T} E\left[\mathcal{Z}_s^4|\mathcal{F}_{s-1}\right] = o_P(1), \text{ and } \sum_{s=1}^{T}\mathcal{Z}_s^2 - \mathbb{V}_{3T} = o_P(1). \tag{S2.16}$$

First, we verify the first part of (S2.16). Observing that $\mathbb{Z} \ge 0$, it suffices to show $\mathbb{Z} = o_P(1)$ by showing that $E(\mathbb{Z}) = o(1)$ by Markov inequality. Let $\phi_{sr} = Z_s^{0\prime}\mathbb{S}_T Z_r^0\varepsilon_r' D_{Q^{-1}}D_{Q^{-1}}'\varepsilon_s$. We have

$$
\begin{aligned}
E(\mathbb{Z}) &= \sum_{s=2}^{T} E\left\{\left[\frac{2}{T h_2^{1/2}} \sum_{r=1}^{s-1} \bar{k}_{sr}\phi_{sr}\right]^4\right\} \\
&= \frac{16}{T^4 h_2^2} \sum_{s=2}^{T} E\left[\sum_{r=1}^{s-1} \bar{k}_{sr}^4 \phi_{sr}^4 + 2 \sum_{1\le r_1<r_2\le s-1} \bar{k}_{sr_1}^2 \bar{k}_{sr_2}^2 \phi_{sr_1}^2 \phi_{sr_2}^2\right. \\
&\quad \left. +4\sum_{t=1}^{s-1} \sum_{1\le r_1<r_2\le s-1} \bar{k}_{st}^2 \bar{k}_{sr_1}\bar{k}_{sr_2}\phi_{st}^2 \phi_{sr_1}\phi_{sr_2} + 4 \sum_{\substack{1\le r_1<r_2\le s-1, \\ 1\le t_1<t_2\le s-1}} \bar{k}_{r_1 s}\bar{k}_{r_2 s}\bar{k}_{t_1 s}\bar{k}_{t_2 s}\phi_{sr_1}\phi_{sr_2}\phi_{st_1}\phi_{st_2}\right] \\
&\equiv \mathbb{Z}_1 + \mathbb{Z}_2 + \mathbb{Z}_3 + \mathbb{Z}_4.
\end{aligned}
$$

Let $\|A\|_q = E\|A\|^q$ for any $q \ge 1$. Noting that $\max_{r<s} \|\phi_{sr}\|_4^4 \le C$ under Assumption A.1(v)-(vi), we can readily show that under Assumption A.2

$$
\begin{aligned}
\mathbb{Z}_1 &\le \max_{r<s} \|\phi_{sr}\|_4^4 \frac{16}{T^4 h_2^2} \sum_{s=2}^{T} \sum_{r=1}^{s-1} \bar{k}_{sr}^4 = O\left(T^{-2} h_2^{-1}\right) \\
\mathbb{Z}_2 &\le \max_{r<s} \|\phi_{sr}\|_4^4 \frac{32}{T^4 h_2^2} \sum_{s=2}^{T} \sum_{1\le r_1<r_2\le s-1} \bar{k}_{sr_1}^2 \bar{k}_{sr_2}^2 = O\left(T^{-1}\right), \\
\mathbb{Z}_3 &\le \max_{r<s} \|\phi_{sr}\|_4^4 \frac{64}{T^4 h_2^2} \sum_{s=2}^{T} \sum_{t=1}^{s-1} \sum_{1\le r_1<r_2\le s-1} \bar{k}_{st}^2 \bar{k}_{sr_1}\bar{k}_{sr_2} = O(h_2).
\end{aligned}
$$

For $\mathbb{Z}_4$, we can apply Assumption A.4(i)–(ii) along with the Davydov inequality to show that

$$
\mathbb{Z}_4 = \frac{64}{T^4 h_2^2} \sum_{s=2}^{T} \sum_{1\le r_1<r_2\le s-1, 1\le t_1<t_2\le s-1} \bar{k}_{r_1 s}\bar{k}_{r_2 s}\bar{k}_{t_1 s}\bar{k}_{t_2 s} E\left(\phi_{sr_1}\phi_{sr_2}\phi_{st_1}\phi_{st_2}\right) = O(h_2).
$$

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

To verify the second part of (S2.16), it suffices to show (I) $\sum_{s=2}^{T} E[\mathcal{Z}_s^2] = \mathbb{V}_{3T} + o(1)$, and (II) $\text{Var}(\sum_{s=2}^{T} \mathcal{Z}_s^2) = o(1)$ by Chebyshev inequality. We first prove (I). Observe that

$$
\begin{aligned}
\text{Var}(M_{31,1}^{(2,1)}) &= \sum_{s=2}^{T} E\left(\mathcal{Z}_s^2\right) = 4T^{-2} h_2^{-1} \sum_{s=2}^{T} \sum_{r=1}^{s-1} \bar{k}_{sr}^2 E\left\{\left[Z_s^{0\prime} \mathbb{S}_T Z_r^0 \varepsilon_r' D_{Q^{-1}} D_{Q^{-1}}' \varepsilon_s\right]^2\right\} \\
&\quad + 8T^{-2} h_2^{-1} \sum_{s=2}^{T} \sum_{1\le r_1<r_2\le s-1} \bar{k}_{sr_1}\bar{k}_{sr_2} E\left[Z_s^{0\prime} \mathbb{S}_T Z_{r_1}^0 \varepsilon_{r_1}' D_{Q^{-1}} D_{Q^{-1}}' \varepsilon_s \varepsilon_s' D_{Q^{-1}} D_{Q^{-1}}' \varepsilon_{r_2} Z_2 Z_{r_2}^{0\prime} \mathbb{S}_T Z_s^0\right]
\end{aligned}
$$

$$\equiv \mathbb{V}_{3T} + 8b_{3T}.$$

Now, we study b_{3T}. For the time indices in the summands of b_{3T}, 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 $b_{3T}^{(l)}$ to denote b_{3T} when the time indices are restricted to case (l) for $l = 1, 2, 3$. In case (1), we apply Lemma S2.1 and the fact that $E(\varepsilon_{r_2}|\varepsilon_{r_2-1}, Z_{r_2}^0, \varepsilon_{r_2-2}, Z_{r_2-1}^0 ...) = 0$ under Assumption A.3(iv) to obtain

$$|b_{3T}^{(1)}| \le CT^{-2}h_2^{-1} \sum_{r_1 < r_2 < s} \bar{k}_{sr_1}\bar{k}_{sr_2}\alpha(T_0)^{\delta/(1+\delta)} = O\left(Th_2\alpha(T_0)^{\delta/(1+\delta)}\right) = o(1)$$

In case (2), we apply Lemma S2.1 and the fact that $E(\varepsilon_{r_1}|\varepsilon_{r_1-1}, Z_{r_1}^0, \varepsilon_{r_1-2}, Z_{r_1-1}^0 ...) = 0$ to obtain

$$|b_{T}^{(2)}| \le CT^{-2}h_2^{-1} \sum_{r_1 < r_2 < s} \bar{k}_{sr_1}\bar{k}_{sr_2}\alpha(T_0)^{\delta/(1+\delta)} = O\left(Th_2\alpha(T_0)^{\delta/(1+\delta)}\right) = o(1)$$

In case (3), we have

$$\begin{aligned} \left|b_{3T}^{(3)}\right| &\le T^{-2}h_2^{-1} \sum_{r_1 < r_2 < s, \text{ case (3)}} \bar{k}_{sr_1}\bar{k}_{sr_2} \left|E\left[Z_s^{0\prime}\mathbb{S}Z_{r_1}^0\varepsilon_{r_1}' D_{Q^{-1}}D_{Q^{-1}}'\varepsilon_s\varepsilon_s' D_{Q^{-1}}D_{Q^{-1}}'\varepsilon_{r_2}Z_{r_2}^{0\prime}\mathbb{S}Z_s^0\right]\right| \\ &\lesssim T^{-2}h_2^{-1} \sum_{r_1 < r_2 < s, \text{ case (3)}} \bar{k}_{sr_1}\bar{k}_{sr_2} = O\left(T^{-1}T_0^2 h_2\right) = o(1), \end{aligned}$$

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

Now, we want to prove (II) by showing that $E(\sum_{s=2}^T \mathcal{Z}_s^2)^2 = (\mathbb{V}_{3T})^2 + o(1)$. Noting that

$$\begin{aligned} E\left(\sum_{s=2}^T \mathcal{Z}_s^2\right)^2 &= \frac{1}{T^4h_2^2}E\left(\sum_{s=2}^T\left[\sum_{r=1}^{s-1}\bar{k}_{sr}\phi_{sr}\right]^2\right)^2 \\ &= \frac{1}{T^4h_2^2}E\left(\sum_{s=2}^T\sum_{r=1}^{s-1}\bar{k}_{sr}^2\phi_{sr}^2\right)^2 + \frac{1}{T^4h_2^2}E\left(\sum_{s=2}^T\sum_{1\le r_1\ne r_2\le s-1}\bar{k}_{sr_1}\bar{k}_{sr_2}\phi_{sr_1}\phi_{sr_2}\right)^2 \\ &\quad + \frac{2}{T^4h_2^2}E\left[\left(\sum_{s=2}^T\sum_{r=1}^{s-1}\bar{k}_{sr}^2\phi_{sr}^2\right)\sum_{s=2}^T\sum_{1\le r_1\ne r_2\le s-1}\bar{k}_{sr_1}\bar{k}_{sr_2}\phi_{sr_1}\phi_{sr_2}\right] \\ &\equiv c_{1T} + c_{2T} + c_{3T}, \text{ say,} \end{aligned}$$

it suffices to show that (a) $c_{1T} = (\mathbb{V}_{3T})^2 + o(1)$ and (b) $c_{2T} = o(1)$, because then $c_{3T} \leq 2\{c_{1T}c_{2T}\}^{1/2} = o_P(1)$ by Cauchy-Schwarz (CS) inequality. Note that $c_{1T} = \frac{1}{T^4h_2^2}\sum_{1\leq r_1<s_1\leq T, 1\leq r_2<s_2\leq T} \bar{k}_{s_1r_1}^2\bar{k}_{s_2r_2}^2 E(\phi_{s_1r_1}^2\phi_{s_2r_2}^2)$ and $(\mathbb{V}_{3T})^2 = \frac{1}{T^4h_2^2}\sum_{1\leq r_1<s_1\leq T, 1\leq r_2<s_2\leq T} \bar{k}_{s_1r_1}^2\bar{k}_{s_2r_2}^2 E\left(\phi_{s_1r_1}^2\right)E\left(\phi_{s_2r_2}^2\right)$. 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 $l \neq t$, 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 $c_{1T}(l)$ and $\mathbb{V}_{3T}^2(l)$ denote c_{1T} and $(\mathbb{V}_{3T})^2$ when the time indices are restricted to lie in $\mathcal{S}_{3,l}$, respectively. Note that $c_{1T} = c_{1T}(1) + c_{1T}(2)$ and $(\mathbb{V}_{3T})^2 = \mathbb{V}_{3T}^2(1) + \mathbb{V}_{3T}^2(2)$. In case (2), we have by Assumptions A.5(i)-(iv),

$$
\begin{aligned}
c_{1T}(2) &\leq \max_{s<r}\left\|\phi_{sr}^2\right\|_2^2 \frac{1}{T^4h_2^2}\sum_{\substack{1\leq r_1<s_1\leq T, 1\leq r_2<s_2\leq T,\\ \text{case (2)}}} \bar{k}_{s_1r_1}^2\bar{k}_{s_2r_2}^2 = O\left(T_0T^{-1}\right) = o(1), \\
\mathbb{V}_{3T}^2(2) &\leq \max_{s<r}\left[E\left(\phi_{sr}^2\right)\right]^2 \frac{1}{T^4h_2^2}\sum_{\substack{1\leq r_1<s_1\leq T, 1\leq r_2<s_2\leq T\\ \text{case (2)}}} \bar{k}_{s_1r_1}^2\bar{k}_{s_2r_2}^2 = O\left(T_0T^{-1}\right) = o(1),
\end{aligned}
$$

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 $c_{1T}(1, v)$ and $\mathbb{V}_{2NT}^2(1, v)$ to denote $c_{1T}(1)$ and $\mathbb{V}_{3T}^2(1)$, 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 S2.1 and Assumptions A.4(i)-(ii),

$$
\begin{aligned}
c_{1NT}(1, a) &= \frac{1}{T^4h_2^2}\sum_{r_1<s_1<r_2<s_2,\ \mathcal{S}_{3,1}} \bar{k}_{s_1r_1}^2\bar{k}_{s_2r_2}^2 E\left(\phi_{s_1r_1}^2\phi_{s_2r_2}^2\right) \\
&\leq \frac{1}{T^4h_2^2}\sum_{r_1<s_1<r_2<s_2,\ \mathcal{S}_{3,1}} \bar{k}_{s_1r_1}^2\bar{k}_{s_2r_2}^2 \left\{E\left(\phi_{s_1r_1}^2\right)E\left(\phi_{s_2r_2}^2\right) + C\alpha\left(T_0\right)^{\delta/(1+\delta)}\right\} \\
&= \mathbb{V}_{3T}^2(1, a) + o(1),
\end{aligned}
$$

where $\sum_{r_1<s_1<r_2<s_2,\ \mathcal{S}_{3,1}}$ 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, $c_{1T}(1, b) = \mathbb{V}_{3T}^2(1, b) + o(1)$. Now, consider subcase (1c). By applying Lemma S2.1 three times, we have

$$
c_{1T}(1, c) = \frac{1}{T^4h_2^2}\sum_{r_1<r_2<s_1<s_2,\mathcal{S}_{3,1}} \bar{k}_{s_1r_1}^2\bar{k}_{s_2r_2}^2 E\left(\phi_{s_1r_1}^2\phi_{s_2r_2}^2\right)
$$

$$
\begin{aligned}
&\leq \frac{1}{T^4 h_2^2} \sum_{r_1<r_2<s_1<s_2, \mathcal{S}_{3,1}} \bar{k}_{s_1 r_1}^2 \bar{k}_{s_2 r_2}^2 \left\{ E\left(\phi_{s_1 r_1}^2\right) E\left(\phi_{s_2 r_2}^2\right) + C\alpha\left(T_0\right)^{\delta/(1+\delta)} \right\} \\
&= \frac{1}{T^4 h_2^2} \sum_{r_1<r_2<s_1<s_2, \mathcal{S}_{3,1}} \bar{k}_{s_1 r_1}^2 \bar{k}_{s_2 r_2}^2 E\left(\phi_{s_1 r_1}^2\right) E\left(\phi_{s_2 r_2}^2\right) + o(1) \\
&= \mathbb{V}_{3T}^2\left(1, c\right) + o(1).
\end{aligned}
$$

It follows that $c_{1T}(1,c) = \mathbb{V}_{3T}^2(1,c) + o(1)$. Analogously, we can show that $c_{1T}(1,v) = \mathbb{V}_{3T}^2(1,v) + o(1)$ for $v = d, e, f$. Consequently, we have $c_{1T}(1) = \mathbb{V}_{3T}^2(1) + o(1)$ and $c_{1T} = (\mathbb{V}_{3T})^2 + o(1)$. Using arguments as used in the analysis of c_{1T} and Lemma S2.1, we can also show that

$$
\begin{aligned}
c_{2T} &= \frac{1}{T^4 h_2^2} \sum_{s_1=2}^{T} \sum_{s_2=2}^{T} \sum_{1\leq r_1 \neq r_2 \leq s_1-1} \sum_{1\leq r_3 \neq r_4 \leq s_2-1} \bar{k}_{s_1 r_1} \bar{k}_{s_1 r_2} \bar{k}_{s_2 r_3} \bar{k}_{s_2 r_4} E\left(\phi_{s_1 r_1} \phi_{s_1 r_2} \phi_{s_2 r_3} \phi_{s_2 r_4}\right) \\
&= O\left(T^{-1} h_2^{-2} + T h_2^2 \alpha\left(T_0\right)^{\delta/(1+\delta)} + T^{-2} T_0^4 + T^{-2} T_0^3 h_2^{-1} + T^{-2} T_0^2 h_2^{-2}\right) = o(1).
\end{aligned}
$$

It follows that $E(\sum_{s=2}^{T} \mathcal{Z}_s^2)^2 = (\mathbb{V}_{3T})^2 + o(1)$ and $\mathrm{Var}(\sum_{s=2}^{T} \mathcal{Z}_s^2) = o(1)$. Then the second part of (S2.16) follows by Chebyshev inequality, and we conclude that $(\mathbb{V}_{3T})^{-1/2} M_{31,1}^{(2,1)} \xrightarrow{d} N(0,1)$. In addition, by straightforward moment calculations, we can show that $M_{31,1}^{(2,\ell)} = o_P(1)$ for $\ell = 2, 3$. It follows that $(\mathbb{V}_{3T})^{-1/2} M_{31,1}^{(2)} \xrightarrow{d} N(0,1)$.

For $M_{31,1}^{(3)}$, we have

$$
\begin{aligned}
M_{31,1}^{(3)} &= \frac{h_2^{1/2}}{T^2} \sum_{t=1}^{T} \sum_{s=1}^{T} \sum_{r=1}^{T} k_{h_2,st}^{\dagger} k_{h_2,rt}^{\dagger} \mathrm{tr}\left(\varepsilon_s^{\dagger}\left(\tilde{Z}_s - Z_s^{\dagger}\right)' \tilde{S}_{ZZ,t}^{-1} \tilde{S}_{ZZ,t}^{-1}\left(\tilde{Z}_r - Z_r^{\dagger}\right) \varepsilon_r^{\dagger\prime}\right) \\
&= h_2^{1/2} \sum_{t=1}^{T} \left\| \tilde{S}_{ZZ,t}^{-1} \frac{1}{T} \sum_{s=1}^{T} k_{h_2,st}^{\dagger}\left(\tilde{Z}_s - Z_s^{\dagger}\right) \varepsilon_s^{\dagger\prime} \right\|^2 \lesssim h_2^{1/2} \sum_{t=1}^{T} \left\| \frac{1}{T} \sum_{s=1}^{T} k_{h_2,st}^{\dagger}\left(\tilde{Z}_s - Z_s^{\dagger}\right) \varepsilon_s^{\dagger\prime} \right\|^2,
\end{aligned}
$$

where the inequality follows from the fact that $\tilde{S}_{ZZ,t}^{-1}$ is asymptotically nonsingular uniform in t.

$$
M_{31,1}^{(3)} \lesssim \sum_{j=1}^{p} h_2^{1/2} \sum_{t=1}^{T} \left\| \frac{1}{T} \sum_{s=1}^{T} k_{h_2,st}^{\dagger}\left(\tilde{F}_{s-j} - H' F_{s-j}\right) \varepsilon_s' \right\|^2 = T h_2^{1/2} O_P(C_{0NT}^{-4} + a_{1NT}^2) = o_P(1),
$$

where the first equality follows from the fact that $\sum_{t=1}^{T} \left\| \frac{1}{T} \sum_{s=1}^{T} k_{h_2,st}^{\dagger}\left(\tilde{F}_{s-j} - H' F_{s-j}\right) \varepsilon_s' \right\|^2 = T[O_P\left(C_{0NT}^{-4}\right) + o_P\left(a_{1NT}^2\right)]$ by arguments as used in proving Lemma A.2 in Su and Wang (2020a).

Similarly, we can show that

$$M_{31,1}^{(4)} = \frac{2h_2^{1/2}}{T^2} \sum_{t=1}^{T} \sum_{s=1}^{T} \sum_{r=1}^{T} k_{h_2,st}^{\dagger} k_{h_2,rt}^{\dagger} \mathrm{tr}\left(\varepsilon_s^{\dagger} Z_s^{\dagger\prime} \tilde{S}_{ZZ,t}^{-1} \tilde{S}_{ZZ,t}^{-1}\left(\tilde{Z}_r - Z_r^{\dagger}\right)\varepsilon_r^{\dagger\prime}\right) = o_P(1).$$

Consequently, we have shown that $(\mathbb{V}_{3T})^{-1/2}\left(M_{31,1} - \mathbb{B}_{3T}^{(1)}\right) \xrightarrow{d} N(0,1)$.

Next, we prove (ii) $M_{31,2} = \Pi_3 + o_P(1)$. Using $a_{2T} = T^{-1/2} h_2^{-1/4}$ and $\tilde{Z}_s = Z_s + (\tilde{Z}_s - Z_s)$, we have

$$\begin{aligned}
M_{31,2} &= h_2^{1/2} a_{2T}^2 \sum_{t=1}^{T} \left\| \tilde{S}_{ZZ,t}^{-1} \frac{1}{T} \sum_{s=1}^{T} k_{h_2,st}^{\dagger} \tilde{Z}_s \tilde{Z}_s' \tilde{g}_2(s/T) \right\|^2 \\
&= \frac{1}{T} \sum_{t=1}^{T} \left\| \tilde{S}_{ZZ,t}^{-1} \frac{1}{T} \sum_{s=1}^{T} k_{h_2,st}^{\dagger} \tilde{Z}_s \tilde{Z}_s' \tilde{g}_2(t/T) \right\|^2 + \frac{1}{T} \sum_{t=1}^{T} \left\| \tilde{S}_{ZZ,t}^{-1} \frac{1}{T} \sum_{s=1}^{T} k_{h_2,st}^{\dagger} \tilde{Z}_s \tilde{Z}_s' \left[\tilde{g}_2(s/T) - \tilde{g}_2(t/T)\right] \right\|^2 \\
&+ \frac{2}{T} \sum_{t=1}^{T} \mathrm{tr}\left(\tilde{S}_{ZZ,t}^{-1} \frac{1}{T} \sum_{s=1}^{T} k_{h_2,st}^{\dagger} \tilde{Z}_s \tilde{Z}_s' \tilde{g}_2(t/T) \frac{1}{T} \sum_{r=1}^{T} k_{h_2,rt}^{\dagger} \left[\tilde{g}_2(s/T) - \tilde{g}_2(t/T)\right]' \tilde{Z}_r \tilde{Z}_r' \tilde{S}_{ZZ,t}^{-1} \right) \\
&\equiv M_{31,2}^{(1)} + M_{31,2}^{(2)} + 2M_{31,2}^{(3)}.
\end{aligned}$$

Noting that $\tilde{S}_{ZZ,t} = \frac{1}{T}\sum_{s=1}^{T} k_{h_2,st}^{\dagger} \tilde{Z}_s \tilde{Z}_s'$ and $\tilde{g}_2(s/T)' = D_H' g_2(s/T)'(\mathbb{I}_p \otimes D_H'^{-1})$, we have

$$\begin{aligned}
M_{31,2}^{(1)} &= \frac{1}{T} \sum_{s=1}^{T} \|\tilde{g}_2(s/T)\|^2 = \frac{1}{T} \sum_{s=1}^{T} \mathrm{tr}\left\{ D_H' g_2(s/T)'(\mathbb{I}_p \otimes D_H'^{-1} D_H^{-1}) g_2(s/T) D_H \right\} \\
&= \frac{1}{T} \sum_{s=1}^{T} \mathrm{tr}\left\{ D_H D_H' g_2(s/T)'(\mathbb{I}_p \otimes D_H'^{-1} D_H^{-1}) g_2(s/T) \right\} \\
&= \left[\mathrm{vec}\,(D_H D_H')\right]' \left[\frac{1}{T} \sum_{s=1}^{T} g_2(s/T)' \otimes g_2(s/T)' \right] \mathrm{vec}\left(\mathbb{I}_p \otimes D_H'^{-1} D_H^{-1}\right) \\
&= \left[\mathrm{vec}\left(D_{Q^{-1}} D_{Q^{-1}}'\right)\right]' \left[\int_0^1 g_2(u)' \otimes g_2(u)' du \right] \mathrm{vec}\left(\mathbb{I}_p \otimes D_{Q^{-1}}'^{-1} D_{Q^{-1}}^{-1}\right) + o_P(1) \\
&\equiv \Pi_3 + o_P(1),
\end{aligned}$$

where the third equality follows from the fact that $\mathrm{tr}(ABCD) = [\mathrm{vec}(A)]'(B \otimes D')\mathrm{vec}(C')$ for any four conformable matrices A, B, C and D. For $M_{31,2}^{(2)}$, we have

$$M_{31,2}^{(2)} = \frac{1}{T} \sum_{t=1}^{T} \left\| \tilde{S}_{ZZ,t}^{-1} \frac{1}{T} \sum_{s=1}^{T} k_{h_2,st}^{\dagger} \tilde{Z}_s \tilde{Z}_s' \left[\tilde{g}_2(s/T) - \tilde{g}_2(t/T)\right] \right\|^2$$

$$
\begin{aligned}
&\lesssim \frac{1}{T}\sum_{t=1}^{T}\left\|\frac{1}{T}\sum_{s=1}^{T}k_{h_2,st}^{\dagger}Z_s^{\dagger}Z_s^{\dagger\prime}\left[\tilde{g}_2(s/T)-\tilde{g}_2(t/T)\right]\right\|^2 \\
&\quad+\frac{1}{T}\sum_{t=1}^{T}\left\|\frac{1}{T}\sum_{s=1}^{T}k_{h_2,st}^{\dagger}\left(\tilde{Z}_s\tilde{Z}_s^{\prime}-Z_s^{\dagger}Z_s^{\dagger\prime}\right)\left[\tilde{g}_2(s/T)-\tilde{g}_2(t/T)\right]\right\|^2 \\
&= O_P\left((Th_2)^{-1}+h_2^2\right)=o_P(1).
\end{aligned}
$$

By CS inequality, $M_{31,2}^{(3)} \leq \{M_{31,2}^{(1)}M_{31,2}^{(2)}\}^{1/2} = o_P(1)$. In sum, $M_{31,2} = \Pi_3 + o_P(1)$.

Now, we prove (iii). Here we focus on the proof that $M_{31,\ell} = o_P(1)$ for $\ell = 3, 4, 5$, which, along with the fact that $M_{31,2} = O_P(1)$ and the CS inequality, implies that $M_{31,\ell} = o_P(1)$ for $\ell = 10, 11, ..., 15$. One can also show that $M_{31,\ell} = o_P(1)$ for $\ell = 6, 7, 8, 9$ by following the arguments as used in the analyses of $M_{31,\ell}$ for $\ell = 1, 2, ..., 5$. For $M_{31,3}$, we can follow the proof of Lemma A.2 in Su and Wang (2020a) and show that

$$
\begin{aligned}
M_{31,3} &= h_2^{1/2}\sum_{t=1}^{T}\left\|\tilde{S}_{ZZ,t}^{-1}\frac{1}{T}\sum_{s=1}^{T}k_{h_2,st}^{\dagger}\tilde{Z}_s(\tilde{W}_s-W_s^{\dagger})^{\prime}\right\|^2 \\
&\lesssim h_2^{1/2}\sum_{t=1}^{T}\left\|\frac{1}{T}\sum_{s=1}^{T}k_{h_2,st}^{\dagger}Z_s^{\dagger}(\tilde{W}_s-W_s^{\dagger})^{\prime}\right\|^2+h_2^{1/2}\sum_{t=1}^{T}\left\|\frac{1}{T}\sum_{s=1}^{T}k_{h_2,st}^{\dagger}(\tilde{Z}_s-Z_s^{\dagger})(\tilde{W}_s-W_s^{\dagger})^{\prime}\right\|^2 \\
&= Th_2^{1/2}\left[O_P\left(C_{0NT}^{-2}\right)+o_P(a_{1NT})\right]^2=o_P(1).
\end{aligned}
$$

By the same token,

$$
\begin{aligned}
M_{31,4} &\lesssim h_2^{1/2}\sum_{t=1}^{T}\left\|\frac{1}{T}\sum_{s=1}^{T}k_{h_2,st}^{\dagger}\tilde{Z}_s(\tilde{Z}_s-Z_s^{\dagger})^{\prime}\right\|^2 \\
&\lesssim h_2^{1/2}\sum_{t=1}^{T}\left\|\frac{1}{T}\sum_{s=1}^{T}k_{h_2,st}^{\dagger}Z_s^{\dagger}(\tilde{Z}_s-Z_s^{\dagger})^{\prime}\right\|^2+h_2^{1/2}\sum_{t=1}^{T}\left\|\frac{1}{T}\sum_{s=1}^{T}k_{h_2,st}^{\dagger}(\tilde{Z}_s-Z_s^{\dagger})(\tilde{Z}_s-Z_s^{\dagger})^{\prime}\right\|^2 \\
&= Th_2^{1/2}\left[O_P\left(C_{0NT}^{-2}\right)+o_P(a_{1NT})\right]^2=o_P(1), \text{ and} \\
M_{31,5} &\lesssim a_{2T}^2h_2^{1/2}\sum_{t=1}^{T}\left\|\frac{1}{T}\sum_{s=1}^{T}k_{h_2,st}^{\dagger}\tilde{Z}_s(\tilde{Z}_s-Z_s^{\dagger})^{\prime}g_2(s/T)\right\|^2=a_{2T}^2Th_2^{1/2}\left[O_P\left(C_{0NT}^{-2}\right)+o_P(a_{1NT})\right]^2 \\
&= o_P(1).
\end{aligned}
$$

This completes the proof of Proposition S1.1. ■

Proof of Proposition S1.2. Let $\tilde{S}_{ZZ} \equiv \frac{1}{T}\sum_{s=1}^{T}\tilde{Z}_s\tilde{Z}_s^{\prime}$. Noting that $\tilde{W}_t = \Psi_t^{\dagger\prime}\tilde{Z}_t + U_t^{\dagger} =$

$\Psi_0'\tilde{Z}_t + a_{2T}\tilde{g}_2(t/T)'\tilde{Z}_t + U_t^\dagger$ by (S2.10) and $U_t^\dagger = (\tilde{W}_t - W_t^\dagger) - \Psi_t^{\dagger\prime}(\tilde{Z}_t - Z_t^\dagger) + \varepsilon_t^\dagger$ by (S2.12), we have

$$
\begin{aligned}
\tilde{\Psi}_0 - \Psi_0 &= \left(\frac{1}{T}\sum_{s=1}^{T}\tilde{Z}_s\tilde{Z}_s'\right)^{-1}\frac{1}{T}\sum_{s=1}^{T}\tilde{Z}_s\tilde{W}_s' - \Psi_0 \\
&= \tilde{S}_{ZZ}^{-1}\frac{1}{T}\sum_{s=1}^{T}\tilde{Z}_s\left[\Psi_0'\tilde{Z}_s + a_{2T}\tilde{g}_2(s/T)'\tilde{Z}_s + U_s^\dagger\right]' - \Psi_0 \\
&= \tilde{S}_{ZZ}^{-1}\frac{1}{T}\sum_{s=1}^{T}\tilde{Z}_s U_s^{\dagger\prime} + a_{2T}\tilde{S}_{ZZ}^{-1}\frac{1}{T}\sum_{s=1}^{T}\tilde{Z}_s\tilde{Z}_s'\tilde{g}_2(s/T) \\
&= \tilde{S}_{ZZ}^{-1}\frac{1}{T}\sum_{s=1}^{T}\tilde{Z}_s\left[\left(\tilde{W}_s - W_s^\dagger\right) - \left(\Psi_0 + a_{2T}\tilde{g}_2\left(s/T\right)\right)'\left(\tilde{Z}_s - Z_s^\dagger\right) + \varepsilon_s^\dagger\right]' \\
&+ a_{2T}\tilde{S}_{ZZ}^{-1}\frac{1}{T}\sum_{s=1}^{T}\tilde{Z}_s\tilde{Z}_s' g_2(s/T) \\
&= \tilde{S}_{ZZ}^{-1}\frac{1}{T}\sum_{s=1}^{T}\tilde{Z}_s\varepsilon_s^{\dagger\prime} + a_{2T}\tilde{S}_{ZZ}^{-1}\frac{1}{T}\sum_{s=1}^{T}\tilde{Z}_s\tilde{Z}_s'\tilde{g}_2(s/T) + \tilde{S}_{ZZ}^{-1}\frac{1}{T}\sum_{s=1}^{T}\tilde{Z}_s(\tilde{W}_s - W_s^\dagger)' \\
&\quad - \tilde{S}_{ZZ}^{-1}\frac{1}{T}\sum_{s=1}^{T}\tilde{Z}_s(\tilde{Z}_s - Z_s^\dagger)'\Psi_0 - a_{2T}\tilde{S}_{ZZ}^{-1}\frac{1}{T}\sum_{s=1}^{T}\tilde{Z}_s(\tilde{Z}_s - Z_s^\dagger)' g_2(s/T) \\
&\equiv D_1 + D_2 + D_3 + D_4 + D_5. \qquad \text{(S2.17)}
\end{aligned}
$$

Then

$$
\begin{aligned}
M_{32} &= Th^{1/2}\left\|\tilde{\Psi}_0 - \Psi_0\right\|^2 \\
&= Th^{1/2}[\|D_1\|^2 + \|D_2\|^2 + \|D_3\|^2 + \|D_4\|^2 + \|D_5\|^2 + 2\mathrm{tr}(D_1D_2') + 2\mathrm{tr}(D_1D_3') + 2\mathrm{tr}(D_1D_4') \\
&+ 2\mathrm{tr}(D_1D_5') + 2\mathrm{tr}(D_2D_3') + 2\mathrm{tr}(D_2D_4') + 2\mathrm{tr}(D_2D_5') + 2\mathrm{tr}(D_3D_4') + 2\mathrm{tr}(D_3D_5') + 2\mathrm{tr}(D_4D_5')] \\
&\equiv M_{32,1} + M_{32,2} + M_{32,3} + M_{32,4} + M_{32,5} + 2M_{32,6} + 2M_{32,7} + 2M_{32,8} + 2M_{32,9} \\
&\quad + 2M_{32,10} + 2M_{32,11} + 2M_{32,12} + 2M_{32,13} + 2M_{32,14} + 2M_{32,15}.
\end{aligned}
$$

We prove the proposition by showing that (i) $M_{32,1} = \mathbb{B}_{3T}^{(2)} + o_P(1)$, (ii) $M_{32,2} = o_P(1)$, (iii) $M_{32,j} = o_P(1)$ for $j = 3, \cdots, 15$.

We first prove (i). We decompose $M_{32,1}$ as follows:

$$
M_{32,1} = Th_2^{1/2}\left\|D_1\right\|^2 = Th_2^{1/2}\left\|\tilde{S}_{ZZ}^{-1}\frac{1}{T}\sum_{s=1}^{T}\tilde{Z}_s\varepsilon_s^{\dagger\prime}\right\|^2
$$

$$
\begin{aligned}
&= T^{-1}h_2^{1/2}\sum_{s=1}^{T}\sum_{r=1}^{T}\tilde{Z}_s'\tilde{S}_{ZZ}^{-1}\tilde{S}_{ZZ}^{-1}\tilde{Z}_r\varepsilon_r^{\dagger\prime}\varepsilon_s^{\dagger} \\
&= \frac{h_2^{1/2}}{T}\sum_{s=1}^{T}\sum_{r=1}^{T}\varepsilon_s^{\dagger}Z_s^{\dagger\prime}\tilde{S}_{ZZ}^{-1}\tilde{S}_{ZZ}^{-1}Z_r^{\dagger}U_r' + \frac{h_2^{1/2}}{T}\sum_{s=1}^{T}\sum_{r=1}^{T}(\tilde{Z}_s - Z_s^{\dagger})'\tilde{S}_{ZZ}^{-1}\tilde{S}_{ZZ}^{-1}\left(\tilde{Z}_r - Z_r^{\dagger}\right)\varepsilon_r^{\dagger\prime}\varepsilon_s^{\dagger} \\
&\quad + \frac{2h_2^{1/2}}{T}\sum_{s=1}^{T}\sum_{r=1}^{T}Z_s^{\dagger\prime}\tilde{S}_{ZZ}^{-1}\tilde{S}_{ZZ}^{-1}\left(\tilde{Z}_r - Z_r^{\dagger}\right)\varepsilon_r^{\dagger\prime}\varepsilon_s^{\dagger} \\
&= \frac{h_2^{1/2}}{T}\sum_{s=1}^{T}Z_s^{\dagger\prime}\tilde{S}_{ZZ}^{-1}\tilde{S}_{ZZ}^{-1}Z_s^{\dagger}\varepsilon_s^{\dagger\prime}\varepsilon_s^{\dagger} + \frac{2h_2^{1/2}}{T}\sum_{1\le r<s\le T}\varepsilon_s^{\dagger}Z_s^{\dagger\prime}\tilde{S}_{ZZ}^{-1}\tilde{S}_{ZZ}^{-1}Z_r^{\dagger}\varepsilon_r^{\dagger\prime}\varepsilon_s^{\dagger} \\
&+ \frac{h_2^{1/2}}{T}\sum_{s=1}^{T}\sum_{r=1}^{T}\left(\tilde{Z}_s - Z_s^{\dagger}\right)'\tilde{S}_{ZZ}^{-1}\tilde{S}_{ZZ}^{-1}\left(\tilde{Z}_r - Z_r^{\dagger}\right)\varepsilon_r^{\dagger\prime}\varepsilon_s^{\dagger} + \frac{2h_2^{1/2}}{T}\sum_{s=1}^{T}\sum_{r=1}^{T}Z_s^{\dagger\prime}\tilde{S}_{ZZ}^{-1}\tilde{S}_{ZZ}^{-1}\left(\tilde{Z}_r - Z_r^{\dagger}\right)\varepsilon_r^{\dagger\prime}\varepsilon_s^{\dagger} \\
&\equiv M_{32,1}^{(1)} + M_{32,1}^{(2)} + M_{32,1}^{(3)} + M_{32,1}^{(4)}.
\end{aligned}
$$

First, we observe that $M_{32,1}^{(1)} = \frac{h_2^{1/2}}{T}\sum_{s=1}^{T}Z_s^{\dagger\prime}\tilde{S}_{ZZ}^{-1}\tilde{S}_{ZZ}^{-1}Z_s^{\dagger}\varepsilon_s^{\dagger\prime}\varepsilon_s^{\dagger} = \mathbb{B}_{3T}^{(2)}$. We keep $\mathbb{B}_{3T}^{(2)}$ as a bias term to be corrected in finite samples despite the fact that it is $O_P(h_2^{1/2})$ and thus asymptotically negligible. Next, following the analysis of $M_{31,1}^{(2)}$ in the proof of Proposition S1.1, we can readily show that $M_{32,1}^{(2)} = o_P(h_2^{1/2})$. In addition, we can readily show that

$$
\begin{aligned}
\left|M_{32,1}^{(3)}\right| &\lesssim Th_2^{1/2}\left\|\frac{1}{T}\sum_{r=1}^{T}\left(\tilde{Z}_r - Z_r^{\dagger}\right)\varepsilon_r^{\dagger\prime}\right\|^2 \lesssim Th_2^{1/2}\sum_{j=1}^{p}\left\|\frac{1}{T}\sum_{r=1}^{T}\left(\tilde{F}_r - HF_r\right)\varepsilon_r'\right\|^2 \\
&= Th_2^{1/2}\left[O_P(C_{0NT}^{-4}) + o_P(a_{1NT}^2)\right] = o_P(1).
\end{aligned}
$$

By the CS inequality, $\left|M_{32,1}^{(4)}\right| \le \left\{M_{32,1}^{(23)}(M_{32,1}^{(1)} + M_{32,1}^{(2)})\right\}^{1/2} = o_P(1)$. In sum, we conclude that $M_{32,1} = \mathbb{B}_{3T}^{(2)} + o_P(1)$.

Then, we prove (ii). Using $a_{2T} = T^{-1/2}h_2^{-1/4}$ and $\tilde{Z}_s = Z_s^{\dagger} + (\tilde{Z}_s - Z_s^{\dagger})$, we have

$$
\begin{aligned}
M_{32,2} &= Th_2^{1/2}\|D_2\|^2 = Th^{1/2}a_{2T}^2\left\|\tilde{S}_{ZZ}^{-1}\frac{1}{T}\sum_{s=1}^{T}\tilde{Z}_s\tilde{Z}_s'\tilde{g}_2(s/T)\right\|^2 \\
&= \left\|\tilde{S}_{ZZ}^{-1}\frac{1}{T}\sum_{s=1}^{T}Z_s^{\dagger}Z_s^{\dagger\prime}\tilde{g}_2(s/T)\right\|^2 + \left\|\tilde{S}_{ZZ}^{-1}\frac{1}{T}\sum_{s=1}^{T}\left(\tilde{Z}_s\tilde{Z}_s' - Z_s^{\dagger}Z_s^{\dagger\prime}\right)\tilde{g}_2(s/T)\right\|^2 \\
&\quad + 2\mathrm{tr}\left(\tilde{S}_{ZZ}^{-1}\frac{1}{T}\sum_{s=1}^{T}Z_s^{\dagger}Z_s^{\dagger\prime}\tilde{g}_2(s/T)\frac{1}{T}\sum_{r=1}^{T}\tilde{g}_2'(r/T)\left(\tilde{Z}_r\tilde{Z}_r' - Z_r^{\dagger}Z_r^{\dagger\prime}\right)\tilde{S}_{ZZ}^{-1}\right) \equiv \sum_{\ell=1}^{3}M_{32,2}^{(\ell)}.
\end{aligned}
$$

First,

$$
\begin{aligned}
\tilde{S}_{ZZ} &= \frac{1}{T}\sum_{s=1}^{T}\tilde{Z}_s\tilde{Z}_s' = \frac{1}{T}\sum_{s=1}^{T}Z_s^{\dagger}Z_s^{\dagger\prime} + O_P(C_{0NT}^{-1}) \\
&= (\mathbb{I}_p\otimes D_H')\frac{1}{T}\sum_{s=1}^{T}Z_s^0Z_s^{0\prime}(\mathbb{I}_p\otimes D_H) + O_P(C_{0NT}^{-1}) \\
&= (\mathbb{I}_p\otimes D_H')\frac{1}{T}\sum_{s=1}^{T}E\left(Z_s^0Z_s^{0\prime}\right)(\mathbb{I}_p\otimes D_H) + O_P(C_{0NT}^{-1}) \\
&\equiv S_T + O_P(C_{0NT}^{-1}).
\end{aligned}
$$

Noting that $\tilde{g}_2(s/T)' = D_H' g_2(s/T)'(\mathbb{I}_p\otimes D_H'^{-1})$, we have

$$
\begin{aligned}
M_{32,2}^{(1)} &= \left\|\tilde{S}_{ZZ}^{-1}\frac{1}{T}\sum_{s=1}^{T}Z_s^{\dagger}Z_s^{\dagger\prime}\tilde{g}_2(s/T)\right\|^2 \\
&= \left\|\tilde{S}_{ZZ}^{-1}(\mathbb{I}_p\otimes D_H')\frac{1}{T}\sum_{s=1}^{T}Z_s^0Z_s^{0\prime}(\mathbb{I}_p\otimes D_H)(\mathbb{I}_p\otimes D_H^{-1})g_2(s/T)D_H\right\|^2 \\
&= \left\|\tilde{S}_{ZZ}^{-1}(\mathbb{I}_p\otimes D_H')\frac{1}{T}\sum_{s=1}^{T}Z_s^0Z_s^{0\prime}g_2(s/T)D_H\right\|^2 \\
&= \left\|\left[S_T^{-1}+O_P(C_{0NT}^{-1})\right](\mathbb{I}_p\otimes D_H')\frac{1}{T}\sum_{s=1}^{T}E\left(Z_s^0Z_s^{0\prime}\right)g_2(s/T)D_H\right\|^2 + O_P(T^{-1/2}) \\
&= \left\|\left[(\mathbb{I}_p\otimes D_H')\frac{1}{T}\sum_{s=1}^{T}E\left(Z_s^0Z_s^{0\prime}\right)(\mathbb{I}_p\otimes D_H)\right]^{-1}(\mathbb{I}_p\otimes D_H')\frac{1}{T}\sum_{s=1}^{T}E\left(Z_s^0Z_s^{0\prime}\right)g_2(s/T)D_H\right\|^2 \\
&\quad + O_P\left(C_{0NT}^{-1}\right) \\
&= \left\|\left(\mathbb{I}_p\otimes D_H^{-1}\right)\left[\frac{1}{T}\sum_{s=1}^{T}E\left(Z_s^0Z_s^{0\prime}\right)\right]^{-1}\frac{1}{T}\sum_{s=1}^{T}E\left(Z_s^0Z_s^{0\prime}\right)g_2(s/T)D_H\right\|^2 + O_P\left(C_{0NT}^{-1}\right).
\end{aligned}
$$

Under the local alternative $\mathbb{H}_A^{(1)}$, we can show that $E\left(Z_s^0Z_s^{0\prime}\right) = E^0\left(Z_s^0Z_s^{0\prime}\right) + O\left(a_{1NT}+a_{2T}\right)$ $\equiv \Sigma_{ZZ} + O\left(a_{1NT}+a_{2T}\right)$ uniformly in s, where E^0 denotes expectation taken under $\mathbb{H}_0^{(1)}$. This implies that the leading term in the last displayed line is given by

$$
\left\|\left(\mathbb{I}_p\otimes D_H^{-1}\right)\Sigma_{ZZ}^{-1}\frac{1}{T}\sum_{s=1}^{T}\Sigma_{ZZ}g_2(s/T)D_H\right\|^2 + O_P\left(a_{1NT}+a_{2T}\right)
$$

$$= \left\| \left(\mathbb{I}_p \otimes D_H^{-1}\right) \frac{1}{T} \sum_{s=1}^{T} g_2(s/T) D_H \right\|^2 + O_P\left(a_{1NT} + a_{2T}\right) = o_P\left(1\right)$$

where the second equality follows from the fact that $\frac{1}{T}\sum_{s=1}^{T} g_2(s/T) = \int_0^1 g_2(u)du + O(T^{-1})$ and $\int_0^1 g_2(u)du = 0$ by the normalization condition. Next, we can readily show that

$$M_{32,2}^{(2)} \leq \left\| \tilde{S}_{ZZ}^{-1} \frac{1}{T} \sum_{s=1}^{T} \left(\tilde{Z}_s \tilde{Z}_s' - Z_s^{\dagger} Z_s^{\dagger\prime} \right) \tilde{g}_2(s/T) \right\|^2 = O_P\left(C_{0NT}^{-2}\right).$$

In addition, $\left| M_{32,2}^{(3)} \right| \leq \{M_{32,2}^{(1)} M_{32,2}^{(2)}\}^{1/2} = o_P\left(1\right)$ by the CS inequality.

Finally, we prove (iii). For $M_{32,3}$, we can show that

$$\begin{aligned} M_{32,3} &= Th_2^{1/2} \left\| \tilde{S}_{ZZ}^{-1} \frac{1}{T} \sum_{s=1}^{T} \tilde{Z}_s (\tilde{W}_s - W_s^{\dagger})' \right\|^2 \\ &\lesssim Th_2^{1/2} \left\| \frac{1}{T} \sum_{s=1}^{T} Z_s^{\dagger} (\tilde{W}_s - W_s^{\dagger})' \right\|^2 + Th_2^{1/2} \left\| \frac{1}{T} \sum_{s=1}^{T} \left(\tilde{Z}_s - Z_s^{\dagger} \right) (\tilde{W}_s - W_s^{\dagger})' \right\|^2 \\ &= Th_2^{1/2} \left[O_P\left(C_{0NT}^{-2}\right) + o_P\left(a_{1NT}\right) \right]^2 = o_P\left(1\right). \end{aligned}$$

Analogously, we have

$$M_{32,4} = Th_2^{1/2} \left\| \tilde{S}_{ZZ}^{-1} \frac{1}{T} \sum_{s=1}^{T} \tilde{Z}_s (\tilde{Z}_s - Z_s^{\dagger})' \Psi_0 \right\|^2 = Th_2^{1/2} \left[O_P\left(C_{0NT}^{-2}\right) + o_P\left(a_{1NT}\right) \right]^2 = o_P\left(1\right),$$

and

$$M_{32,5} = a_{2T}^2 Th_2^{1/2} \left\| \tilde{S}_{ZZ}^{-1} \frac{1}{T} \sum_{s=1}^{T} \tilde{Z}_s (\tilde{Z}_s - Z_s^{\dagger})' g_2(s/T) \right\|^2 = O_P\left(C_{0NT}^{-2}\right) + o_P\left(a_{1NT}\right) = o_P\left(1\right).$$

These results, along with the results in (i) and (ii), imply that $M_{31,\ell} = o_P\left(1\right)$ for $\ell = 6, 7, ..., 15$ by the CS inequality.

Consequently, we have shown that $M_{32} - \mathbb{B}_{3T}^{(2)} = o_P(1)$ under $\mathbb{H}_A^{(2)}(a_{2T})$. ■

Proof of Proposition S1.3. We note that

$$M_{33} = h_2^{1/2} \sum_{t=1}^{T} \operatorname{tr}\left[\left(\breve{\Psi}_t - \Psi_0 \right) \left(\tilde{\Psi}_0 - \Psi_0 \right)' \right]$$

$$= h_2^{1/2} \sum_{t=1}^{T} \mathrm{tr}\left[\left(D_{1t} + D_{2t} + D_{3t} + D_{4t} + D_{5t}\right)\left(D_1 + D_2 + D_3 + D_4 + D_5\right)'\right]$$
$$= \sum_{\ell=1}^{5} \sum_{\hbar=1}^{5} h_2^{1/2} \sum_{t=1}^{T} \mathrm{tr}\left(D_{\ell t} D_{\hbar}'\right) \equiv \sum_{\ell=1}^{5} \sum_{\hbar=1}^{5} M_{33,\ell\hbar}.$$

We prove the proposition by showing that (i) $M_{33,11} = \mathbb{B}_{3T}^{(3)} + o_p(1)$; (ii) $M_{33,22} = o_p(1)$; (iii) $M_{33,\ell\hbar} = o_P(1)$ for all other combinations of ℓ and $\hbar$.

We first prove (i) by making the following decomposition:

$$\begin{aligned}
M_{33,11} &= h_2^{1/2} \sum_{t=1}^{T} \mathrm{tr}\left(D_{1t} D_1'\right) \\
&= h_2^{1/2} \sum_{t=1}^{T} \mathrm{tr}\left[\left(\tilde{S}_{ZZ,t}^{-1} \frac{1}{T} \sum_{s=1}^{T} k_{h_2,st}^{\dagger} \tilde{Z}_s \varepsilon_s^{\dagger\prime}\right)\left(\tilde{S}_{ZZ}^{-1} \frac{1}{T} \sum_{r=1}^{T} \tilde{Z}_r \varepsilon_r^{\dagger\prime}\right)'\right] \\
&= \frac{h_2^{1/2}}{T^2} \sum_{t=1}^{T} \sum_{s=1}^{T} \mathrm{tr}\left(\tilde{S}_{ZZ,t}^{-1} k_{h_2,st}^{\dagger} Z_s^{\dagger} \varepsilon_s^{\dagger\prime} \varepsilon_s^{\dagger} Z_s^{\dagger\prime} \tilde{S}_{ZZ}^{-1}\right) + \frac{2h_2^{1/2}}{T^2} \sum_{t=1}^{T} \sum_{1 \le r < s \le T} k_{h_2,st}^{\dagger} \mathrm{tr}\left(\tilde{S}_{ZZ,t}^{-1} Z_s^{\dagger} \varepsilon_s^{\dagger\prime} \varepsilon_r^{\dagger} Z_r^{\dagger\prime} \tilde{S}_{ZZ}^{-1}\right) \\
&\quad + h_2^{1/2} \sum_{t=1}^{T} \mathrm{tr}\left[\left(\tilde{S}_{ZZ,t}^{-1} \frac{1}{T} \sum_{s=1}^{T} k_{h_2,st}^{\dagger}\left(\tilde{Z}_s - Z_s^{\dagger}\right) \varepsilon_s^{\dagger\prime}\right)\left(\tilde{S}_{ZZ}^{-1} \frac{1}{T} \sum_{r=1}^{T} \left(\tilde{Z}_r - Z_r^{\dagger}\right) \varepsilon_r^{\dagger\prime}\right)'\right] \\
&\quad + h_2^{1/2} \sum_{t=1}^{T} \mathrm{tr}\left[\left(\tilde{S}_{ZZ,t}^{-1} \frac{1}{T} \sum_{s=1}^{T} k_{h_2,st}^{\dagger} Z_s^{\dagger} \varepsilon_s^{\dagger\prime}\right)\left(\tilde{S}_{ZZ}^{-1} \frac{1}{T} \sum_{r=1}^{T} \left(\tilde{Z}_r - Z_r^{\dagger}\right) \varepsilon_r^{\dagger\prime}\right)'\right] \\
&\quad + h_2^{1/2} \sum_{t=1}^{T} \mathrm{tr}\left[\left(\tilde{S}_{ZZ,t}^{-1} \frac{1}{T} \sum_{s=1}^{T} k_{h_2,st}^{\dagger}\left(\tilde{Z}_s - Z_s^{\dagger}\right) \varepsilon_s^{\dagger\prime}\right)\left(\tilde{S}_{ZZ}^{-1} \frac{1}{T} \sum_{r=1}^{T} Z_r^{\dagger} \varepsilon_r^{\dagger\prime}\right)'\right] \equiv \sum_{\ell=1}^{5} M_{33,11}^{(\ell)}.
\end{aligned}$$

First, we note that $M_{33,11}^{(1)} = \frac{h_2^{1/2}}{T^2} \sum_{t=1}^{T} \sum_{s=1}^{T} k_{h_2,st}^{\dagger} Z_s^{\dagger\prime} \tilde{S}_{ZZ}^{-1} \tilde{S}_{ZZ,t}^{-1} Z_s^{\dagger} \varepsilon_s^{\dagger\prime} \varepsilon_s^{\dagger} = \mathbb{B}_{3T}^{(3)}$. Like $\mathbb{B}_{3T}^{(2)}$, $\mathbb{B}_{3T}^{(3)}$ is $O_P(h_2^{1/2})$ and thus asymptotically negligible, but we keep it to be corrected in finite samples. For $M_{33,12}^{(2)}$, we can follow the analysis of $M_{31,1}^{(2)}$ in the proof of Proposition S1.1 and show that $M_{33,11}^{(2)} = o_P(h_2^{1/2})$. Similarly, we can follow the analysis of $M_{31,1}^{(3)}$ and $M_{31,1}^{(4)}$ and show that $M_{33,11}^{(j)} = o_P(1)$ for $j = 3, 4, 5$. Consequently, we have shown that $M_{33,1} = \mathbb{B}_{3T}^{(3)} + o_P(1)$.

Then, we prove (ii). The $M_{33,22}$ term could be decomposed as the follows:

$$\begin{aligned}
M_{33,2} &= h_2^{1/2} \sum_{t=1}^{T} \mathrm{tr}\left(D_{2t} D_{20}'\right) \\
&= a_{2T}^2 h_2^{1/2} \sum_{t=1}^{T} \mathrm{tr}\left[\left(\tilde{S}_{ZZ,t}^{-1} \frac{1}{T} \sum_{s=1}^{T} k_{h_2,st}^{\dagger} \tilde{Z}_s \tilde{Z}_s' \tilde{g}_2(s/T)\right)\left(\tilde{S}_{ZZ}^{-1} \frac{1}{T} \sum_{r=1}^{T} \tilde{Z}_r \tilde{Z}_r' \tilde{g}_2(r/T)\right)'\right]
\end{aligned}$$

$$
\begin{aligned}
&= \operatorname{tr}\left[\left(T^{-1}\sum_{t=1}^{T}\tilde{S}_{ZZ,t}^{-1}\frac{1}{T}\sum_{s=1}^{T}k_{h_2,st}^{\dagger}\tilde{Z}_s\tilde{Z}_s'\tilde{g}_2(s/T)\right)\frac{1}{T}\sum_{r=1}^{T}\tilde{g}_2(r/T)'\tilde{Z}_r\tilde{Z}_r'\tilde{S}_{ZZ}^{-1}\right] \\
&= \operatorname{tr}\left[\left(T^{-1}\sum_{t=1}^{T}\tilde{S}_{ZZ,t}^{-1}\frac{1}{T}\sum_{s=1}^{T}k_{h_2,st}^{\dagger}\tilde{Z}_s\tilde{Z}_s'\tilde{g}_2(t/T)\right)\frac{1}{T}\sum_{r=1}^{T}\tilde{g}_2(r/T)'\tilde{Z}_r\tilde{Z}_r'\tilde{S}_{ZZ}^{-1}\right] \\
&+ \operatorname{tr}\left[\left(T^{-1}\sum_{t=1}^{T}\tilde{S}_{ZZ,t}^{-1}\frac{1}{T}\sum_{s=1}^{T}k_{h_2,st}^{\dagger}\tilde{Z}_s\tilde{Z}_s'\left[\tilde{g}_2(s/T)-\tilde{g}_2(t/T)\right]\right)\frac{1}{T}\sum_{r=1}^{T}\tilde{g}_2(r/T)'\tilde{Z}_r\tilde{Z}_r'\tilde{S}_{ZZ}^{-1}\right] \\
&\equiv M_{33,22}^{(1)}+M_{33,22}^{(2)}.
\end{aligned}
$$

Noting that $T^{-1}\sum_{t=1}^{T}\tilde{g}_2(t/T) = \left(\mathbb{I}_p\otimes D_H^{-1}\right)\frac{1}{T}\sum_{s=1}^{T}g_2(s/T)D_H = \left(\mathbb{I}_p\otimes D_H^{-1}\right)[\int_0^1 g_2(u)du + O(T^{-1})]D_H = O_P(T^{-1})$, we have

$$
M_{33,22}^{(1)} = \operatorname{tr}\left[\left(T^{-1}\sum_{t=1}^{T}\tilde{g}_2(t/T)\right)\frac{1}{T}\sum_{r=1}^{T}\tilde{g}_2(r/T)'\tilde{Z}_r\tilde{Z}_r'\tilde{S}_{ZZ}^{-1}\right] = O_P(T^{-1}).
$$

In addition, we can readily show that $M_{33,22}^{(2)} = O_P((Th_2)^{-1/2}+h_2^2) = o_P(1)$. It follows that $M_{33,22} = o_P(1)$.

Finally, we show (iii). By the results in the proofs of Propositions S1.1–S1.2, we have

$$
\begin{aligned}
M_{31,1} &= O_P(h_2^{-1/2}),\ M_{31,2}=O_P(1),\ M_{31,\ell}=o_P(1) \text{ for } \ell\in\{3,4,5\}, \text{ and} \\
M_{31,\hbar} &= o_P(1) \text{ for } \hbar\in\{1,2,3,4,5\}.
\end{aligned}
$$

Then by the CS inequality $|M_{33,\ell\hbar}| \le \{M_{31,\ell}M_{32,\hbar}\}^{1/2} = o_P(1)$ for $\ell\in\{2,3,4,5\}$ and $\hbar\in\{1,2,3,4,5\}$. For the other combinations of ℓ and $\hbar$ that are not considered here or in (i) and (ii), we can apply arguments as used above to show that $M_{33,\ell\hbar}=o_P(1)$ directly.

In sum, we have shown that $M_{33}-\mathbb{B}_{3T}^{(3)} = o_P(1)$ under $\mathbb{H}_A^{(2)}(a_{2T})$. ■

Proof of Proposition S1.4. We note that

$$
\mathbb{B}_{3T} = \mathbb{B}_{3T}^{(1)}+\mathbb{B}_{3T}^{(2)}-2\mathbb{B}_{3T}^{(3)} = \frac{h_2^{1/2}}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}Z_s^{\dagger\prime}\left(k_{h_2,st}^{\dagger}\tilde{S}_{ZZ,t}^{-1}-\tilde{S}_{ZZ}^{-1}\right)\left(k_{h_2,st}^{\dagger}\tilde{S}_{ZZ,t}^{-1}-\tilde{S}_{ZZ}^{-1}\right)Z_s^{\dagger}\varepsilon_s^{\dagger\prime}\varepsilon_s^{\dagger},
$$

where we recall that $\tilde{S}_{ZZ,t}=\frac{1}{T}\sum_{s=1}^{T}k_{h_2,st}^{\dagger}\tilde{Z}_s\tilde{Z}_s'$ and $\tilde{S}_{ZZ}=T^{-1}\sum_{s=1}^{T}\tilde{Z}_s\tilde{Z}_s'$. Note that

$$
\hat{\mathbb{B}}_{3T} = \frac{h_2^{1/2}}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\tilde{Z}_s'\left(k_{h_2,st}^{\dagger}\tilde{S}_{ZZ,t}^{-1}-\tilde{S}_{ZZ}^{-1}\right)\left(k_{h_2,st}^{\dagger}\tilde{S}_{ZZ,t}^{-1}-\tilde{S}_{ZZ}^{-1}\right)\tilde{Z}_s\tilde{U}_s'\tilde{U}_s.
$$

Let $L_{st} = \left(k^{\dagger}_{h_2,st}\tilde{S}^{-1}_{ZZ,t} - \tilde{S}^{-1}_{ZZ}\right)\left(k^{\dagger}_{h_2,st}\tilde{S}^{-1}_{ZZ,t} - \tilde{S}^{-1}_{ZZ}\right)$. Using $\tilde{U}_s'\tilde{U}_s - \varepsilon^{\dagger\prime}_s\varepsilon^{\dagger}_s = (\tilde{U}_s - \varepsilon^{\dagger}_s)'(\tilde{U}_s - \varepsilon^{\dagger}_s) + 2(\tilde{U}_s - \varepsilon^{\dagger}_s)'\varepsilon^{\dagger}_s$, we have

$$\begin{aligned}
\hat{\mathbb{B}}_{3T} - \mathbb{B}_{3T} &= \frac{h_2^{1/2}}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\left(\tilde{Z}_s'L_{st}\tilde{Z}_s\tilde{U}_s'\tilde{U}_s - Z^{\dagger\prime}_sL_{st}Z^{\dagger}_s\varepsilon^{\dagger\prime}_s\varepsilon^{\dagger}_s\right) \\
&= \frac{h_2^{1/2}}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\left[\tilde{Z}_s'L_{st}\tilde{Z}_s\left(\tilde{U}_s'\tilde{U}_s - \varepsilon^{\dagger\prime}_s\varepsilon^{\dagger}_s\right) + \left(\tilde{Z}_s'L_{st}\tilde{Z}_s - Z^{\dagger\prime}_sL_{st}Z^{\dagger}_s\right)\varepsilon^{\dagger\prime}_s\varepsilon^{\dagger}_s\right] \\
&= \frac{h_2^{1/2}}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\Big[\tilde{Z}_s'L_{st}\tilde{Z}_s(\tilde{U}_s - \varepsilon^{\dagger}_s)'(\tilde{U}_s - \varepsilon^{\dagger}_s) + 2\tilde{Z}_s'L_{st}\tilde{Z}_s(\tilde{U}_s - \varepsilon^{\dagger}_s)'\varepsilon^{\dagger}_s \\
&\quad + \left(\tilde{Z}_s'L_{st}\tilde{Z}_s - Z^{\dagger\prime}_sL_{st}Z^{\dagger}_s\right)\varepsilon^{\dagger\prime}_s\varepsilon^{\dagger}_s\Big] \equiv B_{31} + 2B_{32} + B_{33}.
\end{aligned}$$

It is suffices to show that (i) $B_{31} = o_P(1)$;(ii) $B_{32} = o_P(1)$;(iii) $B_{33} = o_P(1)$.

We first show (i). Noting that $\Psi^{\dagger}_s = \Psi_0 + a_{2T}\tilde{g}_2(s/T)$ by (S2.10), $\tilde{W}_s = \Psi^{\dagger\prime}_s\tilde{Z}_s + U^{\dagger}_s = \Psi_0'\tilde{Z}_s + a_{2T}\tilde{g}_2(s/T)'\tilde{Z}_s + U^{\dagger}_s$ by (S2.11), and $U^{\dagger}_s \equiv (\tilde{W}_s - W^{\dagger}_s) - \Psi^{\dagger\prime}_s(\tilde{Z}_s - Z^{\dagger}_t) + \varepsilon^{\dagger}_s$ by (S2.12), we have

$$\begin{aligned}
\tilde{U}_s - \varepsilon^{\dagger}_s &= (\tilde{W}_s - \tilde{\Psi}_0'\tilde{Z}_s) - \varepsilon^{\dagger}_s \\
&= \left[\Psi_0'\tilde{Z}_s + a_{2T}\tilde{g}_2(s/T)'\tilde{Z}_s + U^{\dagger}_s - \tilde{\Psi}_0'\tilde{Z}_s\right] - \varepsilon^{\dagger}_s \\
&= (\Psi_0 - \tilde{\Psi}_0)'\tilde{Z}_s + a_{2T}\tilde{g}_2(s/T)'\tilde{Z}_s + (\tilde{W}_s - W^{\dagger}_s) - \Psi^{\dagger\prime}_s(\tilde{Z}_s - Z^{\dagger}_t) \\
&\equiv b_{1s} + b_{2s} + b_{3s} + b_{4s}. \qquad \text{(S2.18)}
\end{aligned}$$

By the CS inequality, we have $B_{31} \le 4\sum_{\ell=1}^{4}\frac{h_2^{1/2}}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\tilde{Z}_s'L_{st}\tilde{Z}_sb_{\ell s}'b_{\ell s} \equiv 4\sum_{\ell=1}^{4}B_{31,\ell}$. For $B_{31,1}$, it suffices to consider the rough probability bound:

$$\begin{aligned}
B_{31,1} &= \frac{h_2^{1/2}}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\tilde{Z}_s'L_{st}\tilde{Z}_s\tilde{Z}_s'\left(\tilde{\Psi}_0 - \Psi_0\right)\left(\tilde{\Psi}_0 - \Psi_0\right)'\tilde{Z}_s \\
&= \frac{h_2^{1/2}}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\left\|\left(k^{\dagger}_{h_2,st}\tilde{S}^{-1}_{ZZ,t} - \tilde{S}^{-1}_{ZZ}\right)\tilde{Z}_s\tilde{Z}_s'\left(\tilde{\Psi}_0 - \Psi_0\right)\right\|^2 \\
&\le h_2^{1/2}\left\|\tilde{\Psi}_0 - \Psi_0\right\|^2\left\{\max_s\frac{1}{T}\sum_{t=1}^{T}\left\|k^{\dagger}_{h_2,st}\tilde{S}^{-1}_{ZZ,t} - \tilde{S}^{-1}_{ZZ}\right\|^2\right\}\frac{1}{T}\sum_{s=1}^{T}\left\|\tilde{Z}_s\tilde{Z}_s'\right\|^2 \\
&= h_2^{1/2}o_P(T^{-1}h_2^{-1/2})O_P\left(h_2^{-1}\right)O_P(1) = o_P(T^{-1}h_2^{-1}) = o_P(1),
\end{aligned}$$

where we use the fact that $Th_2^{1/2}\left\|\tilde{\Psi}_0 - \Psi_0\right\|^2 = o_P(1)$ by the proof of Proposition S1.2. Next, we study $B_{31,2}$. By the matrix version of CS inequality and the fact that $a_{2T}^2 h_2^{1/2} = T^{-1}$, we have

$$\begin{aligned} B_{31,2} &= \frac{a_{2T}^2 h_2^{1/2}}{T^2} \sum_{t=1}^{T} \sum_{s=1}^{T} \tilde{Z}_s' \left(k_{h_2,st}^{\dagger} \tilde{S}_{ZZ,t}^{-1} - \tilde{S}_{ZZ}^{-1} \right) \left(k_{h_2,st}^{\dagger} \tilde{S}_{ZZ,t}^{-1} - \tilde{S}_{ZZ}^{-1} \right) \tilde{Z}_s \tilde{Z}_s' \tilde{g}_2(s/T) \tilde{g}_2(s/T)' \tilde{Z}_s \\ &\leq \frac{2}{T^3} \sum_{t=1}^{T} \sum_{s=1}^{T} \tilde{Z}_s' \left(k_{h_2,st}^{\dagger 2} \tilde{S}_{ZZ,t}^{-1} \tilde{S}_{ZZ,t}^{-1} + \tilde{S}_{ZZ}^{-1} \tilde{S}_{ZZ}^{-1} \right) \tilde{Z}_s \tilde{Z}_s' \tilde{g}_2(s/T) \tilde{g}_2(s/T)' \tilde{Z}_s \equiv 2 \left(B_{31,2}^{(1)} + B_{31,2}^{(2)} \right). \end{aligned}$$

For the first term, we have

$$\begin{aligned} B_{31,2}^{(1)} &= \frac{1}{T^3} \sum_{t=1}^{T} \sum_{s=1}^{T} \tilde{Z}_s' k_{h_2,st}^{\dagger 2} \tilde{S}_{ZZ,t}^{-1} \tilde{S}_{ZZ,t}^{-1} \tilde{Z}_s \tilde{Z}_s' \tilde{g}_2(s/T) \tilde{g}_2(s/T)' \tilde{Z}_s \\ &\leq \max_t \left\| \tilde{S}_{ZZ,t}^{-1} \tilde{S}_{ZZ,t}^{-1} \right\| \frac{1}{T^3} \sum_{t=1}^{T} \sum_{s=1}^{T} k_{h_2,st}^{\dagger 2} \left\| \tilde{Z}_s \tilde{Z}_s' \tilde{g}_2(s/T) \right\|^2 = O_P(1)\, O_P(T^{-1} h_2^{-1}). \end{aligned}$$

Similarly, we can show that $B_{31,2}^{(2)} = O_P(T^{-1})$. Then $B_{31,2} = o_P(1)$. For $B_{31,3}$, we use the fact that $\tilde{S}_{ZZ,t}$ is asymptotically nonsingular uniformly in t and that $\tilde{S}_{ZZ}^{-1}$ is asymptotically nonsingular as demonstrated in the proofs of Propositions S1.1–S1.2,

$$\begin{aligned} B_{31,3} &= \frac{h_2^{1/2}}{T^2} \sum_{t=1}^{T} \sum_{s=1}^{T} \left\| \left(k_{h_2,st}^{\dagger} \tilde{S}_{ZZ,t}^{-1} - \tilde{S}_{ZZ}^{-1} \right) \tilde{Z}_s (\tilde{W}_s - W_s)' \right\|^2 \\ &\lesssim \frac{h_2^{1/2}}{T^2} \sum_{t=1}^{T} \sum_{s=1}^{T} \left(k_{h_2,st}^{\dagger 2} + 1 \right) \left\| \tilde{Z}_s (\tilde{W}_s - W_s)' \right\|^2. \end{aligned}$$

Following the proof of Lemma A.6(iv) in Su and Wang (2020a), we can readily show that $\frac{1}{T^2} \sum_{s=1}^{T} ||\tilde{Z}_s\ (\tilde{W}_s - W_s)'||^2 = O_P(C_{0NT}^{-2} + TN^{-2})$ and $\frac{1}{T^2} \sum_{t=1}^{T} \sum_{s=1}^{T} k_{h_2,st}^{\dagger 2} ||\tilde{Z}_s (\tilde{W}_s - W_s)'||^2 = O_P(h_2^{-1}(C_{0NT}^{-2} + TN^{-2}))$. Then $B_{31,3} = O_P(h_2^{-1/2}(C_{0NT}^{-2} + TN^{-2})) = o_P(1)$ under Assumption A.5. Analogously, we can show that $B_{31,4} = O_P(h_2^{-1/2}(C_{0NT}^{-2} + TN^{-2})) = o_P(1)$. In sum, we have shown that $B_{31} = o_P(1)$.

Now, we show (ii) by making the following decomposition:

$$B_{32} = \sum_{\ell=1}^{4} \frac{h_2^{1/2}}{T^2} \sum_{t=1}^{T} \sum_{s=1}^{T} \tilde{Z}_s' \left(k_{h_2,st}^{\dagger} \tilde{S}_{ZZ,t}^{-1} - \tilde{S}_{ZZ}^{-1} \right) \left(k_{h_2,st}^{\dagger} \tilde{S}_{ZZ,t}^{-1} - \tilde{S}_{ZZ}^{-1} \right) \tilde{Z}_s b_{\ell s}' U_s \equiv \sum_{\ell=1}^{4} B_{32,\ell}.$$

For $B_{32,1}$ and $B_{32,2}$, it suffices to consider the rough probability bounds:

$$\begin{aligned}
|B_{32,1}| &= \frac{h_2^{1/2}}{T^2}\left|\sum_{t=1}^{T}\sum_{s=1}^{T}\left[\tilde{Z}_s'\left(k_{h_2,st}^{\dagger}\tilde{S}_{ZZ,t}^{-1}-\tilde{S}_{ZZ}^{-1}\right)\left(k_{h_2,st}^{\dagger}\tilde{S}_{ZZ,t}^{-1}-\tilde{S}_{ZZ}^{-1}\right)\tilde{Z}_s\tilde{Z}_s'\left(\tilde{\Psi}_0-\Psi_0\right)\varepsilon_s^{\dagger}\right]\right| \\
&\leq h_2^{1/2}\max_s\frac{1}{T}\sum_{t=1}^{T}\left\|k_{h_2,st}^{\dagger}\tilde{S}_{ZZ,t}^{-1}-\tilde{S}_{ZZ}^{-1}\right\|^2\left\|\tilde{\Psi}_0-\Psi_0\right\|\frac{1}{T}\sum_{s=1}^{T}\left\|\tilde{Z}_s\right\|^3\left\|\varepsilon_s^{\dagger}\right\| \\
&= h_2^{1/2}O_P(h_2^{-1})O_P(T^{-1/2}h_2^{-1/4})O_P(1)=O_P(T^{-1/2}h_2^{-3/4})=o_P(1),
\end{aligned}$$

and

$$\begin{aligned}
|B_{32,2}| &= \frac{a_{2T}h_2^{1/2}}{T^2}\left|\sum_{t=1}^{T}\sum_{s=1}^{T}\tilde{Z}_s'\left(k_{h_2,st}^{\dagger}\tilde{S}_{ZZ,t}^{-1}-\tilde{S}_{ZZ}^{-1}\right)\left(k_{h_2,st}^{\dagger}\tilde{S}_{ZZ,t}^{-1}-\tilde{S}_{ZZ}^{-1}\right)\tilde{Z}_s\tilde{Z}_s'\tilde{g}_2(s/T)\varepsilon_s^{\dagger}\right| \\
&= \frac{h_2^{1/4}}{T^{5/2}}\left|\sum_{t=1}^{T}\sum_{s=1}^{T}\tilde{Z}_s'\left(k_{h_2,st}^{\dagger}\tilde{S}_{ZZ,t}^{-1}-\tilde{S}_{ZZ}^{-1}\right)\left(k_{h_2,st}^{\dagger}\tilde{S}_{ZZ,t}^{-1}-\tilde{S}_{ZZ}^{-1}\right)\tilde{Z}_s\tilde{Z}_s'\tilde{g}_2(s/T)\varepsilon_s^{\dagger}\right| \\
&\leq \frac{h_2^{1/4}}{T^{1/2}}\max_s\frac{1}{T}\sum_{t=1}^{T}\left\|k_{h_2,st}S_{Tt}^{-1}-S_{T0}^{-1}\right\|^2\frac{1}{T}\sum_{s=1}^{T}\left\|\tilde{g}_2(s/T)\right\|\left\|\tilde{Z}_s\right\|^3\left\|\varepsilon_s^{\dagger}\right\| \\
&= T^{-1/2}h_2^{1/4}O_P(h_2^{-1})=o_P(1).
\end{aligned}$$

For $B_{32,3}$, we have

$$\begin{aligned}
|B_{32,3}| &= \frac{h_2^{1/2}}{T^2}\left|\sum_{t=1}^{T}\sum_{s=1}^{T}\tilde{Z}_s'\left(k_{h_2,st}^{\dagger}\tilde{S}_{ZZ,t}^{-1}-\tilde{S}_{ZZ}^{-1}\right)\left(k_{h_2,st}^{\dagger}\tilde{S}_{ZZ,t}^{-1}-\tilde{S}_{ZZ}^{-1}\right)\tilde{Z}_s(\tilde{W}_s-W_s^{\dagger})'\varepsilon_s^{\dagger}\right| \\
&\leq h_2^{1/2}\max_s\frac{1}{T}\sum_{t=1}^{T}\left\|k_{h_2,st}^{\dagger}\tilde{S}_{ZZ,t}^{-1}-\tilde{S}_{ZZ}^{-1}\right\|^2\frac{1}{T}\sum_{s=1}^{T}(\left\|\tilde{W}_s-W_s\right\|\left\|\varepsilon_s^{\dagger}\right\|\left\|\tilde{Z}_s\right\|^2 \\
&\leq h_2^{1/2}\max_s\frac{1}{T}\sum_{t=1}^{T}\left\|k_{h_2,st}^{\dagger}\tilde{S}_{ZZ,t}^{-1}-\tilde{S}_{ZZ}^{-1}\right\|^2\left\{\frac{1}{T}\sum_{s=1}^{T}\left\|\tilde{W}_s-W_s\right\|^2\right\}^{1/2}\left\{\frac{1}{T}\sum_{s=1}^{T}\left\|\varepsilon_s^{\dagger}\right\|^2\left\|\tilde{Z}_s\right\|^4\right\}^{1/2} \\
&= h_2^{1/2}O_P(h_2^{-1})O_P(C_{0NT}^{-1})=O_P(h_2^{-1/2}C_{0NT}^{-1})=o_P(1).
\end{aligned}$$

Analogously, $B_{32,4}=O_P(h_2^{-1/2}C_{0NT}^{-1})=o_P(1)$. It follows that $B_{32}=o_P(1)$.

Finally, we show (iii) $B_{33}=o_P(1)$ by making the following decomposition:

$$B_{33} = \frac{h_2^{1/2}}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\left(\tilde{Z}_s'L_{st}\tilde{Z}_s-Z_s^{\dagger\prime}L_{st}Z_s^{\dagger}\right)\varepsilon_s^{\dagger\prime}\varepsilon_s^{\dagger}$$

$$= \frac{h_2^{1/2}}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\left[\left(\tilde{Z}_s - Z_s^\dagger\right)' L_{st}\left(\tilde{Z}_s - Z_s^\dagger\right) + 2\left(\tilde{Z}_s - Z_s^\dagger\right)' L_{st} Z_s^\dagger\right]\varepsilon_s^{\dagger\prime}\varepsilon_s^\dagger \equiv B_{33,1} + 2B_{33,2}.$$

Note that

$$\begin{aligned}
|B_{33,1}| &= \frac{h_2^{1/2}}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\left(\tilde{Z}_s - Z_s^\dagger\right)'\left(k_{h_2,st}^\dagger \tilde{S}_{ZZ,t}^{-1} - \tilde{S}_{ZZ}^{-1}\right)\left(k_{h_2,st}^\dagger \tilde{S}_{ZZ,t}^{-1} - \tilde{S}_{ZZ}^{-1}\right)\left(\tilde{Z}_s - Z_s^\dagger\right)\varepsilon_s^{\dagger\prime}\varepsilon_s^\dagger \\
&\lesssim \frac{h_2^{1/2}}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\left(k_{h_2,st}^{\dagger 2} + 1\right)\left(\tilde{Z}_s - Z_s^\dagger\right)'\left(\tilde{Z}_s - Z_s^\dagger\right)\varepsilon_s^{\dagger\prime}\varepsilon_s^\dagger \\
&\lesssim \frac{h_2^{-1/2}}{T}\sum_{s=1}^{T}\left(\tilde{Z}_s - Z_s^\dagger\right)'\left(\tilde{Z}_s - Z_s^\dagger\right)\varepsilon_s^{\dagger\prime}\varepsilon_s^\dagger \le h_2^{-1/2}\max_s \left\|\varepsilon_s^\dagger\right\|^2 \frac{1}{T}\sum_{s=1}^{T}\left\|\tilde{Z}_s - Z_s^\dagger\right\|^2 \\
&= h_2^{-1/2} O_P(T^{2/(8+\sigma)}) O_P(C_{0NT}^{-2}) = o_P(1),
\end{aligned}$$

where the second inequality follows from the fact that $\max_t \frac{1}{T}\sum_{t=1}^{T}\left(k_{h_2,st}^{\dagger 2} + 1\right) = O(h_2^{-1})$ and the next to last equality holds by the fact that $E\left\|\varepsilon_s\right\|^{8+\sigma} \le C$ and the Markov inequality. For $B_{33,2}$, we have

$$\begin{aligned}
|B_{33,2}| &= \frac{h_2^{1/2}}{T^2}\left|\sum_{t=1}^{T}\sum_{s=1}^{T}\left(\tilde{Z}_s - Z_s^\dagger\right)'\left(k_{h_2,st}^\dagger \tilde{S}_{ZZ,t}^{-1} - \tilde{S}_{ZZ}^{-1}\right)\left(k_{h_2,st}^\dagger \tilde{S}_{ZZ,t}^{-1} - \tilde{S}_{ZZ}^{-1}\right) Z_s^\dagger \varepsilon_s^{\dagger\prime}\varepsilon_s^\dagger\right| \\
&\le h_2^{1/2}\max_s \frac{1}{T}\sum_{t=1}^{T}\left\|k_{h_2,st}^\dagger \tilde{S}_{ZZ,t}^{-1} - \tilde{S}_{ZZ}^{-1}\right\|^2 \frac{1}{T}\sum_{s=1}^{T}\left\|\tilde{Z}_s - Z_s^\dagger\right\|\left\|Z_s^\dagger\right\|\left|\varepsilon_s^{\dagger\prime}\varepsilon_s^\dagger\right| \\
&\le h_2^{1/2}\max_s \frac{1}{T}\sum_{t=1}^{T}\left\|k_{h_2,st}^\dagger \tilde{S}_{ZZ,t}^{-1} - \tilde{S}_{ZZ}^{-1}\right\|^2 \left\{\frac{1}{T}\sum_{s=1}^{T}\left\|\tilde{Z}_s - Z_s^\dagger\right\|^2\right\}^{1/2}\left\{\frac{1}{T}\sum_{s=1}^{T}\left\|Z_s^\dagger\right\|^2\left|\varepsilon_s^{\dagger\prime}\varepsilon_s^\dagger\right|^2\right\}^{1/2} \\
&= h_2^{1/2} O_P(h_2^{-1}) O_P(C_{0NT}^{-1}) O_P(1) = o_P(1).
\end{aligned}$$

In sum, we have shown that $B_{33} = o_P(1)$.

Combining the above results yields that $\hat{\mathbb{B}}_{3T} - \mathbb{B}_{3T} = o_P(1)$. ■

Proof of Proposition S1.5. Recall that $\hat{\mathbb{S}}_T = \frac{1}{T}\sum_{t=1}^{T}\tilde{S}_{ZZ,t}^{-1}\tilde{S}_{ZZ,t}^{-1}$. By (S2.14) in the proof of Proposition S1.1, we can readily show that $\hat{\mathbb{S}}_T = (\mathbb{I}_p \otimes D_{Q^{-1}})\frac{1}{T}\sum_{t=1}^{T} S_{Tt}^{-1}S_{Tt}^{-1}(\mathbb{I}_p \otimes D_{Q^{-1}}') + O_P((Th_2/\ln T)^{-1/2})$. Observe that $\hat{\mathbb{V}}_{3T} - \mathbb{V}_{3T} = \mathbb{V}_{3T,1} + \mathbb{V}_{3T,2} + \mathbb{V}_{3T,3}$, where

$$\mathbb{V}_{3T,1} = 4T^{-2}h_2^{-1}\sum_{s=2}^{T}\sum_{r=1}^{s-1}\bar{k}_{sr}^2\left[\left(\tilde{Z}_s'\mathbb{S}_T\tilde{Z}_r\tilde{U}_r'\tilde{U}_s\right)^2 - \left(Z_s^{0\prime}\mathbb{S}_T Z_r^0\varepsilon_r^{\dagger\prime}\varepsilon_s^\dagger\right)^2\right],$$

$$\mathbb{V}_{3T,2} = 4T^{-2}h_2^{-1}\sum_{s=2}^{T}\sum_{r=1}^{s-1}\bar{k}_{sr}^2\left[\left(Z_s^{0\prime}\mathbb{S}_T Z_r^0\varepsilon_r^{\dagger\prime}\varepsilon_s^{\dagger}\right)^2 - \left(Z_s^{0\prime}\mathbb{S}_T Z_r^0\varepsilon_r^{\prime}D_{Q^{-1}}D_{Q^{-1}}^{\prime}\varepsilon_s\right)^2\right], \text{ and}$$

$$\mathbb{V}_{3T,3} = 4T^{-2}h_2^{-1}\sum_{s=2}^{T}\sum_{r=1}^{s-1}\bar{k}_{sr}^2\left[\left(Z_s^{0\prime}\mathbb{S}_T Z_r^0\varepsilon_r^{\prime}D_{Q^{-1}}D_{Q^{-1}}^{\prime}\varepsilon_s\right)^2 - E\left(Z_s^{0\prime}\mathbb{S}_T Z_r^0\varepsilon_r^{\prime}D_{Q^{-1}}D_{Q^{-1}}^{\prime}\varepsilon_s\right)^2\right].$$

It suffices to show that under $\mathbb{H}_A^{(2)}(a_{2T})$, we have: (i) $\mathbb{V}_{3T,1} = o_P(1)$, (ii) $\mathbb{V}_{3T,2} = o_P(1)$ and (iii) $\mathbb{V}_{3T,3} = o_P(1)$.

To show (i), we make the following decomposition:

$$\begin{aligned}
\mathbb{V}_{3T,1} &= 4T^{-2}h_2^{-1}\sum_{s=2}^{T}\sum_{r=1}^{s-1}\bar{k}_{sr}^2\left[\left(\tilde{Z}_s^{\prime}\hat{\mathbb{S}}_T\tilde{Z}_r\right)^2\left(\tilde{U}_r^{\prime}\tilde{U}_s\right)^2 - \left(Z_s^{0\prime}\mathbb{S}_T Z_r^0\right)^2\left(\varepsilon_r^{\dagger\prime}\varepsilon_s^{\dagger}\right)^2\right] \\
&= 4T^{-2}h_2^{-1}\sum_{s=2}^{T}\sum_{r=1}^{s-1}\bar{k}_{sr}^2\left(\tilde{Z}_s^{\prime}\hat{\mathbb{S}}_T\tilde{Z}_r\right)^2\left(\tilde{U}_r^{\prime}\tilde{U}_s - \varepsilon_r^{\dagger\prime}\varepsilon_s^{\dagger}\right)^2 \\
&\quad + 8T^{-2}h_2^{-1}\sum_{s=2}^{T}\sum_{r=1}^{s-1}\bar{k}_{sr}^2\left(\tilde{Z}_s^{\prime}\hat{\mathbb{S}}_T\tilde{Z}_r\right)^2\left(\tilde{U}_r^{\prime}\tilde{U}_s - \varepsilon_r^{\dagger\prime}\varepsilon_s^{\dagger}\right)\varepsilon_r^{\dagger\prime}\varepsilon_s^{\dagger} \\
&\quad + 4T^{-2}h_2^{-1}\sum_{s=2}^{T}\sum_{r=1}^{s-1}\bar{k}_{sr}^2\left[\left(\tilde{Z}_s^{\prime}\hat{\mathbb{S}}_T\tilde{Z}_r\right)^2 - \left(Z_s^{0\prime}\mathbb{S}_T Z_r^0\right)^2\right]\left(\varepsilon_r^{\dagger\prime}\varepsilon_s^{\dagger}\right)^2 \equiv 4\mathbb{V}_{3T,1}^{(1)} + 8\mathbb{V}_{3T,1}^{(2)} + 4\mathbb{V}_{3T,1}^{(3)}.
\end{aligned}$$

Using the decomposition given by (S2.18) and following the analysis in the proof of Proposition S1.4, we can readily show that $\mathbb{V}_{3T,1}^{(\ell)} = o_P(1)$ for $\ell = 1, 2$. For $\mathbb{V}_{3T,1}^{(3)}$, we use the fact that $a^2 - b^2 = (a-b)^2 + 2(a-b)b$ to make further decomposition:

$$\begin{aligned}
\mathbb{V}_{3T,1}^{(3)} &= T^{-2}h_2^{-1}\sum_{s=2}^{T}\sum_{r=1}^{s-1}\bar{k}_{sr}^2\left(\tilde{Z}_s^{\prime}\hat{\mathbb{S}}_T\tilde{Z}_r - Z_s^{0\prime}\mathbb{S}_T Z_r^0\right)^2\left(\varepsilon_r^{\dagger\prime}\varepsilon_s^{\dagger}\right)^2 \\
&\quad + 2T^{-2}h_2^{-1}\sum_{s=2}^{T}\sum_{r=1}^{s-1}\bar{k}_{sr}^2\left(\tilde{Z}_s^{\prime}\hat{\mathbb{S}}_T\tilde{Z}_r - Z_s^{0\prime}\mathbb{S}_T Z_r^0\right)Z_s^{0\prime}\mathbb{S}_T Z_r^0\left(\varepsilon_r^{\dagger\prime}\varepsilon_s^{\dagger}\right)^2 \equiv \mathbb{V}_{3T,1}^{(3,1)} + 2\mathbb{V}_{3T,1}^{(3,2)}.
\end{aligned}$$

To study $\mathbb{V}_{3T,1}^{(3,1)}$, we make the use of the following decomposition:

$$\begin{aligned}
&\tilde{Z}_s^{\prime}\hat{\mathbb{S}}_T\tilde{Z}_r - Z_s^{0\prime}\mathbb{S}_T Z_r^0 \\
&= \tilde{Z}_s^{\prime}\hat{\mathbb{S}}_T\tilde{Z}_r - Z_s^{0\prime}(\mathbb{I}_p\otimes D_{Q^{-1}})\frac{1}{T}\sum_{t=1}^{T}S_{Tt,1}^{-1}S_{Tt,1}^{-1}(\mathbb{I}_p\otimes D_{Q^{-1}}^{\prime})Z_r^0 \\
&= \tilde{Z}_s^{\prime}\hat{\mathbb{S}}_T\left[\tilde{Z}_r - (\mathbb{I}_p\otimes D_{Q^{-1}}^{\prime})Z_r^0\right] + \tilde{Z}_s^{\prime}\left[\hat{\mathbb{S}}_T - \frac{1}{T}\sum_{t=1}^{T}S_{Tt,1}^{-1}S_{Tt,1}^{-1}\right](\mathbb{I}_p\otimes D_{Q^{-1}}^{\prime})Z_r^0
\end{aligned}$$

$$
\begin{aligned}
&+\left[\tilde{Z}_s-(\mathbb{I}_p\otimes D'_{Q^{-1}})Z_s^0\right]\frac{1}{T}\sum_{t=1}^{T}S_{Tt,1}^{-1}S_{Tt,1}^{-1}(\mathbb{I}_p\otimes D'_{Q^{-1}})Z_r^0\\
\equiv\quad & I_1+I_2+I_3.
\end{aligned}
$$

Note that

$$
\mathbb{V}_{3T,1}^{(3,1)}\leq 3\sum_{\ell=1}^{3}T^{-2}h_2^{-1}\sum_{s=2}^{T}\sum_{r=1}^{s-1}\bar{k}_{sr}^2 I_\ell^2\left(\varepsilon_r^{\dagger\prime}\varepsilon_s^{\dagger}\right)^2\equiv 3\sum_{\ell=1}^{3}\mathbb{V}_{3T,1}^{(3,1)}(\ell).
$$

Noting that $\frac{1}{T}\sum_{r=1}^{T}\left\|\tilde{F}_r-Q'^{-1}F_r\right\|^2\leq\frac{2}{T}\sum_{r=1}^{T}\left\|\tilde{F}_r-H'F_r\right\|^2+\frac{1}{T}\sum_{r=1}^{T}\left\|(H-Q^{-1})'F_r\right\|^2=O_P\left(C_{0NT}^{-2}\right)+O_P\left(C_{0NT}^{-2}\right)$, we have

$$
\begin{aligned}
\mathbb{V}_{3T,1}^{(3,1)}(1) &= T^{-2}h_2^{-1}\sum_{s=2}^{T}\sum_{r=1}^{s-1}\bar{k}_{sr}^2\left(\tilde{Z}_s'\hat{\mathbb{S}}_T\left[\tilde{Z}_r-(\mathbb{I}_p\otimes D'_{Q^{-1}})Z_r^0\right]\right)^2\left(\varepsilon_r^{\dagger\prime}\varepsilon_s^{\dagger}\right)^2\\
&\lesssim T^{-2}h_2^{-1}\sum_{j=1}^{p}\sum_{s=2}^{T}\sum_{r=1}^{s-1}\bar{k}_{sr}^2\left\|\tilde{Z}_s\right\|^2\left\|\tilde{F}_{r-j}-Q'^{-1}F_{r-j}\right\|^2\left(\varepsilon_r^{\dagger\prime}\varepsilon_s^{\dagger}\right)^2\\
&\leq \max_r\left\|\varepsilon_r^{\dagger}\right\|^2\frac{1}{T}\sum_{j=1}^{p}\sum_{r=j+1}^{T}\left\|\tilde{F}_{r-j}-Q'^{-1}F_{r-j}\right\|^2\frac{1}{Th_2}\sum_{s=2}^{T}\bar{k}_{sr}^2\left\|\tilde{Z}_s\right\|^2\left\|\varepsilon_s^{\dagger}\right\|^2\\
&= O_P\left(T^{2/(8+\sigma)}\right)O_P\left(C_{0NT}^{-2}\right)O_P(1)=o_P(1).
\end{aligned}
$$

Analogously, we can show that

$$
\begin{aligned}
\mathbb{V}_{3T,1}^{(3,1)}(3) &= T^{-2}h_2^{-1}\sum_{s=2}^{T}\sum_{r=1}^{s-1}\bar{k}_{sr}^2\left\{\left[\tilde{Z}_s-(\mathbb{I}_p\otimes D'_{Q^{-1}})Z_s^0\right]\frac{1}{T}\sum_{t=1}^{T}S_{Tt,1}^{-1}S_{Tt,1}^{-1}(\mathbb{I}_p\otimes D'_{Q^{-1}})Z_r^0\right\}^2\left(\varepsilon_r^{\dagger\prime}\varepsilon_s^{\dagger}\right)^2\\
&\lesssim \max_r\left\|\varepsilon_r^{\dagger}\right\|^2\frac{1}{T}\sum_{j=1}^{p}\sum_{s=j+1}^{T}\left\|\tilde{F}_{s-j}-Q'^{-1}F_{s-j}\right\|^2=O_P\left(T^{2/(8+\sigma)}\right)O_P\left(C_{0NT}^{-2}\right)=o_P(1).
\end{aligned}
$$

In addition.

$$
\begin{aligned}
\mathbb{V}_{3T,1}^{(3,1)}(2) &= T^{-2}h_2^{-1}\sum_{s=2}^{T}\sum_{r=1}^{s-1}\bar{k}_{sr}^2\left\{\tilde{Z}_s'\left[\hat{\mathbb{S}}_T-\frac{1}{T}\sum_{t=1}^{T}S_{Tt,1}^{-1}S_{Tt,1}^{-1}\right](\mathbb{I}_p\otimes D'_{Q^{-1}})Z_r^0\right\}^2\left(\varepsilon_r^{\dagger\prime}\varepsilon_s^{\dagger}\right)^2\\
&\lesssim \left\|\hat{\mathbb{S}}_T-\frac{1}{T}\sum_{t=1}^{T}S_{Tt,1}^{-1}S_{Tt,1}^{-1}\right\|T^{-2}h_2^{-1}\sum_{s=2}^{T}\sum_{r=1}^{s-1}\bar{k}_{sr}^2\left\|\tilde{Z}_s\right\|^2\left\|Z_r^0\right\|^2\left\|\varepsilon_r^{\dagger}\right\|^2\left\|\varepsilon_s^{\dagger}\right\|^2\\
&= O_P((Th_2/\ln T)^{-1/2})=o_P(1).
\end{aligned}
$$

So $\mathbb{V}_{3T,1}^{(3,1)}=o_P(1)$. By the same token, $\mathbb{V}_{3T,1}^{(3,2)}=o_P(1)$. Then $\mathbb{V}_{3T,1}^{(3)}=o_P(1)$ and $\mathbb{V}_{3T,1}=o_P(1)$.

Next, we show (ii). Using $a^2 - b^2 = (a-b)^2 + 2(a-b)b$, we have

$$\begin{aligned}
\mathbb{V}_{3T,2} &= 4T^{-2}h_2^{-1}\sum_{s=2}^{T}\sum_{r=1}^{s-1}\bar{k}_{sr}^2\left(Z_s^{0\prime}\mathbb{S}_T Z_r^0\left(\varepsilon_r^{\dagger\prime}\varepsilon_s^{\dagger} - \varepsilon_r' D_{Q^{-1}}D_{Q^{-1}}'\varepsilon_s\right)\right)^2 \\
&\quad + 8T^{-2}h_2^{-1}\sum_{s=2}^{T}\sum_{r=1}^{s-1}\bar{k}_{sr}^2 Z_s^{0\prime}\mathbb{S}_T Z_r^0\left(\varepsilon_r^{\dagger\prime}\varepsilon_s^{\dagger} - \varepsilon_r' D_{Q^{-1}}D_{Q^{-1}}'\varepsilon_s\right)\varepsilon_r' D_{Q^{-1}}D_{Q^{-1}}'\varepsilon_s \\
&\equiv 4\mathbb{V}_{3T,2}^{(1)} + 8\mathbb{V}_{3T,2}^{(2)}.
\end{aligned}$$

Using $\varepsilon_r^{\dagger\prime}\varepsilon_s^{\dagger} - \varepsilon_r' D_{Q^{-1}}D_{Q^{-1}}'\varepsilon_s = \varepsilon_r'\left(D_H D_H' - D_{Q^{-1}}D_{Q^{-1}}'\right)\varepsilon_s$, we can readily show that

$$\begin{aligned}
\left|\mathbb{V}_{3T,2}^{(1)}\right| &= T^{-2}h_2^{-1}\sum_{s=2}^{T}\sum_{r=1}^{s-1}\bar{k}_{sr}^2\left(Z_s^{0\prime}\mathbb{S}_T Z_r^0\left(\varepsilon_r^{\dagger\prime}\varepsilon_s^{\dagger} - \varepsilon_r' D_{Q^{-1}}D_{Q^{-1}}'\varepsilon_s\right)\right)^2 \\
&\lesssim T^{-2}h_2^{-1}\sum_{s=2}^{T}\sum_{r=1}^{s-1}\bar{k}_{sr}^2\left(Z_s^{0\prime}\mathbb{S}_T Z_r^0\right)^2\|\varepsilon_r\|^2\|\varepsilon_s\|^2\left\|D_H D_H' - D_{Q^{-1}}D_{Q^{-1}}'\right\|^2 \\
&\lesssim \left\|H - Q^{-1}\right\|^2 = o_P(1)
\end{aligned}$$

and similarly $\left|\mathbb{V}_{3T,2}^{(2)}\right| \lesssim \|H - Q^{-1}\| = o_P(1)$. It follows that $\mathbb{V}_{3T,2} = o_P(1)$.

Now, we show (iii). Noting that $E[\mathbb{V}_{3T,3}] = 0$ and $4T^{-2}h_2^{-1}\sum_{s=2}^{T}\sum_{r=1}^{s-1}\bar{k}_{sr}^2(Z_s^{0\prime}\mathbb{S}_T Z_r^0\varepsilon_r' D_{Q^{-1}}D_{Q^{-1}}'\varepsilon_s)^2$ is the leading term in the expansion of $\sum_{s=1}^{T}\mathcal{Z}_s^2$, $\mathrm{Var}(\mathbb{V}_{3T,3}) = o(1)$ by the proof of part 2 in (S2.16). Then $\mathbb{V}_{3T,3} = o_P(1)$ by Chebyshev inequality.

In sum, we have shown that $\hat{\mathbb{V}}_{3T} = \mathbb{V}_{3T} + o_P(1)$. ■

S3 Outline of the Study of $\widehat{SM}_{3S}$ in Section 4

Recall that $L_1 = \begin{pmatrix} \mathbb{I}_K & 0_{K\times R} \end{pmatrix}$, $L = \begin{pmatrix} \mathbb{I}_K & 0_{K\times R} \\ 0_{R\times K} & 0_{R\times R} \end{pmatrix}$, $\mathbb{L}_1 = \mathbb{I}_p \otimes L_1$ and $\mathbb{L} = \mathbb{I}_p \otimes L$. Note that $L_1' L_1 = L$, $\mathbb{L}_1'\mathbb{L}_1 = \mathbb{L}$, $L^2 = L$ and $\mathbb{L}^2 = \mathbb{L}$. Then

$$\begin{aligned}
Th_2^{1/2}M_{3S} &= h_2^{1/2}\sum_{t=1}^{T}\left\|\mathbb{L}_1(\breve{\Psi}_t - \tilde{\Psi}_0)L_1'\right\|^2 \\
&= h_2^{1/2}\sum_{t=1}^{T}\mathrm{tr}\left((\breve{\Psi}_t - \tilde{\Psi}_0)L(\breve{\Psi}_t - \tilde{\Psi}_0)'\mathbb{L}\right)
\end{aligned}$$

$$= h_2^{1/2}\sum_{t=1}^{T}\text{tr}\left((\breve{\Psi}_t-\Psi_0)L(\breve{\Psi}_t-\Psi_0)'\mathbb{L}\right)+h_2^{1/2}\sum_{t=1}^{T}\text{tr}\left((\tilde{\Psi}_0-\Psi_0)L(\tilde{\Psi}_0-\Psi_0)'\mathbb{L}\right)$$
$$-2h_2^{1/2}\sum_{t=1}^{T}\text{tr}\left((\breve{\Psi}_t-\Psi_0)L(\tilde{\Psi}_0-\Psi_0)'\mathbb{L}\right)$$
$$\equiv M_{31S}+M_{32S}-2M_{33S}.$$

Following the proofs of Propositions S1.1–S1.3, we can readily show that

(i) $M_{31S}-\mathbb{B}^{(1)}_{3T,S}-\Pi_{3T,S}\xrightarrow{d}N(0,\mathbb{V}_{3T})$, (ii) $M_{32S}=\mathbb{B}^{(2)}_{3T,S}+o_P(1)$, and (iii) $M_{33S}=\mathbb{B}^{(1)}_{3T,S}+o_P(1)$,

where

$$\mathbb{B}^{(1)}_{3T,S}=\frac{h_2^{1/2}}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}k^{\dagger 2}_{h_2,st}Z_s^{\dagger\prime}\tilde{S}^{-1}_{ZZ,t}\mathbb{L}\tilde{S}^{-1}_{ZZ,t}Z_s^{\dagger}\varepsilon_s^{\dagger\prime}L\varepsilon_s^{\dagger},$$
$$\mathbb{B}^{(2)}_{3T,S}=\frac{h_2^{1/2}}{T}\sum_{s=1}^{T}Z_s^{\dagger\prime}\tilde{S}^{-1}_{ZZ}\mathbb{L}\tilde{S}^{-1}_{ZZ}Z_s^{\dagger}\varepsilon_s^{\dagger\prime}L\varepsilon_s^{\dagger},$$
$$\mathbb{B}^{(3)}_{3T,S}=\frac{h_2^{1/2}}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}k^{\dagger}_{h_2,st}Z_s^{\dagger\prime}\tilde{S}^{-1}_{ZZ,t}\mathbb{L}\tilde{S}^{-1}_{ZZ}Z_s^{\dagger}\varepsilon_s^{\dagger\prime}L\varepsilon_s^{\dagger},$$
$$\Pi_{3T,S}=\left[\text{vec}\left(D_{Q^{-1}}LD'_{Q^{-1}}\right)\right]'\left[\int_0^1 g_2(u)'\otimes g_2(u)'du\right]\text{vec}\left(\mathbb{I}_p\otimes D'^{-1}_{Q^{-1}}LD^{-1}_{Q^{-1}}\right),$$
$$\mathbb{V}_{3T,S}=\frac{4}{T^2h_2}\sum_{s=2}^{T}\sum_{r=1}^{s-1}\bar{K}\left(\frac{s-r}{Th_2}\right)^2E\left\{\left[Z_s^{0\prime}\mathbb{S}_{T,S}Z_r^0\varepsilon_r'D_{Q^{-1}}LD'_{Q^{-1}}\varepsilon_s\right]^2\right\},\tag{S3.1}$$

with $\mathbb{S}_{T,S}=\mathbb{S}_T=(\mathbb{I}_p\otimes D_{Q^{-1}})\frac{1}{T}\sum_{t=1}^{T}S^{-1}_{ZZ,t,1}\mathbb{L}S^{-1}_{ZZ,t,1}(\mathbb{I}_p\otimes D'_{Q^{-1}})$. Note that

$$\begin{aligned}
\mathbb{B}_{3T,S} &\equiv \mathbb{B}^{(1)}_{3T,S}+\mathbb{B}^{(2)}_{3T,S}-2\mathbb{B}^{(3)}_{3T,S}\\
&= \frac{h_2^{1/2}}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}Z_s^{\dagger\prime}\left(\tilde{S}^{-1}_{ZZ,t}k^{\dagger}_{h_2,st}-\tilde{S}^{-1}_{ZZ}\right)\mathbb{L}\left(\tilde{S}^{-1}_{ZZ,t}k^{\dagger}_{h_2,st}-\tilde{S}^{-1}_{ZZ}\right)Z_s^{\dagger}\varepsilon_s^{\dagger\prime}L\varepsilon_s^{\dagger},\\
\hat{\mathbb{B}}_{3T,S} &= \frac{h_2^{1/2}}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\tilde{Z}_s'\left(k^{\dagger}_{h_2,st}\tilde{S}^{-1}_{ZZ,t}-\tilde{S}^{-1}_{ZZ}\right)\mathbb{L}\left(k^{\dagger}_{h_2,st}\tilde{S}^{-1}_{ZZ,t}-\tilde{S}^{-1}_{ZZ}\right)\tilde{Z}_s\tilde{U}_s'L\tilde{U}_s,\\
\hat{\mathbb{V}}_{3T,S} &= \frac{4}{T^2h_2}\sum_{s=2}^{T}\sum_{r=1}^{s-1}\bar{K}\left(\frac{s-r}{Th_2}\right)^2\left[\tilde{Z}_s'\hat{\mathbb{S}}_{T,S}\tilde{Z}_r\tilde{U}_r'L\tilde{U}_s\right]^2,
\end{aligned}$$

Table S1: The IC-based determination of the lag order

DGP\(N, T)	RMSE				Percentage of correct selection			
	(100,100)	(100,200)	(200,100)	(200,200)	(100,100)	(100,200)	(200,100)	(200,200)
1	0.000	0.000	0.000	0.000	1.000	1.000	0.998	1.000
2	0.233	0.000	0.189	0.000	0.994	1.000	0.996	1.000
3	0.195	0.000	0.232	0.000	0.994	1.000	0.994	1.000
4	0.000	0.000	0.044	0.000	1.000	1.000	0.998	1.000
5	0.000	0.000	0.000	0.000	1.000	1.000	0.998	1.000
6	0.189	0.000	0.189	0.000	0.996	1.000	0.996	1.000
7	0.134	0.000	0.000	0.000	0.998	1.000	1.000	1.000
8	0.219	0.000	0.134	0.000	0.994	1.000	0.998	1.000

Note: Numbers in the main entries are the results based on 1000 replications.

where $\hat{\mathbb{S}}_{T,S} = \frac{1}{T}\sum_{t=1}^{T} \tilde{S}_{ZZ,t}^{-1}\mathbb{L}\tilde{S}_{ZZ,t}^{-1}$. One can readily follow the proofs of Propositions S1.4–S1.5 and show that $\hat{\mathbb{B}}_{3T,S} - \mathbb{B}_{3T,S} = o_P(1)$ and $\hat{\mathbb{V}}_{3T,S} - \mathbb{V}_{3T,S} = o_P(1)$.

S4 Some Additional Simulation Results

In this section we report some additional results in Sections 5 and 6.

S4.1 Determination of the Lag Order

In this subsection, we assess the performance of the IC in determining the lag order p. In particular, we consider the BIC-type information criterion given by (3.1) with $\rho_T = \log(T)/T$.

We use two measures to evaluate the information criterion, i.e., the root-mean-squared error (RMSE) and the empirical probability of correct selection over 1000 replications. Table S1 reports the results of IC for various DGPs with i.i.d. error terms. The results for the heteroskedastic error terms and the cross-sectional dependent error terms are quite similar. As shown in Table S1, our IC works fairly well for all the DGPs under investigation. The RMSE tends to zero, and the empirical probability of correct selection tends to 1 as the sample size T increases.

S4.2 Simulation Results for the Performance of Estimation in Term of SSR

Besides the reported values of $RMSE$ and $RMSE_\sigma$, we also evaluate the considered estimators using the sum of squared residuals (SSR) defined as the following: $SSR = \frac{1}{M(T-p)}\sum_{l=1}^{M}\sum_{t=p+1}^{T}\left(\hat{\varepsilon}_{1t}^{(l)}\right)^2$,

Table S2: Performance on the estimation in terms of SSR and $RMSE_\sigma$

DGP\(N,T)	Bai and Ng (2006) (100,100)	(100,200)	(200,100)	(200,200)	This paper (100,100)	(100,200)	(200,100)	(200,200)
	SSR							
1	**0.9643**	**0.9874**	**0.9501**	**0.9941**	0.8438	0.9169	0.8310	0.9260
2	1.5686	1.6463	1.5633	1.6370	**0.9135**	**0.9754**	**0.9051**	**0.9826**
3	1.6095	1.6641	1.6068	1.6795	**0.9508**	**0.9987**	**0.9432**	**1.0061**
4	1.5882	1.6580	1.6195	1.6534	**0.9179**	**0.9931**	**0.9263**	**0.9875**
5	1.6414	1.6709	1.6265	1.6541	**0.9674**	**1.0215**	**0.9730**	**1.0193**
6	1.3445	1.4221	1.3786	1.4114	**0.9493**	**1.0112**	**0.9604**	**1.0062**
7	1.9118	1.9699	1.8993	1.9554	**1.2018**	**1.2265**	**1.2118**	**1.2289**
8	1.8331	1.8458	1.7475	1.8423	**1.1658**	**1.1630**	**1.1262**	**1.1754**

Note: The main entries report the results based on 1000 replications. The bold entries highlight the better performance in each case.

where M is the number of replications (which is 1000 here) and $\{\hat{\varepsilon}_{1t}^{(l)}\}_{t=1,\ldots,T,l=1,\ldots,M}$ is the residual corresponding to the error term ε_{1t} at the l-th replication. We define the SSR as the average of the squared residuals, which can be regarded as an estimator for the variance of the error term. Recall that $\varepsilon_{1t} \sim i.i.d.N(0,1)$. Hence, the closer the SSR is to 1, the better the estimation result is.

Table S2 reports the SSRs for the estimates of Bai and Ng (2006) and ours with i.i.d. error terms based on 1000 replications. As mentioned above, since SSR is an estimator for the variance of the error terms, the closer the SSR is to 1, the better the estimator is. Clearly, Bai and Ng's (2006) estimator outperforms our estimator under DGP 1 in terms of SSR, which describes a time-invariant FAVAR model. However, it does not perform as well as our estimator under the other DGPs. This conclusion is also consistent with the results reported in the paper based on the other criteria.

S4.3 Simulation Results for the Stochastic Evolutions and Recurrent Changes

Since our time-varying (TV) factor loadings and VAR coefficients are specified as a deterministic function of t/T, we do not consider the stochastic specification in our theoretical and simulation studies. In our simulation studies, we have considered the cases of the single structural break, the multiple structural breaks, and the smooth structural changes. Here we follow the suggestion of a referee and consider the following additional setups for the factor loadings $\lambda_{it} = (\lambda_{it,1}, \lambda_{it,2})'$ and the VAR coefficients ϕ_t:

Table S3: Empirical rejection rates (i.i.d. errors)

			$\widehat{SM}_1$		$\widehat{SM}_2$		$\widehat{SM}_3$		$\widehat{SM}_{3S}$	
DGP	N	T	5%	10%	5%	10%	5%	10%	5%	10%
A1	100	100	0.384	0.510	0.986	0.996	0.094	0.152	0.250	0.334
	100	200	0.674	0.800	0.996	1.000	0.192	0.272	0.682	0.808
	200	100	0.380	0.520	0.984	0.998	0.108	0.152	0.250	0.340
	200	200	0.664	0.822	1.000	1.000	0.234	0.318	0.718	0.782
A2	100	100	0.330	0.452	0.694	0.758	0.046	0.086	0.236	0.390
	100	200	0.870	0.932	0.964	0.976	0.110	0.198	0.546	0.678
	200	100	0.412	0.524	0.804	0.826	0.042	0.100	0.248	0.362
	200	200	0.958	0.980	0.986	0.990	0.106	0.180	0.578	0.722

Note: The main entries report the empirical rejection rates under each GDP based on 500 iterations.

- DGP.A1: (truncated random walk evolution in both factor loadings and VAR coefficients)

  $\tilde{\lambda}_{it} = \tilde{\lambda}_{i,t-1} + v_{it}$, $\lambda_{it} = \text{sign}(\tilde{\lambda}_{it}) \min\{|\tilde{\lambda}_{it}|, 1\}$;

  $\phi_t = (\phi_t^{(1,1)}, \phi_t^{(1,2)}, \phi_t^{(1,3)}; 0, 0.6, 0; 0, 0, 0.3)$, with $\tilde{\phi}_t^{(1,j)} = \tilde{\phi}_{t-1}^{(1,j)} + u_{j,t}$ and $\phi_t^{(1,j)} = \text{sign}(\tilde{\phi}_t^{(1,j)}) \cdot \min\{|\tilde{\phi}_t^{(1,j)}|, 1\}$, where $v_{it} \sim i.i.d.N(0, 0.1^2)$ and $u_{j,t} \sim i.i.d.N(0, 0.1^2)$.

- DGP.A2: (recurrent changes)

  $\lambda_{it,1} = \mu_i + \sin(4\pi t/T)$, $\mu_i \sim i.i.d.U(0,1)$, $\lambda_{it,2} = \lambda_{i0,2} \sim i.i.d.N(0,1)$, where $\mu_i \sim i.i.d.U(0,1)$;

  $\phi_t = (\phi_t^{(1,1)}, \phi_t^{(1,2)}, \phi_t^{(1,3)}; 0, 0.6, 0; 0, 0, 0.3)$, with $\phi_t^{(1,1)} = \sin(6\pi t/T)$, $\phi_t^{(1,2)} = \sin(4\pi t/T)$ and $\phi_t^{(1,3)} = \cos(4\pi t/T)$.

DGP.A1 is the FAVAR model with the factor loadings and the VAR coefficients being a truncated version of the unit root processes. Since an AR(1) process with the autoregressive coefficient larger than 1 will result in an explosive process, which rarely exists in economic applications, we truncate the unit root process to guarantee that the absolute value of the auto-coefficient is less than or equal to 1. Such a treatment is necessary for Assumption A.3(ii) to hold. DGP.A2 is the FAVAR model with recurrent changing factor loadings and VAR coefficients.

Table S3 reports the empirical rejection rates of our tests at both 5% and 10% significance levels with i.i.d. error terms, using the bootstrap critical values. The smoothing parameters, kernel functions, and other pre-settings for the proposed tests are the same as in Section 5 of the paper. The results reported in Table S3 show that the empirical rejection rates of $\widehat{SM}_1$ and

$\widehat{SM}_2$ increase quickly with the sample sizes (N, T). The empirical rejection rates of $\widehat{SM}_3$ are slightly low. As mentioned in the paper, the $\widehat{SM}_3$ test is only valid when the null hypothesis $\mathbb{H}_0^{(2)}$ holds. Since these two DGPs show time-varying factor loadings, the $\widehat{SM}_3$ test is not applicable. However, our $\widehat{SM}_{3S}$ test is still available even when $\mathbb{H}_0^{(2)}$ is violated. We note that the empirical rejection rates of $\widehat{SM}_{3S}$ increase as the sample size T grows. Therefore, although our tests are designed for detecting deterministic structural changes, they can still capture the stochastic TV features such as the truncated random walk process given by DGP.A1. In addition, our tests are also powerful to capture the recurrent changes given by DGP.A2, which is actually a special case of smooth structural changes.

S4.4 Comparison of Our Estimations and Predictions with Those Based on the Kalman Filter

We note that if the factor loadings are TV with random walk evolutions, the estimation procedure is quite involved. Although existing literature provides some approaches to estimate the TV factor model with random walk evolutions, the asymptotic theory is not formally established. For example, Eickmeier et al. (2015) suggest ignoring the TV behavior of the factor loadings and estimate the common factors by conventional PCA in the first step, and then treat the factors as observable and estimate the TV factor loadings using the Kalman-filter-based maximum likelihood estimation (MLE). However, they do not provide any asymptotic theory for their estimators. Mikkelsen et al. (2019) establish some asymptotic results for the TV factor model with *stationary* stochastic evolutions, but they are not applicable to the case of random walk evolutions. Since the estimation of the TV factor model with random walk evolution is far beyond the scope of this paper, we assume the factor loadings in the FAVAR model to be time invariant and only examine the behavior of various estimators for the TV VAR coefficients.

Specifically, we consider the FAVAR(1) model given in Section 5.1 of our revised paper. The time-invariant factor loadings are generated by $\lambda_{i0} \sim i.i.d.N(0, \mathbb{I}_2)$. We consider the following setups for the VAR coefficients ϕ_t.

DGP.B1: (time-invariant VAR coefficients)

$\phi_t = \phi_0 = (0.5, 0.4, 0.3; 0, 0.6, 0; 0, 0, 0.3)$;

DGP.B2: (TV VAR coefficients with smooth structural changes)

$\phi_t = (\phi_t^{(1,1)}, \phi_t^{(1,2)}, \phi_t^{(1,3)}; 0, 0.6, 0; 0, 0, 0.3)$ with $\phi_t^{(1,1)} = -0.2 + \Xi(10t/T; 1, 5)$, $\phi_t^{(1,2)} = -0.2 + \Xi(10t/T; 0.3, (4, 8)')$, and $\phi_t^{(1,3)} = 1 - \Xi(10t/T; 0.1, (2, 4, 8)')$.

DGP.B3: (TV VAR coefficients with an abrupt break)

$\phi_t = (\phi_t^{(1,1)}, \phi_t^{(1,2)}, \phi_t^{(1,3)}; 0, 0.6, 0; 0, 0, 0.3)$, where $\phi_t^{(1,j)} = \begin{cases} -0.5 + 0.5\phi_0^{(1,j)}, & \text{for } t \leq T/2 \\ 0.3 + 0.5\phi_0^{(1,j)}, & \text{for } t \geq T/2+1 \end{cases}$ with $\phi_0^{(1,j)} \sim i.i.d.U(0,1)$ for $j = 1,2,3$.

DGP.B4: (TV VAR coefficients following a truncated random walk process with small innovation variance)

$\phi_t = \left(\phi_t^{(1,1)}, \phi_t^{(1,2)}, \phi_t^{(1,3)}; 0, 0.6, 0; 0, 0, 0.3\right)$, with $\tilde{\phi}_t^{(1,j)} = \tilde{\phi}_{t-1}^{(1,j)} + u_{j,t}$, and $\phi_t^{(1,j)} = \text{sign}(\tilde{\phi}_t^{(1,j)})$ $\min\{|\tilde{\phi}_t^{(1,j)}|, 1\}$, where $u_{j,t} \sim i.i.d.N(0, 0.1^2)$.

DGP.B5: (TV VAR coefficients following a truncated random walk process with large innovation variance)

$\phi_t = \left(\phi_t^{(1,1)}, \phi_t^{(1,2)}, \phi_t^{(1,3)}; 0, 0.6, 0; 0, 0, 0.3\right)$, with $\tilde{\phi}_t^{(1,j)} = \tilde{\phi}_{t-1}^{(1,j)} + u_{j,t}$, and $\phi_t^{(1,j)} = \text{sign}(\tilde{\phi}_t^{(1,j)})$ $\min\{|\tilde{\phi}_t^{(1,j)}|, 1\}$, where $u_{j,t} \sim i.i.d.N(0, 1)$.

These DGPs describe various TV patterns in the VAR coefficients. DGPs B.1 to B.3 are the same as DGPs 1 to 3 in our paper, which are FAVAR models with time-invariant factor loadings and various types of VAR coefficients, including the time-invariant VAR coefficients, VAR coefficients with smooth changes, and VAR coefficients with an abrupt structural break, respectively. DGPs B.4 and B.5 specify the TV VAR coefficients as the truncated versions of the random walk with a small and large innovation variance, respectively. Since an AR(1) process with an autoregressive coefficient larger than 1 will result in an explosive process, which rarely exists in economic applications, we truncate the random walk to guarantee that the absolute value of the auto coefficient is less than or equal to 1. For each DGP, we compare our estimation and forecasting results with those estimated by Bai and Ng (2006) and the Kalman-filter-based MLE, respectively. Specifically, for the Kalman-filter-based MLE, we estimate the common factors by the conventional PCA in the first stage and then estimate the VAR coefficients by the Kalman filter in the second stage.

Table S4.4 reports the $RMSE$, SSR, and $RMSE_\sigma$ for the estimators of Bai and Ng (2006), the Kalman filter-based MLE and ours with i.i.d. error term based on 1000 replications. As mentioned above, since SSR is an estimator for the variance of the error terms, the closer the SSR is to 1, the better the estimator is. While for the $RMSE$ and $RMSE_\sigma$, we prefer to the estimator that achieves the smallest value. As shown in the table, the results with different criteria are generally consistent. First, Bai and Ng's (2006) estimator outperforms the other two estimators under DGP.B1, which describes a time-invariant FAVAR model. However, it is not as good as our estimator and the Kalman-filter-based MLE estimator under the other DGPs. Second, our local smoothing estimator achieves the best result for DGPs B2 and B3, which have smooth structural change and an abrupt structural break in the VAR coefficients. Third, the results for the truncated random walk process given by DGPs B4 and B5 are mixed. Under

Table S4: Performance of the estimation in terms of $RMSE$, SSR and $RMSE_{\sigma}$

DGP	(N, T)	$RMSE$			SSR			$RMSE_{\sigma}$		
		BN	LS	KF	BN	LS	KF	BN	LS	KF
B1	(100,100)	**0.0750**	0.1878	0.2613	**0.9643**	0.8438	0.7230	**0.0166**	0.0391	0.1481
	(100,200)	**0.0568**	0.1416	0.1993	**0.9874**	0.9169	0.7615	**0.0088**	0.0146	0.1397
	(200,100)	**0.0819**	0.2038	0.2575	**0.9501**	0.8310	0.7546	**0.0232**	0.0466	0.1203
	(200,200)	**0.0544**	0.1300	0.2000	**0.9941**	0.9260	0.7390	**0.0096**	0.0142	0.1571
B2	(100,100)	0.4331	**0.1812**	0.2730	1.5109	**0.9259**	0.7596	0.3266	**0.0262**	0.0879
	(100,200)	0.4419	**0.1353**	0.2055	1.6023	**0.9947**	0.8040	0.4129	**0.0133**	0.0509
	(200,100)	0.4367	**0.1862**	0.2842	1.4992	**0.9272**	0.7461	0.3223	**0.0291**	0.1050
	(200,200)	0.4399	**0.1245**	0.2134	1.6012	**0.9928**	0.8022	0.3987	**0.0104**	0.0517
B3	(100,100)	0.4750	**0.2480**	0.3415	1.6095	**0.9508**	0.7587	0.4403	**0.0261**	0.0900
	(100,200)	0.4863	**0.1917**	0.2519	1.6641	**0.9987**	0.8187	0.4793	**0.0127**	0.0462
	(200,100)	0.4777	**0.2519**	0.3367	1.6068	**0.9432**	0.7465	0.4419	**0.0267**	0.0960
	(200,200)	0.4866	**0.1816**	0.2591	1.6795	**1.0061**	0.8180	0.5006	**0.0127**	0.0455
B4	(100,100)	0.3828	**0.2367**	0.3130	1.5877	**1.0338**	0.8151	0.5148	**0.0335**	0.0893
	(100,200)	0.4951	**0.2523**	0.3065	2.1950	1.2116	**0.8770**	2.3514	0.1221	**0.0912**
	(200,100)	0.3776	**0.2376**	0.3165	1.5805	**1.0534**	0.8236	0.5256	**0.0439**	0.0990
	(200,200)	0.4663	**0.2441**	0.2885	2.0625	1.1937	**0.8516**	1.8403	0.0836	**0.0668**
B5	(100,100)	0.8096	0.7569	**0.6829**	6.2630	4.3413	**1.7030**	33.6340	13.7550	**1.7165**
	(100,200)	0.8119	0.7806	**0.6462**	7.0717	5.3991	**1.5578**	41.2870	21.8970	**0.8162**
	(200,100)	0.8084	0.7589	**0.6734**	6.2642	4.3025	**1.4977**	33.3180	13.5310	**0.9067**
	(200,200)	0.8141	0.7758	**0.6425**	7.1050	5.4531	**1.5088**	41.7660	22.3070	**0.7546**

Notes: (i) BN, LS, and KF denote Bai and Ng's (2006) estimation, the local smoothing estimation proposed in this paper, and the Kalman-filter-based MLE, respectively; (ii) The main entries report the results based on 1000 replications; (iii) The bold entries highlight the best performance in each case.

DGP.B4, the proposed local smoothing estimator has the smallest RMSE, which shows that our estimator for the $(1, 1)$st element of the VAR coefficients is the most accurate. However, suppose we use SSR and MSE_{σ}, which evaluate the explanatory power of an estimation method. In those cases, our estimator outperforms the Kalman-filter-based MLE estimator when the sample size T is small but underperforms when T is large. In contrast, under DGP.B5, which specifies the truncated random walk TV VAR coefficients with a larger innovation variance, the Kalman-filter-based MLE performs the best.

Table S5 reports the MSFE under DGPs B.1 to B.5 with 1000 replications. As shown in the table, Bai and Ng's (2006) approach performs the best under DGP.B1, our approach performs the best under DGP.B2. The Kalman-filter-based MLE outperforms the other two methods under DGP.B3 with abrupt structural changes. In addition, we note that our approach achieves the smallest MSFE under DGP.B4. While under DGP.B5, which specifies a truncated random walk process with large variance, Bai and Ng's (2006) approach performs the best. It implies that when the VAR coefficients evolve as random walk processes with a large volatility, the TV methods, such as our local smoothing approach and the Kalman-filter-based MLE, may not provide more accurate predictions than that of Bai and Ng (2006).

S4.5 Additional Simulation Results for the Proposed Tests: Different Cases of Error Terms

In this subsection, we report some additional simulation results for the proposed tests in Section 4 when the error terms are heteroskedastic, spatially correlated, or exhibiting stochastic volatilities.

Tables S6 and S7 report the testing results when the error terms are heteroskedastic and spatially correlated, respectively. We summarize some key findings. First, regarding the size performance, the results in these two tables are comparable with those in Table 3 in the main body of the paper. Second, the presence of heteroskedasticity and spatial correlation has some impact on the power performance of the tests for some DGPs. This is particularly true for DGPs 7 and 8. But this may be due to the fact that we do not control the signal-noise ratio in these DGPs when the errors are i.i.d., heteroskedastic, or spatially correlated to ensure that they are comparable. Third, overall speaking, our tests have reasonable size and power performance across the eight DGPs and the different settings for the error terms.

Tables S4.5 and S4.5 report the empirical rejection rates of our tests at both 5% and 10% significance levels when the error terms are given by SV1 and SV2. These results are qualitatively similar to those reported in Table 3, indicating that our tests are robust to the

Table S5: Performance of the prediction in terms of MSFE

DGP	N	T	BN	LS	KF
B1	100	100	**0.9829**	1.2357	2.0609
	100	200	**0.9996**	1.1416	2.0467
	200	100	**0.9964**	1.2397	2.0807
	200	200	**0.9877**	1.1423	2.0218
B2	100	100	2.3750	**1.4829**	1.9408
	100	200	2.4374	**1.2960**	1.4394
	200	100	2.4748	**1.5000**	2.0290
	200	200	2.4434	**1.3070**	1.4631
B3	100	100	1.6335	1.3201	**1.2776**
	100	200	1.6435	1.2007	**1.1647**
	200	100	1.6795	1.3224	**1.2724**
	200	200	1.6281	1.2118	**1.1344**
B4	100	100	2.2490	**1.7274**	2.3281
	100	200	2.7615	1.7034	**1.6957**
	200	100	2.1641	**1.7160**	2.1998
	200	200	3.0490	**1.7367**	1.7913
B5	100	100	**8.3186**	8.6803	8.8970
	100	200	**7.5919**	7.9289	7.7588
	200	100	**7.1349**	7.7366	8.1627
	200	200	**7.3761**	7.5470	7.6327

Notes: (i) BN, LS, and KF denote Bai and Ng's (2006) estimation, the local smoothing estimation proposed in this paper, and the Kalman-filter-based MLE, respectively; (ii) The main entries report the results based on 1000 replications; (iii) The bold entries highlight the best performance in each case.

Table S6: Empirical rejection rates of the tests (heteroscedastic errors)

			$\widehat{SM}_1$		$\widehat{SM}_2$		$\widehat{SM}_3$		$\widehat{SM}_{3S}$	
DGP	N	T	5%	10%	5%	10%	5%	10%	5%	10%
1	100	100	0.058	0.120	0.062	0.126	0.042	0.100	0.060	0.120
	100	200	0.080	0.156	0.062	0.118	0.082	0.132	0.078	0.148
	200	100	0.070	0.158	0.082	0.122	0.056	0.132	0.056	0.124
	200	200	0.060	0.114	0.050	0.102	0.066	0.128	0.052	0.114
2	100	100	0.610	0.726	0.062	0.126	0.612	0.730	0.584	0.708
	100	200	0.984	0.994	0.062	0.118	0.988	0.998	0.986	0.992
	200	100	0.614	0.730	0.082	0.122	0.626	0.722	0.590	0.716
	200	200	0.966	0.990	0.050	0.102	0.980	0.994	0.984	0.994
3	100	100	0.522	0.670	0.068	0.140	0.536	0.690	0.506	0.618
	100	200	0.984	0.994	0.052	0.098	0.992	0.998	0.954	0.976
	200	100	0.554	0.678	0.072	0.118	0.582	0.712	0.524	0.644
	200	200	0.968	0.990	0.046	0.106	0.980	0.994	0.964	0.982
4	100	100	0.804	0.882	0.564	0.662	0.666	0.792	0.576	0.712
	100	200	1.000	1.000	0.872	0.902	0.988	0.992	0.984	0.994
	200	100	0.842	0.908	0.758	0.824	0.670	0.756	0.640	0.750
	200	200	0.996	1.000	0.960	0.972	0.982	0.990	0.978	0.992
5	100	100	0.806	0.898	0.624	0.716	0.668	0.802	0.542	0.654
	100	200	0.990	0.996	0.882	0.918	0.982	0.992	0.968	0.988
	200	100	0.804	0.886	0.714	0.786	0.638	0.758	0.530	0.668
	200	200	1.000	1.000	0.962	0.978	0.988	0.998	0.964	0.980
6	100	100	0.610	0.774	0.564	0.662	0.370	0.506	0.550	0.658
	100	200	0.982	0.992	0.872	0.902	0.894	0.954	0.972	0.986
	200	100	0.746	0.850	0.758	0.824	0.406	0.534	0.554	0.674
	200	200	0.994	0.998	0.960	0.972	0.886	0.944	0.968	0.988
7	100	100	0.506	0.644	0.624	0.716	0.232	0.342	0.092	0.202
	100	200	0.958	0.982	0.882	0.918	0.866	0.920	0.596	0.744
	200	100	0.576	0.722	0.714	0.786	0.210	0.366	0.142	0.210
	200	200	0.994	0.998	0.962	0.978	0.886	0.942	0.620	0.760
8	100	100	0.454	0.636	0.824	0.900	0.104	0.228	0.098	0.172
	100	200	0.990	0.996	0.994	0.996	0.766	0.892	0.554	0.702
	200	100	0.530	0.694	0.890	0.950	0.120	0.232	0.094	0.172
	200	200	0.996	1.000	1.000	1.000	0.784	0.882	0.578	0.742

Note: The main entries report the empirical rejection rate under each GDP based on 500 iterations.

Table S7: Empirical rejection rates of the tests (spatially correlated errors)

			$\widehat{SM}_1$		$\widehat{SM}_2$		$\widehat{SM}_3$		$\widehat{SM}_{3S}$	
DGP	N	T	5%	10%	5%	10%	5%	10%	5%	10%
1	100	100	0.046	0.100	0.058	0.096	0.058	0.096	0.048	0.096
	100	200	0.088	0.142	0.068	0.110	0.074	0.148	0.074	0.142
	200	100	0.064	0.124	0.060	0.108	0.058	0.112	0.054	0.114
	200	200	0.056	0.118	0.050	0.100	0.060	0.120	0.046	0.122
2	100	100	0.096	0.182	0.058	0.096	0.106	0.188	0.274	0.430
	100	200	0.336	0.484	0.068	0.110	0.368	0.536	0.816	0.904
	200	100	0.114	0.218	0.060	0.108	0.108	0.198	0.298	0.456
	200	200	0.392	0.518	0.050	0.100	0.410	0.558	0.842	0.918
3	100	100	0.120	0.230	0.058	0.096	0.128	0.236	0.286	0.396
	100	200	0.520	0.716	0.068	0.110	0.600	0.756	0.814	0.898
	200	100	0.160	0.270	0.060	0.108	0.154	0.274	0.262	0.426
	200	200	0.580	0.736	0.050	0.100	0.618	0.808	0.856	0.930
4	100	100	0.162	0.310	0.240	0.364	0.088	0.174	0.270	0.406
	100	200	0.748	0.866	0.680	0.772	0.418	0.584	0.816	0.892
	200	100	0.288	0.446	0.436	0.572	0.122	0.228	0.316	0.480
	200	200	0.906	0.966	0.936	0.972	0.442	0.608	0.828	0.912
5	100	100	0.222	0.374	0.240	0.364	0.134	0.248	0.250	0.392
	100	200	0.850	0.942	0.680	0.772	0.598	0.782	0.824	0.904
	200	100	0.324	0.490	0.436	0.572	0.178	0.298	0.284	0.428
	200	200	0.942	0.988	0.936	0.972	0.602	0.768	0.816	0.900
6	100	100	0.144	0.278	0.240	0.364	0.072	0.128	0.304	0.426
	100	200	0.640	0.792	0.680	0.772	0.292	0.438	0.758	0.840
	200	100	0.242	0.374	0.436	0.572	0.072	0.168	0.352	0.458
	200	200	0.854	0.938	0.936	0.972	0.304	0.446	0.776	0.876
7	100	100	0.112	0.188	0.240	0.364	0.066	0.110	0.034	0.096
	100	200	0.612	0.798	0.680	0.772	0.286	0.458	0.306	0.498
	200	100	0.130	0.272	0.436	0.572	0.050	0.116	0.050	0.090
	200	200	0.828	0.946	0.936	0.972	0.258	0.434	0.322	0.480
8	100	100	0.178	0.308	0.728	0.842	0.028	0.074	0.030	0.074
	100	200	0.882	0.964	0.994	1.000	0.128	0.246	0.280	0.424
	200	100	0.206	0.378	0.848	0.916	0.030	0.074	0.030	0.076
	200	200	0.920	0.974	1.000	1.000	0.118	0.254	0.264	0.432

Note: The main entries report the empirical rejection rate under each GDP based on 500 iterations.

stochastic volatility type error terms.

S4.6 Estimation and Prediction Based on the Test Results

We follow one referee's suggestion to select a suitable model based on the specification tests first and then estimate and predict using the selected model. Specifically, we consider the following three cases based on the results of our tests $\widehat{SM}_2$ and $\widehat{SM}_3$:

- If neither $\widehat{SM}_2$ nor $\widehat{SM}_3$ rejects the null hypotheses of no structural change in the factor loadings and the VAR coefficients, we use Bai and Ng's (2006) procedure to estimate the model and conduct the one-step-ahead prediction.
- If the $\widehat{SM}_2$ test fails to reject the null hypothesis of no structural change in the factor loadings, but the $\widehat{SM}_3$ test rejects the null hypothesis of no structural change in the VAR coefficients, we estimate the factor model using the conventional PCA procedure in the first stage and estimate the VAR coefficients using a local smoothing procedure in the second stage.
- If the $\widehat{SM}_2$ test rejects the null hypothesis of no structural change in factor loadings, we use the proposed two-step local smoothing approach to estimate the model and conduct the one-step-ahead prediction based on the estimation results. Please note that if the factor model suffers from structural changes, we should adopt the TV-FAVAR model no matter whether the VAR coefficients are TV or not.

Table S10 reports the $RMSE$, SSR and $RMSE_\sigma$ for the Bai and Ng's (2006) estimator (BN), the test-based estimator (TE), and the proposed local smoothing estimator (LS) with i.i.d. error terms based on 500 replications. First, as shown in the table, Bai and Ng's (2006) estimator outperforms the other two under DGP 1, which depicts a time-invariant FAVAR model. Second, according to $RMSE$ and $RMSE_\sigma$, the proposed local smoothing estimator performs the best among the three estimators in most cases under DGPs 2 to 8. Third, the results based on SSR are mixed under DGPs 4 to 6. Note that SSR, as an estimator for the error variance, may not be a good measure for the goodness of fit of the model. Recall that the true error variance in our model is 1, the estimated error variance may be greater or smaller than 1. Hence the bias terms may be cancelled with each other among M replications. Fourth, although the test based estimator does not achieve the best performance under any DGP, it performs similar to the best estimator for all the DGPs under our investigation. This indicates

Table S8: Empirical rejection rates of the proposed tests (SV1 errors)

			$\widehat{SM}_1$		$\widehat{SM}_2$		$\widehat{SM}_3$		$\widehat{SM}_{3S}$	
DGP	N	T	5%	10%	5%	10%	5%	10%	5%	10%
1	100	100	0.072	0.116	0.040	0.090	0.066	0.118	0.032	0.068
	100	200	0.076	0.142	0.054	0.106	0.080	0.134	0.024	0.046
	200	100	0.050	0.096	0.040	0.082	0.054	0.094	0.040	0.072
	200	200	0.054	0.106	0.042	0.102	0.056	0.098	0.022	0.040
2	100	100	0.286	0.396	0.062	0.104	0.282	0.398	0.424	0.556
	100	200	0.768	0.862	0.054	0.102	0.786	0.870	0.878	0.924
	200	100	0.260	0.410	0.062	0.122	0.262	0.394	0.398	0.536
	200	200	0.776	0.862	0.048	0.120	0.794	0.900	0.880	0.938
3	100	100	0.218	0.330	0.068	0.114	0.210	0.354	0.344	0.452
	100	200	0.696	0.840	0.042	0.088	0.722	0.854	0.802	0.870
	200	100	0.224	0.340	0.050	0.102	0.236	0.374	0.306	0.420
	200	200	0.766	0.860	0.058	0.096	0.786	0.888	0.778	0.866
4	100	100	0.622	0.768	0.880	0.938	0.222	0.366	0.370	0.534
	100	200	0.996	0.998	0.996	0.998	0.844	0.902	0.868	0.922
	200	100	0.730	0.824	0.940	0.970	0.250	0.414	0.430	0.552
	200	200	0.998	0.998	0.998	0.998	0.854	0.934	0.884	0.938
5	100	100	0.648	0.780	0.836	0.894	0.300	0.450	0.312	0.430
	100	200	0.996	0.998	0.996	0.998	0.888	0.940	0.786	0.866
	200	100	0.758	0.852	0.958	0.974	0.288	0.438	0.276	0.424
	200	200	0.998	1	0.998	1	0.878	0.952	0.774	0.864
6	100	100	0.552	0.702	0.870	0.910	0.138	0.236	0.440	0.536
	100	200	0.994	0.998	0.998	1	0.582	0.746	0.898	0.932
	200	100	0.648	0.776	0.942	0.966	0.164	0.256	0.380	0.506
	200	200	0.988	0.996	0.996	0.998	0.582	0.752	0.904	0.938
7	100	100	0.424	0.572	0.862	0.914	0.048	0.116	0.072	0.148
	100	200	0.986	0.996	0.998	1	0.502	0.646	0.490	0.640
	200	100	0.524	0.694	0.938	0.962	0.038	0.092	0.086	0.150
	200	200	0.980	0.998	0.998	0.998	0.476	0.644	0.478	0.600
8	100	100	0.202	0.318	0.644	0.788	0.054	0.106	0.084	0.142
	100	200	0.836	0.922	0.970	0.992	0.372	0.544	0.494	0.636
	200	100	0.200	0.334	0.692	0.832	0.038	0.086	0.076	0.126
	200	200	0.842	0.918	0.964	0.988	0.378	0.500	0.486	0.634

Note: The main entries report the empirical rejection rates under each DGP based on 500 iterations.

Table S9: Empirical rejection rates of the proposed tests (SV2 errors)

			$\widehat{SM}_1$		$\widehat{SM}_2$		$\widehat{SM}_3$		$\widehat{SM}_{3S}$	
DGP	N	T	5%	10%	5%	10%	5%	10%	5%	10%
1	100	100	0.084	0.134	0.052	0.116	0.078	0.128	0.034	0.074
	100	200	0.088	0.122	0.056	0.106	0.080	0.130	0.022	0.054
	200	100	0.044	0.094	0.060	0.112	0.046	0.080	0.046	0.078
	200	200	0.048	0.094	0.050	0.118	0.044	0.098	0.030	0.066
2	100	100	0.252	0.356	0.058	0.104	0.246	0.388	0.402	0.530
	100	200	0.788	0.878	0.064	0.118	0.818	0.902	0.886	0.938
	200	100	0.218	0.360	0.042	0.096	0.220	0.366	0.386	0.508
	200	200	0.770	0.874	0.040	0.090	0.796	0.898	0.888	0.938
3	100	100	0.214	0.322	0.054	0.094	0.222	0.330	0.340	0.426
	100	200	0.742	0.850	0.056	0.096	0.770	0.870	0.802	0.868
	200	100	0.238	0.338	0.058	0.114	0.222	0.360	0.340	0.446
	200	200	0.752	0.840	0.042	0.106	0.780	0.866	0.764	0.856
4	100	100	0.428	0.608	0.458	0.608	0.274	0.428	0.402	0.544
	100	200	0.980	0.996	0.942	0.958	0.858	0.928	0.882	0.946
	200	100	0.466	0.636	0.578	0.692	0.282	0.416	0.410	0.536
	200	200	0.982	0.994	0.996	1	0.830	0.900	0.846	0.918
5	100	100	0.426	0.594	0.434	0.544	0.286	0.414	0.320	0.466
	100	200	0.976	0.990	0.950	0.970	0.886	0.940	0.758	0.856
	200	100	0.516	0.676	0.628	0.736	0.290	0.436	0.308	0.436
	200	200	0.994	0.998	0.990	0.996	0.888	0.958	0.802	0.870
6	100	100	0.308	0.454	0.442	0.568	0.160	0.256	0.416	0.544
	100	200	0.926	0.970	0.948	0.956	0.610	0.756	0.904	0.934
	200	100	0.356	0.540	0.660	0.750	0.126	0.228	0.444	0.552
	200	200	0.976	0.990	0.990	0.994	0.608	0.734	0.916	0.958
7	100	100	0.174	0.312	0.418	0.554	0.060	0.144	0.066	0.138
	100	200	0.878	0.944	0.946	0.980	0.440	0.608	0.480	0.616
	200	100	0.208	0.352	0.608	0.718	0.050	0.114	0.098	0.154
	200	200	0.936	0.966	0.986	0.996	0.458	0.630	0.480	0.662
8	100	100	0.144	0.264	0.628	0.762	0.036	0.078	0.072	0.144
	100	200	0.806	0.894	0.970	0.984	0.372	0.520	0.496	0.652
	200	100	0.200	0.330	0.662	0.782	0.032	0.074	0.106	0.184
	200	200	0.826	0.902	0.960	0.980	0.338	0.520	0.512	0.668

Note: The main entries report the empirical rejection rates under each DGP based on 500 iterations.

that in practice one can first test the model specification using our tests and then estimate the model based on the test results.

Table S11 reports the MSFEs of the aforementioned three estimators under DGPs 1 to 8 with 500 replications. First, as shown in the table, the MSFEs using the proposed approach decline as the sample size T increases under all DGPs. Second, as expected, Bai and Ng's (2006) approach performs the best under DGP 1. Our predictions outperform those made by Bai and Ng's (2006) approach under other cases, especially under DGPs 2, 4 and 6, which allow for smooth structural changes in the factor loadings and VAR coefficients. Third, we note that the test-based estimator performs the best under DGPs 2 and 3, which exhibit time-invariant factor loadings and TV VAR coefficients. This implies that it is more appropriate to estimate the factor model using the conventional PCA in the first stage and then estimate the VAR coefficients by the local smoothing method in the second stage. Since our test $\widehat{SM}_2$ cannot reject the null hypothesis while our test $\widehat{SM}_3$ can reject the null hypothesis for most replications, the test-based estimator will be consistent with the estimator under the correct model specification. Thus, the test-based estimator performs the best under DGPs 2 and 3. Fourth, under DGPs 4 to 8 that are FAVAR models with time-varying factor loadings and time-varying VAR coefficients, The proposed local smoothing approach performs the best in most cases, especially when the sample size is large ($T = 200$). Last, we note that under each DGP, the test-based estimator's performance is close to the best one, indicating that based on the test result, one can obtain a stable and reasonable prediction result.

S5 Additional Empirical Results

S5.1 Determination of the Lag Order for the Seven Series

Table S12 reports the IC values for the seven series under investigation. According to the table, we can set the lag order to be 1 for the VAR part for each of the seven series.

S5.2 Results for Different Forecasting Periods

We now report the forecasting results starting from 2000:I, and add graphs of absolute forecasting errors starting from 2000:I for seven economic variables of interest.

Table S13 reports the MSFE for the seven target series of interest with different forecasting periods. We note that our TV-FAVAR model outperforms Bai and Ng's (2006) time-invariant FAVAR model for five out of the seven target series using the predicting sample 2010:I-2019:IV. However, if we use the sample 2000:I-2019:IV, our TV-FAVAR model only outperforms Bai

Table S10: Performance of the estimation in terms of $RMSE$, SSR, and $RMSE_\sigma$

DGP	(N,T)	$RMSE$			SSR			$RMSE_\sigma$		
		BN	LS	TE	BN	LS	TE	BN	LS	TE
1	(100,100)	**0.0821**	0.1989	0.1044	**0.9627**	0.8419	0.9476	**0.0196**	0.0415	0.0233
	(100,200)	**0.0543**	0.1333	0.0684	**0.9846**	0.9154	0.9741	**0.0093**	0.0156	0.0103
	(200,100)	**0.0771**	0.1940	0.1066	**0.9529**	0.8364	0.9336	**0.0211**	0.0435	0.0247
	(200,200)	**0.0561**	0.1330	0.0734	**0.9892**	0.9223	0.9795	**0.0097**	0.0153	0.0105
2	(100,100)	0.4423	**0.2011**	0.2974	1.5647	**0.9147**	1.0882	0.4017	**0.0298**	0.1490
	(100,200)	0.4472	0.1319	**0.1281**	1.6265	**0.9701**	0.9447	0.4334	**0.0126**	0.0143
	(200,100)	0.4440	**0.1927**	0.3014	1.5478	**0.9089**	1.0943	0.3693	**0.0326**	0.1538
	(200,200)	0.4460	0.1333	**0.1303**	1.6310	**0.9790**	0.9494	0.4426	**0.0134**	0.0151
3	(100,100)	0.4593	**0.3447**	0.3939	1.6132	**0.9517**	1.1750	0.4483	**0.0266**	0.1892
	(100,200)	0.4717	**0.2561**	0.2583	1.6579	**0.9997**	0.9876	0.4686	**0.0128**	0.0132
	(200,100)	0.4773	**0.2947**	0.3692	1.5863	**0.9411**	1.1402	0.4159	**0.0292**	0.1704
	(200,200)	0.6315	0.3170	**0.3164**	1.6656	**1.0073**	0.9920	0.4796	**0.0127**	0.0124
4	(100,100)	0.4485	**0.1964**	0.2324	1.6039	0.9233	**0.9771**	0.4432	**0.0303**	0.0664
	(100,200)	0.4490	**0.1275**	0.1276	1.6745	**0.9856**	**0.9856**	0.4966	0.0120	**0.0119**
	(200,100)	0.4446	**0.1919**	0.1998	1.5985	0.9278	**0.9374**	0.4351	**0.0281**	0.0344
	(200,200)	0.4493	**0.1318**	0.1318	1.6547	**0.9860**	**0.9860**	0.4739	**0.0120**	**0.0120**
5	(100,100)	0.4634	**0.3142**	0.3300	1.6450	0.9685	**1.0265**	0.4959	**0.0315**	0.0760
	(100,200)	0.5648	0.2623	**0.2622**	1.6772	1.0243	**1.0241**	0.5010	**0.0146**	0.0146
	(200,100)	0.5730	**0.3349**	0.3416	1.6079	**0.9670**	0.9815	0.4533	**0.0288**	0.0376
	(200,200)	0.5756	**0.2682**	0.2682	1.6626	**1.0159**	**1.0159**	0.4764	**0.0122**	**0.0122**
6	(100,100)	0.3542	**0.1854**	0.2218	1.3649	0.9511	**1.0116**	0.1789	**0.0269**	0.0521
	(100,200)	0.3599	**0.1282**	**0.1282**	1.4144	1.0028	**1.0025**	0.1952	**0.0127**	**0.0127**
	(200,100)	0.3533	**0.1881**	0.2003	1.3681	0.9558	**0.9732**	0.1802	**0.0307**	0.0361
	(200,200)	0.3608	**0.1304**	**0.1304**	1.4081	**1.0072**	**1.0072**	0.1906	**0.0127**	**0.0127**
7	(100,100)	0.5349	**0.3430**	0.3876	1.8999	**1.1929**	1.3342	0.9966	**0.0965**	0.2738
	(100,200)	0.6518	**0.3472**	**0.3472**	1.9634	1.2268	**1.2267**	1.0102	**0.0787**	0.0786
	(200,100)	0.4788	**0.3462**	0.3539	1.8880	**1.2003**	1.2431	0.9461	**0.1052**	0.1500
	(200,200)	0.5455	**0.2811**	**0.2811**	1.9650	**1.2295**	**1.2295**	1.0231	**0.0848**	**0.0848**
8	(100,100)	0.4977	**0.3406**	0.3566	1.7769	**1.1350**	1.1898	0.7375	**0.0675**	0.1311
	(100,200)	0.6306	**0.3402**	**0.3402**	1.8313	**1.1615**	**1.1615**	0.7627	**0.0519**	**0.0519**
	(200,100)	0.6204	**0.4053**	0.4184	1.7665	**1.1395**	1.1723	0.7414	**0.0736**	0.1137
	(200,200)	0.4959	**0.2810**	**0.2810**	1.8351	**1.1725**	**1.1725**	0.7650	**0.0531**	**0.0531**

Notes: (i) BN, LS, and TE denote Bai and Ng's (2006) estimation, the local smoothing estimation proposed in this paper, and the test-based estimation, respectively; (ii) The main entries report the results based on 500 replications; (iii) The bold entries highlight the best performance in each case.

Table S11: Performance of the predictions in terms of RMSFE

DGP	N	T	RMSFE			DGP	N	T	RMSFE		
			BN	LS	TE				BN	LS	TE
1	100	100	**0.9829**	1.2357	1.0041	5	100	100	1.6656	**1.3433**	1.3699
	100	200	**0.9996**	1.1416	1.0113		100	200	1.6412	**1.2093**	1.2093
	200	100	**0.9964**	1.2397	1.0305		200	100	1.6592	**1.3425**	1.3530
	200	200	**0.9877**	1.1423	1.0046		200	200	1.6105	**1.2112**	1.2112
2	100	100	2.4027	1.3534	**1.2985**	6	100	100	1.7194	**1.5129**	1.5524
	100	200	2.3923	1.2143	**1.1442**		100	200	1.6609	**1.2971**	1.2971
	200	100	2.4105	1.3538	**1.2790**		200	100	1.6668	**1.4867**	1.4870
	200	200	2.4103	1.2253	**1.1323**		200	200	1.6219	**1.2863**	1.2863
3	100	100	1.6335	1.3201	**1.2239**	7	100	100	1.3323	**1.2451**	1.2647
	100	200	1.6435	1.2007	**1.1216**		100	200	1.3671	**1.1271**	1.1271
	200	100	1.6795	1.3224	**1.2300**		200	100	1.2949	**1.2181**	1.2225
	200	200	1.6281	1.2118	**1.1151**		200	200	1.3821	**1.1365**	1.1365
4	100	100	2.6230	**1.3484**	1.3882	8	100	100	1.2182	1.1830	**1.1798**
	100	200	2.6349	**1.2003**	1.2003		100	200	1.2876	**1.1159**	1.1159
	200	100	2.6793	1.3605	**1.3559**		200	100	1.2432	**1.1931**	1.1933
	200	200	2.7053	**1.2119**	1.2119		200	200	1.2831	**1.1131**	1.1131

Notes: (i) BN, LS, and TE denote Bai and Ng's (2006) estimation, the local smoothing estimation proposed in this paper, and the test-based estimation, respectively; (ii) The main entries report the results based on 500 replications; (iii) The bold entries highlight the best performance in each case.

Table S12: The determination of lag order in the VAR component

	$p = 1$	$p = 2$	$p = 3$	$p = 4$
RGDP	**2.4189**	3.0333	3.5507	3.9474
PCEC	**2.3858**	2.9902	3.4884	3.8810
IP	**2.3598**	2.9369	3.4621	3.8925
GDPDEF	**2.3010**	2.8977	3.3947	3.8028
LCM	**2.3974**	2.9935	3.5083	3.9666
UR	**2.2673**	2.8428	3.3265	3.7385
FedR	**2.2615**	2.8659	3.3445	3.7403

Note: The main entries report the IC values for the lag order $p = 1, 2, 3, 4$, and the bold elements denote the smallest IC values for each series.

Table S13: The out-of-sample MSFE

	2010:I-2019:IV		2000:I-2019:IV		2000:I-2007Q4 &2010Q1-2019:IV	
	BN06	TV-FAVAR	BN06	TV-FAVAR	BN06	TV-FAVAR
RGDP	0.3482	**0.2387**	**0.4067**	0.5179	0.3047	**0.2787**
PCEC	0.3778	**0.3612**	**0.5789**	0.8526	0.3327	**0.3034**
IP	0.3369	**0.1270**	**0.3060**	0.4662	0.2449	**0.1793**
GDPDEF	0.2332	**0.2269**	**0.1723**	0.1736	0.1602	**0.1415**
LCM	**0.5475**	0.5586	**0.9439**	1.3253	**0.7562**	0.7969
UR	0.0190	**0.0187**	**0.0163**	0.0177	0.0142	**0.0142**
FedR	**0.0026**	0.0035	0.0335	**0.0162**	**0.0048**	0.0052

Note: The main entries report the out-of-sample MSFE for Bai and Ng's (2006, BN06) and our TV-FAVAR models. The bold entries highlight the better performance in each case.

and Ng's (2006) time-invariant one for only one series. Since the proposed local smoothing estimators are nonparametric kernel-based, it is most applicable to the smooth changes in the factor loadings and VAR coefficients. Meanwhile, it requires a relatively larger sample size due to its relatively slow convergence rate compared with the conventional PCA and OLS estimators. The relatively poor performance of our TV-FAVAR model may be caused by the small sample size issue that we have only $T = 158$ observations for the first predicted time point 2000:I. It may also be caused by the possible large structural breaks that occurred during the economic and financial crisis period (2008:I-2009:IV). To remove the possible effect of such large structural breaks, we also evaluate the performance of predictions using the forecasting periods 2000:I-2007:IV and 2010:I-2019:IV. The last two columns of Table S13 report the MSFEs for the seven series of interest. We note that by removing the effect of the economic and financial crisis, the performance of forecasting starting at 2000:I is quite similar to that starting at 2010:I.

Figures S5.2 and S5.2 plot the absolute values of forecast errors obtained from the time-invariant FAVAR model and the TV FAVAR model for the seven target series of interest. We start the predictions in 2000:I and shade the period 2008:I-2009:IV. As shown in Figure S5.2, the TV-FAVAR model performs worse than the time-invariant FAVAR during the economic and financial crisis (2008:I-2009:IV) for the economic variables (RGDP, RCEC, IP). However, the TV-FAVAR performs better than the latter for the other periods. According to Figure S5.2, we note that the results of TV-FAVAR are quite similar to those of time-invariant FAVAR during the non-crisis periods (2000:I-2007:IV and 2010:I-2019:IV). These results are consistent with those reported in Table S13.

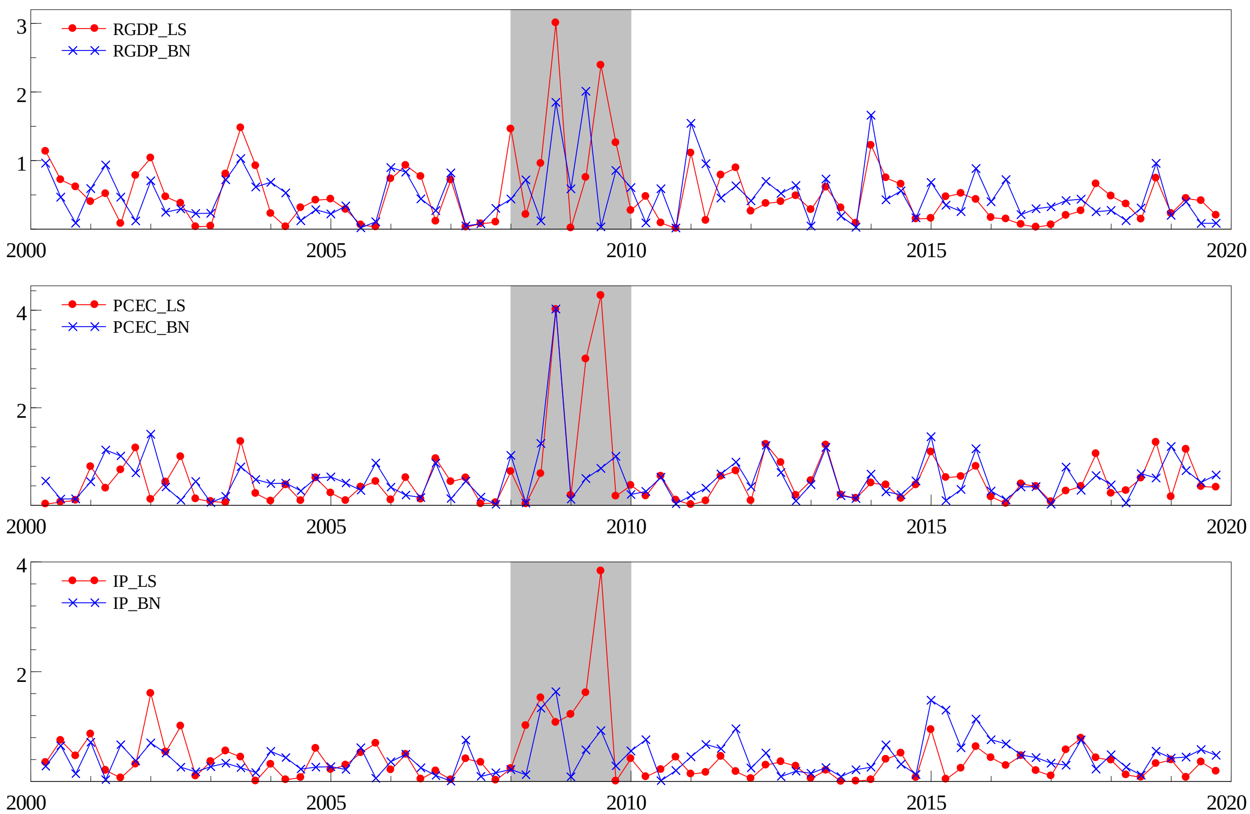

Figure S1: The absolute values of forecast errors for RGDP, PCEC, and IP

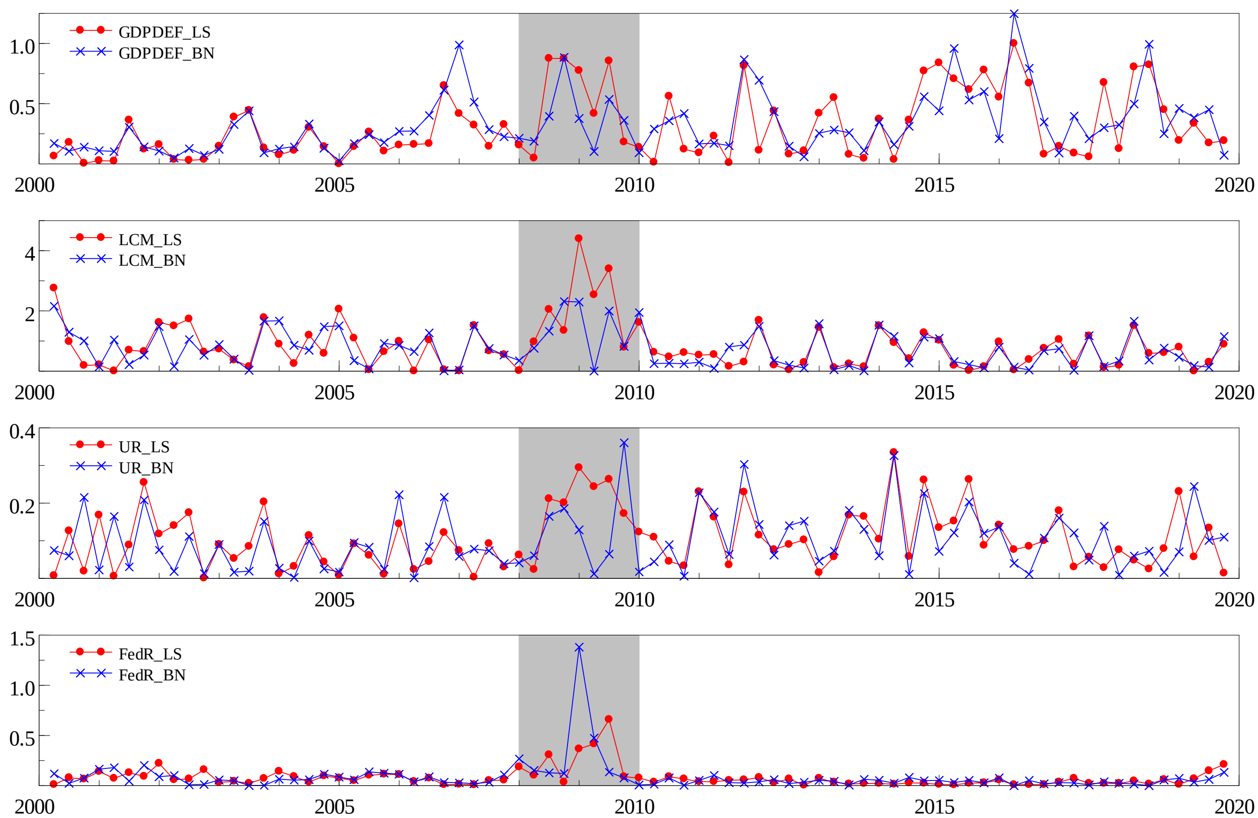

Figure S2: The absolute values of forecast errors for GDPDEF, LCM, UR, FedR

S5.3 Detailed Results for the Forecast Performance and the Test for Predictive Ability

In addition to the MSFE, we also report the ratios of the MSFEs of our local smoothing estimation to those of Bai and Ng's (2006) estimation. In addition, we use the test for predictive ability proposed by Diebold and Mariano (1995) and West (1996) to check the statistical significance of the difference in MSFE. Specifically, letting $d_t = e_{it}^2 - e_{jt}^2$ be the difference in squared forecast errors between model i and model j at time t, $\bar{d} \equiv T^{-1} \sum_{t=1}^{T} d_t$ be the mean difference, and $\rho_\tau \equiv T^{-1} \sum_{t=\tau+1}^{T} (d_t - \bar{d})(d_{t-\tau} - \bar{d})$ be the estimated autocovariance of d_t at lag τ, we compute the test statistic:

$$z = \frac{\bar{d}}{\sqrt{\Omega/T}},$$

where $\Omega \equiv \sum_{l=-L}^{L} (1 - |l|/(L+1)) \rho_l$ with L being the lag order is the Newey-West (1987) Heteroscedasticity and Autocorrelation (HAC) robust estimator of the long-run variance of d_t. Here, we follow Orphanides and Norden (2005) to set $L = 6$ for quarterly data. West (1996) shows that under some conventional assumptions, this statistic is asymptotically normally distributed under the null hypothesis of equal forecasting accuracy. We therefore report the z-statistics and the two-sided p-values using the standard normal distribution in Table S14.

The top panel of Table S14 reports the results for the sample spanning 2010:I to 2019:IV. We observe that our TV-FAVAR model outperforms Bai and Ng's (2006) time-invariant FAVAR model for five out of the seven target series, and the z-statistics are significant for two of the five series at the 5% level. The bottom panel of Table S14 reports the results for the sample spanning 2000:I to 2019:IV. We note that although the time-invariant FAVAR model outperforms our TV-FAVAR model for six out of the seven series, the differences between the two models are insignificant, indicating that there is no significant difference between the two models for the sample period 2000:I to 2019:IV. As mentioned above, the relatively poor performance of our TV-FAVAR model when using 2000:I-2019:IV for out-of-sample prediction may be caused by the possible large structural breaks that occur during the economic and financial crisis period.

Table S14: The out-of-sample MSFE

	BN06	TV-FAVAR	Ratio	$z-$test	p-value
2010:I-2019:IV					
RGDP	0.3482	**0.2387**	0.6856	2.6273	**0.0086**
PCEC	0.3778	**0.3612**	0.9559	0.4423	0.6583
IP	0.3369	**0.1270**	0.3771	2.3723	**0.0177**
GDPDEF	0.2332	**0.2269**	0.9728	0.2388	0.8113
LCM	**0.5475**	0.5586	1.0204	-0.3016	0.7629
UR	0.0190	**0.0187**	0.9816	0.1322	0.8948
FedR	**0.0026**	0.0035	1.3239	-0.6946	0.4873
2000:I-2019:IV					
	BN06	TV-FAVAR	Ratio	$z-$test	p-value
RGDP	**0.4067**	0.5179	1.2735	1.0123	0.3114
PCEC	**0.5789**	0.8526	1.4730	0.9181	0.3586
IP	**0.3060**	0.4662	1.5234	0.7554	0.4500
GDPDEF	**0.1723**	0.1736	1.0079	0.0540	0.9569
LCM	**0.9439**	1.3253	1.4041	1.2801	0.2005
UR	**0.0163**	0.0177	1.0888	0.5812	0.5611
FedR	0.0335	**0.0162**	0.4850	-0.9325	0.3511

Notes: (i) The entries under 'BN06' and 'TV-FAVAR' report the out-of-sample MSFEs for Bai and Ng's (2006, BN06) and the proposed TV-FAVAR models, respectively. The bold entries highlight the better performance in each case. (ii) The entries under 'Ratio' report the ratios of the MSFEs of the local smoothing estimation to those of Bai and Ng's (2006) estimation. (iii) 'z-test' and 'p-value' denote Diebold and Mariano's (1995) z-statistics and the corresponding p-values.

References

Bai J., Ng, S. (2006), "Confidence Intervals for Diffusion Index Forecasts and Inference for Factor-augmented Regressions," Econometrica, 74(4): 1133-1150.

Diebold F. X., Mariano, R. S. (1995), "Comparing predictive accuracy. Journal of Business & economic statistics, 13(3): 253-263.

Newey, W. K., & West, K. D. (1987), "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," Econometrica, 55(3), 703-708.

Orphanides, A., Van Norden, S. (2005), "The Reliability of Inflation Forecasts Based on Output Gap Estimates in Real Time," Journal of Money, Credit and Banking, 583-601.

Su, L., and Wang, X. (2017), "On Time-varying Factor Models: Estimation and Testing," Journal of Econometrics, 198, 84-101.

Su, L., and Wang, X. (2020a), "Testing for Structural Changes in Factor Models via a Nonparametric Regression," Econometric Theory, 36(6), 1127-1158.

Su, L., and Wang, X. (2020b), "Corrigendum to 'On Time-varying Factor Models: Estimation and Testing'," [J. Econometrics 198 (2017) 84-101], Working Paper, Singapore Management University.

Sun, S., and Chiang, C. Y. (1997), "Limiting Behavior of the Perturbed Empirical Distribution Functions Evaluated at U-statistics for Strongly Mixing Sequences of Random Variables," Journal of Applied Mathematics and Stochastic Analysis, 10(1), 3-20.

West K. D. (1996), "Asymptotic Inference about Predictive Ability," Econometrica, 64, 1067-1084.