CENTER FOR POLICY RESEARCH
THE MAXWELL SCHOOL

WORKING PAPER SERIES

Asymptotic Power of the Sphericity Test Under Weak and Strong Factors in a Fixed Effects Panel Data Model

Badi H. Baltagi, Chihwa Kao, and
Fa Wang

Paper No. 189
March 2016

ISSN: 1525-3066

426 Eggers Hall
Syracuse University
Syracuse, NY 13244-1020
(315) 443-3114/email: ctrpol@syr.edu

http://www.maxwell.syr.edu/CPR_Working_Papers.aspx

CENTER FOR POLICY RESEARCH –Spring 2016

Leonard M. Lopoo, Director
Associate Professor of Public Administration and International Affairs (PAIA)

Associate Directors

Margaret Austin
Associate Director
Budget and Administration

John Yinger
Trustee Professor of Economics and PAIA
Associate Director, Metropolitan Studies Program

SENIOR RESEARCH ASSOCIATES

Badi H. Baltagi .... Economics
Robert Bifulco .... PAIA
Leonard Burman .... PAIA
Thomas Dennison .... PAIA
Alfonso Flores-Lagunes .... Economics
Sarah Hamersma .... PAIA
William C. Horrace .... Economics
Yilin Hou .... PAIA
Hugo Jales .... Economics
Duke Kao .... Economics
Jeffrey Kubik .... Economics
Yoonseok Lee .... Economics
Amy Lutz .... Sociology
Yingyi Ma .... Sociology
Jerry Miner .... Economics
Cynthia Morrow .... PAIA
Jan Ondrich .... Economics
John Palmer .... PAIA
David Popp .... PAIA
Stuart Rosenthal .... Economics
Rebecca Schewe .... Sociology
Amy Ellen Schwartz .... PAIA/Economics
Perry Singleton .... Economics
Michael Wasylenko .... Economics
Peter Wilcoxen .... PAIA

GRADUATE ASSOCIATES

Emily Cardon .... PAIA
Brianna Carrier .... PAIA
John T. Creedon .... PAIA
Carlos Diaz .... Economics
Alex Falevich .... Economics
Wancong Fu .... Economics
Boqian Jiang .... Economics
Yusun Kim .... PAIA
Ling Li .... Economics
Michelle Lofton .... PAIA
Judson Murchie .... PAIA
Brian Ohl .... PAIA
Jindong Pang .... Economics
Malcolm Philogene .... PAIA
William Reed .... PAIA
Laura Rodriquez-Ortiz .... PAIA
Fabio Rueda De Vivero .... Economics
Max Ruppersburg .... PAIA
Iuliia Shybalkina .... PAIA
Kelly Stevens .... PAIA
Mary Stottele .... PAIA
Tian Tang .... PAIA
Saied Toossi .... PAIA
Rebecca Wang .... Sociology
Nicole Watson .... Social Science
Katie Wood .... PAIA
Jinqi Ye .... Economics
Pengju Zhang .... PAIA
Xirui Zhang .... Economics

STAFF

Kelly Bogart .... Administrative Specialist
Kathleen Nasto .... Administrative Assistant
Candi Patterson .... Computer Consultant
Mary Santy .... Administrative Assistant
Katrina Wingle .... Administrative Assistant

Abstract

This paper studies the asymptotic power for the sphericity test in a fixed effect panel data model proposed by Baltagi, Feng and Kao (2011), (JBFK). This is done under the alternative hypotheses of weak and strong factors. By weak factors, we mean that the Euclidean norm of the vector of the factor loadings is O(1). By strong factors, we mean that the Euclidean norm of the vector of factor loadings is O(pn), where n is the number of individuals in the panel. To derive the limiting distribution of JBFK under the alternative, we first derive the limiting distribution of its raw data counterpart. Our results show that, when the factor is strong, the test statistic diverges in probability to infinity as fast as Op(nT). However, when the factor is weak, its limiting distribution is a rightward mean shift of the limit distribution under the null. Second, we derive the asymptotic behavior of the difference between JBFK and its raw data counterpart. Our results show that when the factor is strong this difference is as large as Op(n). In contrast, when the factor is weak, this difference converges in probability to a constant. Taken together, these results imply that when the factor is strong, JBFK is consistent, but when the factor is weak, JBFK is inconsistent even though its asymptotic power is nontrivial.

JEL No. C12; C33

Keywords: Asymptotic power; Sphericity; John Test; Weak Factor; Strong Factor; High Dimensional Inference; Panel Data

Badi H. Baltagi, Chihwa Kao, and Fa Wang, Syracuse University, The Center for Policy Research, 426 Eggers Hall, Syracuse, NY 13244

We dedicate this paper in honor of Esfandiar Maasoumi's many contributions to econometrics. We would like to thank the editors Aman Ullah and Peter Phillips and two anonymous referees for useful comments and suggestions.

1. INTRODUCTION

This paper studies the asymptotic power of the John (1972) test for sphericity of the covariance matrix of the error term which was extended by Baltagi, Feng and Kao (2011) to a fixed effects panel data model. We consider the large n large T setup. Typically, the number of cross-sectional units n in a panel is large, while the number of time series observations T could be either large (in macro applications) or small in (micro applications). Labor panels are typical of micro-panels with hundreds of individuals observed over a few time periods. While panels in finance may involve hundreds of stocks observed over hundreds of days. When n tends to infinity jointly with T, generic results in random matrix theory show that the spectral norm of the sample covariance matrix does not converge to that of the population covariance matrix and follows a Tracy–Widom distribution asymptotically, see Geman (1980) and Johnstone (2001). In addition, if $\frac{n}{T} \to c \in (0, \infty)$, the eigenvalues of the sample covariance matrix vary between $(1-\sqrt{c})^2$ and $(1+\sqrt{c})^2$, while the eigenvalues of the population covariance matrix are all one, see Bai (1999). These results indicate that when the dimension tends to infinity jointly with sample size, the sample covariance matrix is no longer consistent for the population covariance matrix, and consequently cast doubt on the consistency of BFK's John test (J_{BFK}) since the latter is based on the sample covariance matrix. Furthermore, BFK's John test is based on the within residuals rather than the real error term, and its consistency is not guaranteed.

Studying the asymptotic power is also empirically motivated. Intuitively, the empirical power should depend on how strong the cross-sectional dependence is. In case the cross-sectional dependence is due to common factors, the cross-sectional dependence would be weak if factors are weak. In case the cross-sectional dependence is due to spatial effects, the cross-sectional dependence would still likely to be weak since spatial effects are typically local and thus can be regarded as weak factors. Asymptotic power derived under the sequence of weak factor alternatives therefore provides better approximation of the empirical power when cross-sectional dependence is weak. The asymptotic scheme under the sequence of weak factor alternatives is also similar to the pitman drift, which is used in Staiger and Stock (1997) to obtain the asymptotic approximation of the finite sample distribution of 2SLS and LIML estimators when the instruments are weak.

In the statistics literature, several papers analyzed the asymptotic power of the test for sphericity in a high dimensional setup. Srivastava (2005) proposed tests for the identity, sphericity and diagonality of the covariance matrix based on estimators of the first and second moments of the spectral distribution of the population covariance matrix. Srivastava derived limit distributions under both the null and alternative. Wang, Cao and Miao (2013) proposed similar tests and derived their limit distributions under both the null and alternative, but these tests were based on estimators of the second and fourth moments rather than the first and second moments. Chen, Zhang and Zhong (2010) proposed U-statistics based tests for the identity and sphericity of the covariance matrix and derived their limit distribution under both the null and alternative. Cai and Ma (2013), on the other hand, studied this problem from a minimax perspective. They characterized the boundary that separates the testable region from the non-testable region by the Frobenius norm when the ratio of the dimension and the sample size is bounded. Using Le Cam's Lemma 1, Onatski, Moreira and Hallin (2013, 2014), hereafter (OMH), established mutual contiguity of the joint distributions of the sample covariance eigenvalues under the null and alternative when the alternative is a low rank perturbation of the null and the norm of perturbation is fixed and less than a threshold. Next, they derived the asymptotic power of all sample covariance eigenvalue based tests using Le Cam's Lemma 3. OMH's result is thought-provoking in the sense that it builds up the connection between high dimensionality and Pitman drift, or roughly speaking, weak identification, although only for a special class of alternatives. A key shortcoming of OMH's result is that it does not allow us to calculate the asymptotic power when the norm of perturbation is greater than the threshold or when it goes to infinity.

This paper studies the asymptotic power of the BFK John test under the alternative hypotheses of weak and strong factors. By weak factors, we mean that the Euclidean norm of the vector of the factor loadings is $O(1)$. By strong factors, we mean that the Euclidean norm of the vector of factor loadings is $O(n)$, where n is the number of individuals in the panel. These correspond to strong and weak cross-sectional dependence, respectively, see Chudik and Pesaran (2013). To derive the limiting distribution of J_{BFK} under the alternative, we first derive the limiting distribution of its raw data counterpart. Our results show that, when the factor is strong, it diverges to infinity in probability as fast as $O_p(nT)$. When the factor is weak, its limiting distribution is a rightward mean shift of the limit distribution under the null. The magnitude of the mean shift is proportional to the norm of variance adjusted factor loadings and the sample size, and inversely proportional to the dimension. This result is in sharp contrast to the fixed dimension case in which the asymptotic power tends to one as the sample size tend to infinity if the norm of perturbation is fixed. This result also indicates that the effect of increasing the dimension on asymptotic power is similar to Pitman drifting the parameter. We then derive the asymptotic behavior of the difference between J_{BFK}

and its raw data counterpart. This difference is due to the additional noise in J_{BFK} resulting from the estimation of the regression coefficients β and the fixed effects μ_i. Our results show that when the factor is strong, this difference is as large as $O_p(n)$. When the factor is weak, this difference converges in probability to a constant, $c/2$. These results also contrast with the fixed dimension case in which the additional noise resulting from $\hat{\beta}-\beta$ and μ_i will be smoothed away as the sample size tends to infinity. In summary, due to the effect of increasing dimension, J_{BFK} is inconsistent under the weak factor alternative, although it still has nontrivial asymptotic power. Under the strong factor alternative, J_{BFK} is consistent, since the cross-sectional dependence is strong enough to outweigh the effect of increasing dimension, i.e., $O_p(nT)$ dominates $O_p(n)$. Our results also shed light on the asymptotic power of the tests for cross-sectional independence in panel data recently proposed in Pesaran (2004, 2012), Pesaran, Ullah and Yamagata (2008) and Baltagi, Feng and Kao (2012). We leave these extensions for a future study.[1]

The organization of this paper is as follows. Section 2 introduces the model, notation and assumptions. Section 3 introduces BFK's John test of sphericity. Section 4 studies the asymptotic power of BFK's John test, and Section 5 concludes. The appendix contains all the proofs and technical details.

2. NOTATION AND PRELIMINARIES

Consider the fixed effects panel data model,

$$y_{it} = x_{it}'\beta + \mu_i + \nu_{it,} \text{ for } i = 1, ..., n \text{ and } t = 1, ..., T, \tag{1}$$

where i is the index of the cross-sectional units, t is the index of the time series observations, μ_i is the time invariant individual effects which could be fixed or random. ν_{it} is the idiosyncratic error term.

Assumption 1 *For any $i, j = 1, ..., n$; and $t, l = 1, ..., T$, the regressors x_{it} and the idiosyncratic error terms ν_{jl} are independent, and x_{it} have finite 4th moments.*

Assumption 2 *Let $\nu_t = (\nu_{1t}, ..., \nu_{nt})'$, the $n \times 1$ vectors $\nu_1, ..., \nu_T$ are iid $N(0, \Sigma_n)$, where Σ_n is an $n \times n$ general population covariance matrix.*

[1] Cross-sectional dependence, due to either spatial or common factor effects, is prevalent in economic data. Chudik and Pesaran (2013) argued that even after controlling for heterogeneity in panel data, cross-sectional dependence still arises. Ignoring cross-sectional dependence may lead to misleading inference and even inconsistent estimation. Therefore, testing the presence and extent of cross-sectional dependence is very important. See also the special issue of Econometric Reviews edited by Baltagi and Maasoumi (2013) which deals with several aspects of dependence in time-series, cross-section and panels.

Assumption 3 $\frac{n}{T} \to c \in (0, \infty)$, *as n and T go to infinity jointly. This is diagonal path asymptotics not joint asymptotics as in Phillips and Moon (1999).*

Assumption 1 is a standard but albeit restrictive requirement for the consistency of the fixed effects estimator. Assumption 2 allows for any form of heteroskedasticity and cross-sectional dependence. The covariance matrix is only required to be stable over time. The restrictive part of Assumption 2 is the normality and no serial correlation over time of the error term. These are assumed to simplify the derivation of the limiting distribution of BFK's John test. Assumption 3 imposes a condition on the relative speed at which n and T go to infinity. More specifically, it should be: $\frac{n_T}{T} \to c \in (0, \infty)$, but we suppress the subscript T hereafter for simplicity. This large n and large T setup is more appropriate than the fixed n and large T setup for macroeconomic applications in which typically n and T are both large and of comparable magnitudes. In model (1), the within estimator of β is

$$\tilde{\beta} = \beta + (\sum_{i=1}^{n} \sum_{t=1}^{T} \tilde{x}_{it}\tilde{x}_{it}')^{-1}(\sum_{i=1}^{n} \sum_{t=1}^{T} \tilde{x}_{it}\tilde{\nu}_{it}), \tag{2}$$

where $\tilde{x}_{it} = x_{it} - \bar{x}_{i\cdot}$ and $\tilde{\nu}_{it} = \nu_{it} - \bar{\nu}_{i\cdot}$, with $\bar{x}_{i\cdot} = \sum_{t=1}^{T} x_{it}/T$, and $\bar{\nu}_{i\cdot} = \sum_{t=1}^{T} \nu_{it}/T$. Under Assumptions 1, 2 and 3, $\tilde{\beta}$ is a consistent estimator of β.

Throughout the paper, trA is the trace of matrix A, $\|A\| = (trAA')^{\frac{1}{2}}$ denotes the Frobenius norm, $\|x\|$ denotes the Euclidean norm of vector x, $\xrightarrow{p}$ denotes convergence in probability, $\xrightarrow{d}$ denotes convergence in distribution, $(N, T) \to \infty$ denotes N and T going to infinity jointly.

3. BFK'S JOHN TEST

This section gives a quick review of BFK's John test for sphericity. In order not to impose any structure on the population covariance matrix, tests for sphericity are based on the sample covariance matrix. It is important to note that when $n > T$ the sample covariance matrix becomes singular, and consequently the likelihood ratio test for sphericity is no longer feasible. As such, John (1971) proposed a sphericity test defined as follows:

$$U = \frac{1}{n} tr[(\frac{1}{n} trS)^{-1} S - I_n]^2 = (\frac{1}{n} trS)^{-2} (\frac{1}{n} trS^2) - 1, \tag{3}$$

where S is sample covariance matrix and I_n is an $n \times n$ identity matrix. Under the null of sphericity and when n is fixed and $T \to \infty$, $\frac{1}{n} trS$ is a consistent estimator of the variance of the error term, σ^2. Hence, $(\frac{1}{n} trS)^{-1} S$ is a normalized sample covariance matrix and $tr[(\frac{1}{n} trS)^{-1} S - I_n]^2$ measures the distance between this normalized sample covariance matrix and the identity matrix. John (1972) showed that under the null with n fixed and $T \to \infty$,

$$J = \frac{nT}{2}U \xrightarrow{d} \chi^2_{\frac{n(n+1)}{2}-1}.$$

However, as n increases the John test is significantly oversized. In fact, it can be shown that as $n \to \infty$, John's test diverges to infinity in probability. To correct the size distortion, Ledoit and Wolf (2002), hereafter (LW), recentered and rescaled John's test as follows:

$$J_{LW} = \frac{TU - n - 1}{2} = \frac{1}{n}(J - \frac{n^2}{2} - \frac{n}{2}). \tag{4}$$

Under the null hypothesis, with $(n, T) \to \infty$ and $\frac{n}{T} \to c \in (0, \infty)$, Ledoit and Wolf (2002) showed that

$$J_{LW} \xrightarrow{d} N(0, 1). \tag{5}$$

Both the John test and the LW's John test are based on the true error term, while in the fixed effects panel data model the test statistics are based on within residuals. In the fixed n and large T setup, the extra noise contained in the within residuals vanishes gradually as $T \to \infty$. Hence, it is reasonable to believe that the test statistics based on the true error term and within residuals should be asymptotically equivalent.

However, this is no longer true when n and T are both large and of comparable magnitudes, since each $\tilde{\nu}_{it}$ contains an extra noise and their number is n. To bridge this gap, Baltagi, Feng and Kao (2011) studied the asymptotic behavior of $\widehat{J}_{LW} - J_{LW}$, where $\widehat{J}_{LW}$ is LW's John test based on within residuals. They proved that under the null hypothesis with $(n, T) \to \infty$ and $\frac{n}{T} \to c \in (0, \infty)$, $\widehat{J}_{LW} - J_{LW} - \frac{n}{2(T-1)} \xrightarrow{p} 0$. It follows that under the null,

$$J_{BFK} = \widehat{J}_{LW} - \frac{n}{2(T-1)} \xrightarrow{d} N(0, 1). \tag{6}$$

4. ASYMPTOTIC POWER OF BFK'S JOHN TEST

This section studies the asymptotic power of BFK's John test under the weak and strong factor alternatives. The null hypothesis is:

$$H_0 : \Sigma_n = \sigma^2 I_n. \tag{7}$$

Under the alternative, $\nu_{it} = \sum_{j=1}^{r} \gamma_{ij} f_{tj} + \epsilon_{it}$, where γ_{ij} is the factor loading of individual i for factor j, f_{tj} is the factor j in period t, r is the known number of factors. Hence, $\Sigma_n = E(\nu_t \nu_t') = E(\sum_{j=1}^{r} \gamma_j f_{tj} + \epsilon_t)(\sum_{j=1}^{r} \gamma_j f_{tj} + \epsilon_t)'$. To simplify the analysis, we make the following assumptions:

Assumption 4 *1. Each factor f_{tj} is iid $N(0,\sigma_j^2)$ across time, and the variance σ_j^2 is bounded.*

2. The idiosyncratic error ϵ_{it} is iid $N(0,\sigma^2)$, and independent of all factors.

3. The correlation coefficient between factors f_{tj} and f_{tk} is zero, for all j, k and t.

4. The vectors of factor loading γ_j are orthogonal to each other.

Although these assumptions are restrictive, Assumption (4) will not lead to loss of generality. Time dependence of the factors is likely present in real data, but as long as such dependence is not strong, the asymptotic power property will not change qualitatively. The idiosyncratic error ϵ_{it} may still have cross-sectional dependence, if cross-sectional dependence in ν_{it} cannot be totally filtered by the factor structure. Nonetheless, adding additional cross-sectional dependence in ϵ_{it} will not change the results as long as such dependence is weak. Parts 3 and 4 in assumption (4) are innocuous since factors and factor loadings are identifiable only up to a rotation, and from this normalization we can always redefine factors and factor loadings so that parts 3 and 4 are satisfied.

Under Assumption (4),

$$E(\sum\nolimits_{j=1}^{r}\gamma_j f_{tj}+\epsilon_t)(\sum\nolimits_{j=1}^{r}\gamma_j f_{tj}+\epsilon_t)'=\sigma^2(I_n+\sum\nolimits_{j=1}^{r}\frac{\sigma_j^2}{\sigma^2}\gamma_j\gamma_j'), \tag{8}$$

where $\gamma_j=(\gamma_{1j},...,\gamma_{nj})'$ is the vector of factor loading. Normalizing γ_j, we get

$$\Sigma_n=\sigma^2(I_n+\sum\nolimits_{j=1}^{r}\frac{\sigma_j^2}{\sigma^2}\left\|\gamma_j\right\|^2\frac{\gamma_j}{\left\|\gamma_j\right\|}\frac{\gamma_j'}{\left\|\gamma_j\right\|})=\sigma^2(I_n+\sum\nolimits_{j=1}^{r}h_je_je_j'), \tag{9}$$

where $h_j=\frac{\sigma_j^2}{\sigma^2}\left\|\gamma_j\right\|^2$, $e_j=\frac{\gamma_j}{\left\|\gamma_j\right\|}$ and $\|e_j\|=1$. Therefore, the sequence of alternative hypothesis is:

$$H_a:\Sigma_n=\sigma^2(I_n+\sum\nolimits_{j=1}^{r}h_je_je_j'). \tag{10}$$

In this expression, the covariance matrix is a rank-r perturbation of sphericity. Each e_je_j' characterizes one direction of perturbation and h_j is the magnitude of the perturbation along this direction. Obviously, the asymptotic power under this sequence of alternatives depends upon how h_j evolves as $(n,T)\to\infty$. We will study two different cases, $h_j/n\to d_j\in(0,\infty)$ and $h_j\to d_j\in(0,\infty)$, which correspond to the strong and weak factor cases considered recently by Bai (2003), Onatski (2012) and Johnstone and Lu (2009). To calculate the asymptotic power of the BFK's John test, we need to derive the limiting distribution of J_{BFK} under the alternative hypothesis. This can be done in two steps. First, we derive the limiting distribution of J_{LW} under the alternative. Second, we derive the asymptotic behavior of $\widehat{J}_{LW} - J_{LW}$ under the alternative. Note that $J_{BFK} = \widehat{J}_{LW} - \frac{n}{2(T-1)}$, once the limiting distribution of $\widehat{J}_{LW}$ is known, that of J_{BFK} follows.

4.1. Asymptotic Power under the Weak Factor Alternative

Theorem 1 *Under Assumptions 2-4, and under the weak factor alternative with* $h_j \to d_j \in (0, \infty)$ *for* $j = 1, ..., r$,

$$J_{LW} - \frac{T\sum_{j=1}^{r} d_j^2}{2n} \xrightarrow{d} N(0,1). \tag{11}$$

or equivalently

$$J_{LW} \xrightarrow{d} N\left(\frac{\sum_{j=1}^{r} d_j^2}{2c}, 1\right) \tag{12}$$

Theorem 1 implies that under the weak factor alternative, the limiting distribution of J_{LW} is a mean shift of its limiting distribution under the null. The magnitude of the mean shift is proportional to the magnitude of variance adjusted factor loadings $\sum_{j=1}^{r} d_j^2$ and the sample size T, and inversely proportional to the dimension n. Here, $\sum_{j=1}^{r} d_j^2$ plays the role of the local parameter in traditional asymptotic optimality analysis. On the one hand, the test statistic gets increasingly sensitive to the underlying parameter as the sample size T goes to infinity. On the other hand, the weak factor alternative gets increasingly difficult to be discriminated as the dimension n goes to infinity. This is because the effect of a perturbation of the covariance matrix with fixed norm on J_{LW}'s distribution gets dissipated as the dimension increases. In other words, the effective distance between the null and weak factor alternative decreases as the dimension increases. Therefore, the limiting distribution under the alternative also depends on the relative speed of n and T and $\frac{\sum_{j=1}^{r} d_j^2}{2c}$ can be interpreted as a discounted local parameter. The detailed proof of this theorem is in the Appendix. This result is also partially proved by Onatski, Moreira and Hallin (2013, 2014) in which they derived the asymptotic power of all sample covariance eigenvalue based tests, including J_{LW}, but only when all h_j are below the threshold $\sqrt{c}$.

Next, we study the asymptotic behavior of $\widehat{J}_{LW} - J_{LW}$ under the weak factor alternative. Let $\hat{S}$ be the sample covariance matrix calculated using the within residuals, it follows that

$$\begin{aligned} \widehat{J}_{LW} - J_{LW} &= \left(\frac{T(\frac{1}{n}tr\hat{S})^{-2}\frac{1}{n}tr\hat{S}^2 - T - n}{2} - \frac{1}{2}\right) - \left(\frac{T(\frac{1}{n}trS)^{-2}\frac{1}{n}trS^2 - T - n}{2} - \frac{1}{2}\right) \\ &= \frac{T[(\frac{1}{n}trS)^2\frac{1}{n}tr\hat{S}^2 - (\frac{1}{n}tr\hat{S})^2\frac{1}{n}trS^2]}{2(\frac{1}{n}tr\hat{S})^2(\frac{1}{n}trS)^2}. \end{aligned} \tag{13}$$

Define $W_1 = \frac{1}{n}tr\hat{S} - \frac{1}{n}trS$ and $W_2 = \frac{1}{n}tr\hat{S}^2 - \frac{1}{n}trS^2$, then

$$\widehat{J}_{LW} - J_{LW} = \frac{TW_2(\frac{1}{n}trS)^2 - 2TW_1\frac{1}{n}trS\frac{1}{n}trS^2 - TW_1^2\frac{1}{n}trS^2}{2(\frac{1}{n}trS + W_1)^2(\frac{1}{n}trS)^2}. \tag{14}$$

From this expression, we can clearly see that the asymptotic behavior of $\widehat{J}_{LW} - J_{LW}$ depends upon the asymptotic behavior of $\frac{1}{n}trS$, $\frac{1}{n}trS^2$, $\frac{1}{n}tr\hat{S} - \frac{1}{n}trS$ and $\frac{1}{n}tr\hat{S}^2 - \frac{1}{n}trS^2$. These, in turn, are studied in the following proposition.

Proposition 1 *Under Assumptions 1-4, and under the weak factor alternative with $h_j \to d_j \in (0, \infty)$ for $j = 1, ..., r$,*

(a) $\frac{1}{n}trS = \sigma^2 + O_p(\frac{1}{\sqrt{nT}})$,

(b) $\frac{1}{n}trS^2 = (\frac{n}{T} + 1)\sigma^4 + O_p(\frac{1}{\sqrt{T}})$,

(c) $\frac{1}{n}tr\hat{S} - \frac{1}{n}trS = -\frac{\sigma^2}{T} + O_p(\frac{1}{T\sqrt{n}})$,

(d) $\frac{1}{n}tr\hat{S}^2 - \frac{1}{n}trS^2 = -\frac{2}{T}\sigma^4 - \frac{n}{T^2}\sigma^4 + O_p(\frac{1}{T\sqrt{T}})$.

Part (a) describes the asymptotic behavior of the average of the sample variance. It implies that, in estimating the population variance, the noise contained in the estimator $\frac{1}{n}trS$ is of magnitude $O_p(\frac{1}{\sqrt{nT}})$. Note that $\frac{1}{n}trS = \frac{1}{n}tr[\frac{1}{T}\sum_{t=1}^{T}\nu_t\nu_t'] = \frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\nu_{it}^2$, so under the null, the above result follows directly from the Central Limit Theorem. Under the alternative, with cross-sectional dependence, $\frac{1}{n}trS$ is no longer the sum of independent random variables. However, weak factor implies weak cross-sectional dependence. Hence $\frac{1}{n}trS$ has the same asymptotic behavior as that obtained under the null.

Part (b) shows that under the weak factor alternative and with $\frac{n}{T} \to c \in (0, \infty)$, $\frac{1}{n}trS^2$ converges in probability to $(c+1)\sigma^4$. This implies that, in the large n and large T setup, $\frac{1}{n}trS^2$ is not a consistent estimator of σ^4. Note that if n is fixed and T tends to infinity, as in deriving the limiting distribution of the Breusch and Pagan (1980) test for cross-sectional dependence, $\frac{1}{n}trS^2$ is consistent.[2] What explains this difference? Note that the number of noisy terms in the expansion of trS^2 is related to n^2. After dividing by n, the number of noisy terms in $\frac{1}{n}trS^2$ is related to n.

[2] One of the early tests for cross-sectional dependence is the traditional Breusch and Pagan (1980) test which relies on fixed n and large T asymptotics. Empirical evidence shows that when n is large, the Breusch-Pagan test is significantly oversized. In the statistics literature, this oversizing phenomenon also appears in the classic likelihood ratio test of the covariance matrix, see Bai, et al. (2009). Several attempts have been made to improve the finite sample properties of the Breusch-Pagan test. In fact, Frees (1995) proposed a nonparametric test based on the spearman's rank correlation coefficient, while Dufour and Khalaf (2002) suggested some Monte Carlo exact tests. The Dufour and Khalaf tests are computationally intensive since they are based on the bootstrap method. Another approach is to correct for the size distortion of the Breusch-Pagan test, see Pesaran (2004), Pesaran, Ullah and Yamagata (2008) and Baltagi, Feng and Kao (2012).

On the other hand, the magnitude of noise in each term is $O_p(\frac{1}{\sqrt{T}})$. As n and T tend to infinity jointly, these noise can not be smoothed away and accumulate into a bias, $\frac{n}{T}\sigma^4$.

Parts (c) and (d) show that, in $\frac{1}{n}tr\hat{S} - \frac{1}{n}trS$, the additional noise contained in the within residuals will accumulate into a term of magnitude $-\frac{\sigma^2}{T} + O_p(\frac{1}{T\sqrt{n}})$, and in $\frac{1}{n}tr\hat{S}^2 - \frac{1}{n}trS^2$, this additional noise will accumulate into a term of magnitude $O_p(\frac{1}{T}) + O_p(\frac{n}{T^2})$. These two results share the same intuition with part (b). Note that $\hat{\nu}_{it} = \nu_{it} - \bar{\nu}_{i\cdot} - \tilde{x}_{it}'(\tilde{\beta} - \beta)$, where ν_{it} is the error term, $\hat{\nu}_{it}$ is the within residual, $\tilde{\beta}$ is the within estimator and $\tilde{x}_{it} = x_{it} - \bar{x}_{i\cdot}$ denote the demeaned regressors. From this expression, it is easy to see that the additional noise comes from $\tilde{\beta} - \beta$ and $\bar{\nu}_{i\cdot}$. $\tilde{\beta}$ is $\sqrt{nT}$ consistent, hence $\tilde{\beta} - \beta$ converges to zero in probability no matter whether n is fixed or tends to infinity jointly with T. $\bar{\nu}_{i\cdot}$ is of magnitude $1/\sqrt{T}$, hence if n is fixed, $\bar{\nu}_{i\cdot}$ would be smoothed away as $T \to \infty$. However, if n goes to infinity jointly with T, although each $\bar{\nu}_{i\cdot}$ converges to zero in probability, the number of $\bar{\nu}_{i\cdot}$ tends to infinity jointly. In the end, how this noise $\bar{\nu}_{i\cdot}$ accumulates depends upon the specific form of the test statistic and the alternative. The detailed proof of this proposition is in the Appendix.

Based on Proposition 1, we have the following theorem.

Theorem 2 *Under Assumptions 1-4, and under the weak factor alternative with* $h_j \to d_j \in (0, \infty)$ *for* $j = 1, ..., r$,

$$\widehat{J}_{LW} - J_{LW} - \frac{n}{2(T-1)} \xrightarrow{p} 0. \tag{15}$$

This theorem implies that for J_{LW} the additional noise contained in the within residuals will accumulate into a constant, $\frac{c}{2}$. Note that this pattern of accumulation relies heavily on the assumption $\frac{n}{T} \to c \in (0, \infty)$ and $h_j \to d_j \in (0, \infty)$ for $j = 1, ..., r$. If $\frac{n}{T} \to \infty$ or $h_j \to \infty$ for some j, the accumulated noise may explode. The detailed proof is in the Appendix.

Note that $J_{BFK} = \widehat{J}_{LW} - \frac{n}{2(T-1)}$, thus Theorem 2 implies $J_{BFK} - J_{LW} \xrightarrow{p} 0$. Combining this with Theorem 1, we have:

Corollary 1 *Under Assumptions 1-4, and under the weak factor alternative with* $h_j \to d_j \in (0, \infty)$ *for* $j = 1, ..., r$,

$$J_{BFK} \xrightarrow{d} N\left(\frac{\sum_{j=1}^{r} d_j^2}{2c}, 1\right) \tag{16}$$

Recall that Baltagi, Feng and Kao (2011) proved that under the null, $J_{BFK} \xrightarrow{d} N(0, 1)$, thus the asymptotic power of J_{BFK} under the weak factor alternative is given in the following theorem:

Theorem 3 *Under Assumptions 1-4, and under the weak factor alternative with* $h_j \to d_j \in (0, \infty)$ *for* $j = 1, ..., r$*, the asymptotic power of* J_{BFK} *is*

$$Power_{J_{BFK}}(d) = 1 - \Phi(\Phi^{-1}(1-\alpha) - \frac{\sum_{j=1}^{r} d_j^2}{2c}), \tag{17}$$

where $\Phi(\cdot)$ *denotes the cdf of a* $N(0,1)$ *and* $d = (d_1, ..., d_r)'$.

Theorem 3 has several important implications. First, BFK's John test is inconsistent in detecting the factor structure when the factors are weak in the sense that $h_j \to d_j \in (0, \infty)$ for $j = 1, ..., r$. Second, BFK's John test still has nontrivial asymptotic power, which is proportional to $\sum_{j=1}^{r} d_j^2$ and inversely proportional to the limit of $\frac{n}{T}$. This result is in sharp contrast with the fixed dimension case in which with fixed magnitude deviation from the null, the asymptotic power tends to one as the sample size tends to infinity. Third, this inconsistency result can also be used to check the extent of cross-sectional dependence due to common factors. If it is reasonable to assume that common factors are the main source of cross-sectional dependence but the power of J_{BFK} is far below one even with large n and large T, then these common factors should be weak.

4.2. Asymptotic Power under the Strong Factor Alternative

Following the same analysis as in Section 4.1, the asymptotic behavior of J_{BFK} under the strong factor alternative is derived in the next theorem.

Theorem 4 *Under Assumptions 2-4, and under the strong factor alternative with* $\frac{h_j}{n} \to d_j \in (0, \infty)$ *for* $j = 1, ..., r$*,*

$$J_{LW} = O_p(nT). \tag{18}$$

Remark 1 *The* $O_p(nT)$ *in this theorem is real, i.e.* $J_{LW} \neq o_p(nT)$*.*

Recall that $J_{LW} = \frac{TU}{2} - \frac{n+1}{2}$, where $U = \frac{1}{n} tr[(\frac{1}{n} trS)^{-1} S - I_n]^2$ measures the distance between the sample covariance matrix and sphericity. With $\frac{h_j}{n} \to d_j \in (0, \infty)$ for $j = 1, ..., r$, as shown in the Appendix, $\frac{1}{n} trS \xrightarrow{p} \sigma^2(1 + \sum_{j=1}^{r} d_j)$ and $\frac{1}{n} trS^2 = O_p(n)$. Hence $U = O_p(n)$ and it follows that $J_{LW} = O_p(nT)$.

Next, we study the asymptotic behavior of $\widehat{J}_{LW} - J_{LW}$ under the strong factor alternative, which as in the weak factor case, depends on the asymptotic behavior of $\frac{1}{n} trS$, $\frac{1}{n} trS^2$, $\frac{1}{n} tr\hat{S} - \frac{1}{n} trS$ and $\frac{1}{n} tr\hat{S}^2 - \frac{1}{n} trS^2$.

Proposition 2 *Under Assumptions 1-4, and under the strong factor alternative with* $\frac{h_j}{n} \to d_j \in (0, \infty)$ *for* $j = 1, ..., r$,

(a) $\frac{1}{n} trS = \sigma^2(1 + \sum_{j=1}^r d_j) + O_p(\frac{1}{\sqrt{T}})$,

(b) $\frac{1}{n} trS^2 = \frac{n(T-1)}{T}\sigma^4[\sum_{j=1}^r d_j^2 - \sum_{i=1}^n(\sum_{j=1}^r d_j e_{i,j}^2)^2] + O_p(\sqrt{n})$,

(c) $\frac{1}{n} tr\hat{S} - \frac{1}{n} trS = O_p(\frac{1}{T})$,

(d) $\frac{1}{n} tr\hat{S}^2 - \frac{1}{n} trS^2 = O_p(\frac{n}{T})$.

Compared to Proposition 1, the stochastic order of part (a) and part (c) remain the same while the stochastic order of part (b) and part (d) are significantly larger. This is because under the strong factor alternative, cross-sectional dependence becomes stronger.

Based on Proposition 2, we have the following theorem.

Theorem 5 *Under Assumptions 1-4, and under the strong factor alternative with* $\frac{h_j}{n} \to d_j \in (0, \infty)$ *for* $j = 1, ..., r$,

$$\widehat{J}_{LW} - J_{LW} = O_p(n). \tag{19}$$

Theorem 5 implies that under the strong factor alternative, the additional noise contained in $\widehat{J}_{LW} - J_{LW}$ is $O_p(n)$. This magnitude is smaller than $O_p(nT)$, the magnitude of J_{LW}, as shown in Theorem 4. Thus $\widehat{J}_{LW} - J_{LW}$ is asymptotically dominated by J_{LW} and this leads us to the consistency of J_{BFK}.

Theorem 6 *Under Assumptions 1-4, and under the strong factor alternative with* $\frac{h_j}{n} \to d_j \in (0, \infty)$ *for* $j = 1, ..., r$, J_{BFK} *is consistent.*

5. CONCLUSION

This paper studies the asymptotic power of BFK's John test for sphericity of the covariance matrix in a fixed effects panel data model under the strong and weak factor alternatives. In the former case, J_{BFK} is consistent, while in the latter case J_{BFK} is inconsistent but has nontrivial asymptotic power. This inconsistency reflects the effect of dimension on the power of statistical tests. From an empirical perspective, the inconsistency also can be used as a model selection scheme to check the extent of cross-sectional dependence resulting from common factors. Several questions are left for future research. First, the normality and no temporal dependence in Assumption 2 are restrictive. Second, for microeconomic applications, one should study the asymptotic power as

$\frac{n}{T} \to \infty$. Third, it would be interesting to study the asymptotic power when the factor is neither strong nor weak in the sense that $\frac{h_j}{n^\delta} \to d_j \in (0, \infty)$ for $0 < \delta < 1$, and when the factors are weak and the number of factors r goes to infinity jointly with n and T.

REFERENCES

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

Bai, Z. (1999). Methodologies in spectral analysis of large dimensional random matrices, a review. *Statistica Sinica* 9:611–677.

Bai, Z., Jiang, D., Yao, J. F., Zheng, S. (2009). Corrections to LRT on large-dimensional covariance matrix by RMT. *Annals of Statistics* 37:3822-3840.

Baltagi, B. H., Feng, Q., Kao, C. (2011). Testing for sphericity in a fixed effects panel data model. *Econometrics Journal* 14:25-47.

Baltagi, B. H., Feng, Q., Kao, C. (2012). A Lagrange Multiplier test for cross-sectional dependence in a fixed effects panel data model. *Journal of Econometrics* 170:164-177.

Baltagi, B. H., Maasoumi, E. (2013). An overview of dependence in cross-section, time-series and panel data. *Econometric Reviews* 32:543-546.

Breusch, T. S., Pagan, A. R. (1980). The Lagrange multiplier test and its applications to model specification in econometrics. *Review of Economic Studies* 47:239-253.

Cai, T. T., Ma, Z. (2013). Optimal hypothesis testing for high dimensional covariance matrices. *Bernoulli* 19:2359-2388.

Chen, S. X., Zhang, L. X., Zhong, P. S. (2010). Tests for high-dimensional covariance matrices. *Journal of the American Statistical Association* 105:810-819.

Chudik, A., Pesaran, M. H. (2013). Large panel data models with cross-sectional dependence: a survey (No. 4371). CESifo Working Paper.

Dufour, J. M., Khalaf, L. (2002). Exact tests for contemporaneous correlation of disturbances in seemingly unrelated regressions. *Journal of Econometrics* 106:143-170.

Frees, E. W. (1995). Assessing cross-sectional correlation in panel data. *Journal of Econometrics* 69:393-414.

Geman, S. (1980). A limit theorem for the norm of random matrices. *Annals of Probability* 8:252–261.

John, S. (1971). Some optimal multivariate tests. *Biometrika* 58:123-127.

John, S. (1972). The distribution of a statistic used for testing sphericity of normal distributions. *Biometrika* 59:169-173.

Johnstone, I. M. (2001). On the distribution of the largest principal component. *Annals of Statistics* 29:295–327.

Johnstone, I. M., Lu, A. Y. (2009). On consistency and sparsity for principal components analysis in high dimensions. *Journal of the American Statistical Association* 104:682-693.

Ledoit, O., Wolf, M. (2002). Some hypothesis tests for the covariance matrix when the dimension is large compared to the sample size. *Annals of Statistics* 30:1081-1102.

Onatski, A. (2012). Asymptotics of the principal components estimator of large factor models with weakly influential factors. *Journal of Econometrics* 168:244-258.

Onatski, A., Moreira, M. J., Hallin, M. (2013). Asymptotic power of sphericity tests for high-dimensional data. *Annals of Statistics*, 41:1204-1231.

Onatski, A., Moreira, M. J., Hallin, M. (2014). Signal detection in high dimension: The multispiked case. *Annals of Statistics* 42:225-254.

Pesaran, M. H. (2004). General diagnostic tests for cross section dependence in panels. Working Paper, University of Cambridge&USC.

Pesaran, M. H. (2012). Testing weak cross-sectional dependence in large panels, CESifo working paper 3800.

Pesaran, M. H., Ullah, A., Yamagata, T. (2008). A bias-adjusted LM test of error cross-section independence. *Econometrics Journal* 11:105-127.

Phillips, P. C. B., Moon, H. R. (1999). Linear regression limit theory for nonstationary panel data. *Econometrica* 67: 1057-1111.

Srivastava, M. S. (2005). Some tests concerning the covariance matrix in high dimensional data. *Journal of the Japan Statistical Society* 35:251-272.

Staiger, D., Stock, J. H. (1997). Instrumental variables regression with weak instruments. *Econometrica* 65: 557-586.

Wang, C., Cao, L., Miao, B. (2013). Asymptotic power of likelihood ratio tests for high dimensional data. arXiv preprint arXiv:1302.3302.

APPENDIX

Lemma 1 *Suppose X_n is a sequence of random variables and $EX_n^2 = O(n^v)$,where v is a constant, then $X_n = O_p(n^{v/2})$.*

Lemma 1 will be used repeatedly in calculating the stochastic order of the cross product of error terms in this appendix.

Lemma 2 *Suppose $\nu \sim N(0, \Sigma_n)$, and let a_{sh} be the typical element of the covariance matrix in the s-th row and h-th column. Then for r, s, h, q,*

(1) $E\nu_s = 0$,

(2) $E\nu_s\nu_h = a_{sh}$,

(3) $E\nu_r\nu_s\nu_h = 0$,

(4) $E\nu_r^2\nu_s\nu_h = 2a_{sr}a_{hr} + a_{rr}a_{sh}$,

(5) $E\nu_s^2\nu_h^2 = a_{ss}a_{hh} + 2a_{sh}^2$,

(6) $E\nu_r\nu_s\nu_h\nu_q = a_{sr}a_{hq} + a_{sq}a_{hr} + a_{sh}a_{rq}$,

(7) $E\nu_r\nu_s\nu_h\nu_p\nu_q = 0$,

(8) $E\nu_s^3\nu_h^3 = 9a_{ss}a_{hh}a_{sh} + 6a_{sh}^3$.

Lemma 2 will be used repeatedly in dealing with cross-sectional dependence under the alternative hypothesis.

Lemma 3 *Define $A_0 = \bar{\nu}.\bar{\nu}'.$, $A_1 = \frac{1}{T}\sum_{t=1}^T \tilde{x}_t(\tilde{\beta} - \beta)\tilde{\nu}_t'$, $A_2 = A_1' = \frac{1}{T}\sum_{t=1}^T \tilde{\nu}_t(\tilde{\beta} - \beta)'\tilde{x}_t'$, $A_3 = \frac{1}{T}\sum_{t=1}^T \tilde{x}_t(\tilde{\beta} - \beta)(\tilde{\beta} - \beta)'\tilde{x}_t'$, and hence $\hat{S} - S = -A_0 - A_1 - A_2 + A_3$.*

Under the weak factor alternative, we have

(a) $\frac{1}{n}tr(SA_1) = O_p(\frac{1}{T^2}) + O_p(\frac{1}{nT}) + O_p(\frac{1}{T\sqrt{nT}})$,

(b) $\frac{1}{n}tr(SA_3) = O_p(\frac{1}{nT})$,

(c) $\frac{1}{n}tr(A_1^2) = O_p(\frac{1}{nT^2})$,

(d) $\frac{1}{n}tr(A_1A_2) = O_p(\frac{1}{T^2})$,

(e) $\frac{1}{n}tr(A_1A_3) = O_p(\frac{1}{nT^2})$,

(f) $\frac{1}{n}tr(A_3^2) = O_p(\frac{1}{nT^2})$,

(g) $\frac{1}{n}tr(SA_0) = \frac{1}{T}\sigma^4 + \frac{n}{T^2}\sigma^4 + O_p(\frac{1}{T\sqrt{T}})$,

(h) $\frac{1}{n}tr(A_0^2) = \frac{n}{T^2}\sigma^4 + O_p(\frac{\sqrt{n}}{T^2})$,

(i) $\frac{1}{n}tr(A_0A_1) = O_p(\frac{1}{T^2})$,

(j) $\frac{1}{n}tr(A_0A_3) = O_p(\frac{1}{nT^2})$.

Under the strong factor alternative, we have

(a) $\frac{1}{n}tr(SA_1) = O_p(\frac{\sqrt{n}}{T})$,

(b) $\frac{1}{n}tr(SA_3) = O_p(\frac{1}{T})$,

(c) $\frac{1}{n}tr(A_1^2) = O_p(\frac{1}{T^2})$,

(d) $\frac{1}{n}tr(A_1A_2) = O_p(\frac{1}{T^2})$,

(e) $\frac{1}{n}tr(A_1A_3) = O_p(\frac{1}{\sqrt{n}T^2})$,

(f) $\frac{1}{n}tr(A_3^2) = O_p(\frac{1}{nT^2})$,

(g) $\frac{1}{n}tr(SA_0) = O_p(\frac{n}{T})$,

(h) $\frac{1}{n}tr(A_0^2) = O_p(\frac{n}{T^2})$,

(i) $\frac{1}{n}tr(A_0A_1) = O_p(\frac{\sqrt{n}}{T^2})$,

(j) $\frac{1}{n}tr(A_0A_3) = O_p(\frac{1}{T^2})$.

This lemma can be proved following the same line of proof as Lemma 3 in the supplementary appendix of Baltagi, Feng and Kao (2011).

A Proof of Theorem 1

Proof. The proof of this theorem is based on Theorem 3.1 of Srivastava (2005). After some notation translation, Srivastava's Theorem 3.1 is equivalent to

$$\frac{T}{2}(\hat{\gamma}_1 - \gamma_1) \xrightarrow{d} N(0, \tau_1^2)$$

provided $T = O(n^\delta), 0 < \delta \leq 1$, and $\frac{tr\Sigma_n^i}{n} \to a_i < \infty$ for $i = 1, ..., 8$, where

$$\gamma_1 = \frac{tr\Sigma_n^2/n}{(tr\Sigma_n/n)^2},$$

$$\tau_1^2 = \frac{2T(a_4a_1^2 - 2a_1a_2a_3 + a_2^3)}{na_1^6} + \frac{a_2^2}{a_1^4},$$

and

$$\hat{\gamma}_1 = \frac{T^2}{(T-1)(T+2)}\left[trS^2/n - \frac{n}{T}(trS/n)^2\right]/(trS/n)^2.$$

Under the current setup with $\frac{n}{T} \to c \in (0, \infty)$ and $h_j \to d_j \in (0, \infty)$ for $j = 1, ..., r$, the two conditions of Srivastava's Theorem 3.1 are satisfied. Hence

$$\frac{T}{2}(\hat{\gamma}_1 - \gamma_1) = \frac{T^2}{(T-1)(T+2)}(J_{LW} + \frac{1}{T} - \frac{(T-1)(T+2)}{2T}\left(\frac{tr\Sigma_n^2/n}{(tr\Sigma_n/n)^2} - 1\right)\Bigg($$

$$\frac{1}{T} - \frac{(T-1)(T+2)}{2T}\left(\frac{tr\Sigma_n^2/n}{(tr\Sigma_n/n)^2} - 1\right) \approx -\frac{T\sum_{j=1}^r d_j^2}{2n} \to -\frac{\sum_{j=1}^r d_j^2}{2c},$$

and

$$\tau_1^2 \to 1.$$

Therefore,

$$J_{LW} \xrightarrow{d} N\left(\frac{\sum_{j=1}^r d_j^2}{2c}, 1\right)$$

■

B Proof of Proposition 1

Proof of part (a). For notation simplicity, we will give the proof for the case where $r = 1$. Using $(\sum_{i=1}^r x_i)^2 \leq r\sum_{i=1}^r x_i^2$ repeatedly, the case where $r > 1$ can be proved similarly, as long as r is fixed. Note that

$$\begin{aligned}
\frac{1}{n}trS &= \frac{1}{n}tr[\frac{1}{T}\sum_{t=1}^T \nu_t\nu_t'] \\
&= \frac{1}{nT}\sum_{t=1}^T\sum_{i=1}^n \nu_{it}^2 \\
&= \sigma^2(1+\frac{h}{n}) + \frac{1}{nT}\sum_{t=1}^T\sum_{i=1}^n(\nu_{it}^2 - E\nu_{it}^2) \\
&= \sigma^2(1+\frac{h}{n}) + O_p(\frac{1}{\sqrt{nT}}) \\
&= \sigma^2 + O_p(\frac{1}{\sqrt{nT}}),
\end{aligned}$$

since

$$\begin{aligned}
& E[\frac{1}{nT}\sum_{t=1}^T\sum_{i=1}^n(\nu_{it}^2 - E\nu_{it}^2)]^2 \\
= {} & \frac{1}{n^2T^2}\sum_{t=1}^T\sum_{i=1}^n\sum_{j=1}^n E(\nu_{it}^2 - E\nu_{it}^2)(\nu_{jt}^2 - E\nu_{jt}^2) \\
= {} & \frac{1}{n^2T^2}[\sum_{t=1}^T\sum_{i=1}^n E(\nu_{it}^2 - E\nu_{it}^2)^2 + \sum_{t=1}^T\sum_{i=1}^n\sum_{j\neq i}^n E(\nu_{it}^2 - E\nu_{it}^2)(\nu_{jt}^2 - E\nu_{jt}^2)] \\
= {} & \frac{1}{n^2T^2}[\sum_{t=1}^T\sum_{i=1}^n 2(\sigma^2 + \sigma^2 he_i^2)^2 + \sum_{t=1}^T\sum_{i=1}^n\sum_{j\neq i}^n(E\nu_{it}^2\nu_{jt}^2 - E\nu_{it}^2E\nu_{jt}^2)] \\
= {} & \frac{1}{n^2T^2}[\sum_{t=1}^T\sum_{i=1}^n 2(\sigma^2 + \sigma^2 he_i^2)^2 + \sum_{t=1}^T\sum_{i=1}^n\sum_{j\neq i}^n 2\sigma^4(he_ie_j)^2] \\
= {} & \frac{1}{n^2T^2}[2T\sum_{i=1}^n(\sigma^4 + 2h\sigma^4e_i^2 + \sigma^4h^2e_i^4) + 2\sigma^4Th(1 - \sum_{i=1}^n e_i^4)] \\
= {} & \frac{1}{n^2T^2}(2Tn\sigma^4 + 4Th\sigma^4 + 2\sigma^4Th^2) \\
= {} & O(\frac{1}{nT}).
\end{aligned}$$

This uses $\sum_{i=1}^{n} e_i^2 = 1$ and $E\nu_s^2\nu_h^2 = a_{ss}a_{hh} + 2a_{sh}^2$. ■

Proof of part (b). Note that

$$\begin{aligned}
& \frac{1}{n}trS^2 \\
= \; & \frac{1}{n}trS^2[(\frac{1}{T}\sum_{t=1}^{T}\nu_t\nu_t')(\frac{1}{T}\sum_{s=1}^{T}\nu_s\nu_s')] = \frac{1}{nT^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\nu_t'\nu_s\nu_s'\nu_t \\
= \; & \frac{1}{nT^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n}\nu_{it}^2\nu_{jt}^2 + \frac{1}{nT^2}\sum_{t=1}^{T}\sum_{s\neq t}^{T}\sum_{i=1}^{n}\sum_{j=1}^{n}\nu_{it}\nu_{is}\nu_{js}\nu_{jt} \\
= \; & \frac{1}{nT^2}\sum_{t=1}^{T}\sum_{i=1}^{n}\nu_{it}^4 + \frac{1}{nT^2}\sum_{t=1}^{T}\sum_{j\neq i}^{n}\sum_{i=1}^{n}\nu_{it}^2\nu_{jt}^2 \\
& + \frac{1}{nT^2}\sum_{t=1}^{T}\sum_{s\neq t}^{T}\sum_{i=1}^{n}\nu_{it}^2\nu_{is}^2 + \frac{1}{nT^2}\sum_{t=1}^{T}\sum_{s\neq t}^{T}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\nu_{it}\nu_{is}\nu_{js}\nu_{jt} \\
= \; & O_p(\frac{1}{T}) + [\frac{n}{T}\sigma^4 + O_p(\frac{1}{\sqrt{T}})] + [\sigma^4 + O_p(\frac{1}{\sqrt{T}})] + O_p(\frac{1}{T}) \\
= \; & (\frac{n}{T}+1)\sigma^4 + O_p(\frac{1}{\sqrt{T}}).
\end{aligned}$$

This uses the following four results:

(1) $\frac{1}{nT^2}\sum_{t=1}^{T}\sum_{i=1}^{n}\nu_{it}^4 = \frac{1}{nT^2}\sum_{t=1}^{T}\sum_{i=1}^{n}E\nu_{it}^4 + \frac{1}{nT^2}\sum_{t=1}^{T}\sum_{i=1}^{n}(\nu_{it}^4 - E\nu_{it}^4) = O_p(\frac{1}{T})$, since

$$\begin{aligned}
& E[\frac{1}{nT^2}\sum_{t=1}^{T}\sum_{i=1}^{n}(\nu_{it}^4 - E\nu_{it}^4)]^2 \\
= \; & \frac{1}{n^2T^4}\sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{s=1}^{T}\sum_{j=1}^{n}E(\nu_{it}^4 - E\nu_{it}^4)(\nu_{js}^4 - E\nu_{js}^4) \\
= \; & \frac{1}{n^2T^4}O_p(n^2T^2) = O_p(\frac{1}{T^2}).
\end{aligned}$$

(2)

$$\begin{aligned}
& \frac{1}{nT^2}\sum_{t=1}^{T}\sum_{j\neq i}^{n}\sum_{i=1}^{n}\nu_{it}^2\nu_{jt}^2 \\
= \; & \frac{1}{nT^2}\sum_{t=1}^{T}\sum_{j\neq i}^{n}\sum_{i=1}^{n}E\nu_{it}^2E\nu_{jt}^2 + \frac{1}{nT^2}\sum_{t=1}^{T}\sum_{j\neq i}^{n}\sum_{i=1}^{n}(\nu_{it}^2\nu_{jt}^2 - E\nu_{it}^2E\nu_{jt}^2) \\
= \; & \frac{1}{nT^2}\sum_{t=1}^{T}\sum_{j\neq i}^{n}\sum_{i=1}^{n}(\sigma^2 + \sigma^2he_i^2)(\sigma^2 + \sigma^2he_j^2) \\
& + \frac{1}{nT^2}\sum_{t=1}^{T}\sum_{j\neq i}^{n}\sum_{i=1}^{n}(\nu_{it}^2\nu_{jt}^2 - E\nu_{it}^2E\nu_{jt}^2) \\
= \; & \frac{n-1}{T}\sigma^4 + \frac{n-1}{nT}2h\sigma^4 + \frac{h^2(1-\sum_{i=1}^{n}e_i^4)}{nT}\sigma^4 + \frac{1}{nT^2}\sum_{t=1}^{T}\sum_{j\neq i}^{n}\sum_{i=1}^{n}(\nu_{it}^2\nu_{jt}^2 - E\nu_{it}^2E\nu_{jt}^2) \\
= \; & \frac{n-1}{T}\sigma^4 + \frac{n-1}{nT}2h\sigma^4 + \frac{h^2(1-\sum_{i=1}^{n}e_i^4)}{nT}\sigma^4 + O_p(\frac{1}{\sqrt{T}}) \\
= \; & \frac{n}{T}\sigma^4 + O_p(\frac{1}{\sqrt{T}}),
\end{aligned}$$

since

$$
\begin{aligned}
& E[\frac{1}{nT^2}\sum\nolimits_{t=1}^{T}\sum\nolimits_{j\neq i}^{n}\sum\nolimits_{i=1}^{n}(\nu_{it}^2\nu_{jt}^2-E\nu_{it}^2E\nu_{jt}^2)]^2 \\
= & \frac{1}{n^2T^4}\sum\nolimits_{t_1=1}^{T}\sum\nolimits_{j_1\neq i_1}^{n}\sum\nolimits_{i_1=1}^{n}\sum\nolimits_{t_2=1}^{T}\sum\nolimits_{j_2\neq i_2}^{n}\sum\nolimits_{i_2=1}^{n}E(\nu_{i_1t_1}^2\nu_{j_1t_1}^2 \\
& -E\nu_{i_1t_1}^2E\nu_{j_1t_1}^2)(\nu_{i_2t_2}^2\nu_{j_2t_2}^2-E\nu_{i_2t_2}^2E\nu_{j_2t_2}^2) \\
= & \frac{1}{n^2T^4}[E(1,\cdot)+E\,(2,\cdot)]=\frac{1}{n^2T^4}[O(n^4T)+O(n^3T^2)] \\
= & O(\frac{n^2}{T^3})+O(\frac{n}{T^2})=O(\frac{1}{T}).
\end{aligned}
$$

Here we used $\frac{n}{T}\to c\in(0,\infty)$ and $E\,(2,\cdot)=E\,(2,4)+E(2,j<4)=O(n^3T^2)$. Hereafter $E(i,j)$ denotes there are i different t-indices and j different n-indices in the summation. By using $E\nu_s^2\nu_h^2=a_{ss}a_{hh}+2a_{sh}^2$,

$$
\begin{aligned}
& E\,(2,4) \\
= & \sum\nolimits_{t_1=1}^{T}\sum\nolimits_{j_1\neq i_1}^{n}\sum\nolimits_{i_1=1}^{n}\sum\nolimits_{t_2=1}^{T}\sum\nolimits_{j_2\neq i_2}^{n}\sum\nolimits_{i_2=1}^{n}E(\nu_{i_1t_1}^2\nu_{j_1t_1}^2 \\
& -E\nu_{i_1t_1}^2E\nu_{j_1t_1}^2)E(\nu_{i_2t_2}^2\nu_{j_2t_2}^2-E\nu_{i_2t_2}^2E\nu_{j_2t_2}^2) \\
= & \sum\nolimits_{t_1=1}^{T}\sum\nolimits_{j_1\neq i_1}^{n}\sum\nolimits_{i_1=1}^{n}\sum\nolimits_{t_2=1}^{T}\sum\nolimits_{j_2\neq i_2}^{n}\sum\nolimits_{i_2=1}^{n}\sigma^8(2h^2e_{i_1}^2e_{j_1}^2)(2h^2e_{i_2}^2e_{j_2}^2) \\
= & 4\sigma^8h^4\sum\nolimits_{t_1=1}^{T}\sum\nolimits_{j_1\neq i_1}^{n}\sum\nolimits_{i_1=1}^{n}\sum\nolimits_{t_2=1}^{T}\sum\nolimits_{j_2\neq i_2}^{n}\sum\nolimits_{i_2=1}^{n}e_{i_1}^2e_{j_1}^2e_{i_2}^2e_{j_2}^2 \\
\leq & 4\sigma^8h^4T^2=O(T^2).
\end{aligned}
$$

There are at most n^3T^2 terms in $E(2,j<4)$, hence $E(2,j<4)=O(n^3T^2)$. Combining these results, we have $E\,(2,\cdot)=O(T^2)+O(n^3T^2)=O(n^3T^2)$.

(3)

$$
\begin{aligned}
& \frac{1}{nT^2}\sum\nolimits_{t=1}^{T}\sum\nolimits_{s\neq t}^{T}\sum\nolimits_{i=1}^{n}\nu_{it}^2\nu_{is}^2 \\
= & \frac{1}{nT^2}\sum\nolimits_{t=1}^{T}\sum\nolimits_{s\neq t}^{T}\sum\nolimits_{i=1}^{n}E\nu_{it}^2E\nu_{is}^2+\frac{1}{nT^2}\sum\nolimits_{t=1}^{T}\sum\nolimits_{s\neq t}^{T}\sum\nolimits_{i=1}^{n}(\nu_{it}^2\nu_{is}^2-E\nu_{it}^2E\nu_{is}^2) \\
= & \frac{1}{nT^2}\sum\nolimits_{t=1}^{T}\sum\nolimits_{s\neq t}^{T}\sum\nolimits_{i=1}^{n}E\nu_{it}^2E\nu_{is}^2+\frac{1}{nT^2}\sum\nolimits_{t=1}^{T}\sum\nolimits_{s\neq t}^{T}\sum\nolimits_{i=1}^{n}(\nu_{it}^2\nu_{is}^2-E\nu_{it}^2E\nu_{is}^2) \\
= & \frac{1}{nT^2}\sum\nolimits_{t=1}^{T}\sum\nolimits_{s\neq t}^{T}\sum\nolimits_{i=1}^{n}(\sigma^2+\sigma^2he_i^2)^2+\frac{1}{nT^2}\sum\nolimits_{t=1}^{T}\sum\nolimits_{s\neq t}^{T}\sum\nolimits_{i=1}^{n}(\nu_{it}^2\nu_{is}^2-E\nu_{it}^2E\nu_{is}^2) \\
= & (\frac{T-1}{T}\sigma^4+\frac{T-1}{nT}2\sigma^4h+\frac{T-1}{nT}\sigma^4h^2\sum\nolimits_{i=1}^{n}e_i^4) \\
& +\frac{1}{nT^2}\sum\nolimits_{t=1}^{T}\sum\nolimits_{s\neq t}^{T}\sum\nolimits_{i=1}^{n}(\nu_{it}^2\nu_{is}^2-E\nu_{it}^2E\nu_{is}^2) \\
= & [\frac{T-1}{T}\sigma^4+O(\frac{1}{n})]+[O_p(\frac{1}{\sqrt{T}})]=\sigma^4+O_p(\frac{1}{\sqrt{T}})
\end{aligned}
$$

since

$$
\begin{aligned}
& E[\frac{1}{nT^2}\sum\nolimits_{t=1}^{T}\sum\nolimits_{s\neq t}^{T}\sum\nolimits_{i=1}^{n}(\nu_{it}^2\nu_{is}^2-E\nu_{it}^2E\nu_{is}^2)]^2 \\
= & \frac{1}{n^2T^4}\sum\nolimits_{t_1=1}^{T}\sum\nolimits_{s_1\neq t_1}^{T}\sum\nolimits_{i_1=1}^{n}\sum\nolimits_{t_2=1}^{T}\sum\nolimits_{s_2\neq t_2}^{T}\sum\nolimits_{i_2=1}^{n}E(\nu_{i_1t_1}^2\nu_{i_1s_1}^2 \\
& -E\nu_{i_1t_1}^2E\nu_{i_1s_1}^2)(\nu_{i_2t_2}^2\nu_{i_2s_2}^2-E\nu_{i_2t_2}^2E\nu_{i_2s_2}^2) \\
= & \frac{1}{n^2T^4}O(n^2T^3)=O(\frac{1}{T}).
\end{aligned}
$$

When s_1,s_2,t_1,t_2 are different from each other, we have

$$
E(\nu_{i_1t_1}^2\nu_{i_1s_1}^2-E\nu_{i_1t_1}^2E\nu_{i_1s_1}^2)(\nu_{i_2t_2}^2\nu_{i_2s_2}^2-E\nu_{i_2t_2}^2E\nu_{i_2s_2}^2)=0.
$$

(4) $\frac{1}{nT^2}\sum_{t=1}^{T}\sum_{s\neq t}^{T}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\nu_{it}\nu_{is}\nu_{js}\nu_{jt}=O_p(\frac{1}{T})$. This is because

$$
\begin{aligned}
& E[\frac{1}{nT^2}\sum\nolimits_{t=1}^{T}\sum\nolimits_{s\neq t}^{T}\sum\nolimits_{i=1}^{n}\sum\nolimits_{j\neq i}^{n}\nu_{it}\nu_{is}\nu_{js}\nu_{jt}]^2 \\
= & \frac{1}{n^2T^4}\sum\nolimits_{t_1=1}^{T}\sum\nolimits_{s_1\neq t_1}^{T}\sum\nolimits_{i_1=1}^{n}\sum\nolimits_{j_1\neq i_1}^{n}\sum\nolimits_{t_2=1}^{T}\sum\nolimits_{s_2\neq t_2}^{T}\sum\nolimits_{i_2=1}^{n}\sum\nolimits_{j_2\neq i_2}^{n}E(\nu_{i_1t_1} \\
& \nu_{i_1s_1}\nu_{j_1s_1}\nu_{j_1t_1}\nu_{i_2t_2}\nu_{i_2s_2}\nu_{j_2s_2}\nu_{j_2t_2}) \\
= & \frac{1}{n^2T^4}[E(4,4)+E(4,3)+E(4,2)+E(3,4)+ \\
& E(3,3)+E(3,2)+E(2,4)+E(2,3)+E(2,2)] \\
= & \frac{1}{n^2T^4}[O(T^4)+O(T^4)+O(T^4)+O(T^3) \\
& +O(T^3\sqrt{n})+O(T^3)+O(T^2)+O(T^2n)+O(T^2n^2)] \\
= & O(\frac{1}{n^2})+O(\frac{1}{n^{\frac{3}{2}}T})+O(\frac{1}{T^2})=O(\frac{1}{T^2}).
\end{aligned}
$$

The above calculation is based on the following results.

$$
\begin{aligned}
E(4,4) &= \sum\nolimits_{t_1=1}^{T}\sum\nolimits_{s_1\neq t_1}^{T}\sum\nolimits_{i_1=1}^{n}\sum\nolimits_{j_1\neq i_1}^{n}\sum\nolimits_{t_2=1}^{T}\sum\nolimits_{s_2\neq t_2}^{T}\sum\nolimits_{i_2=1}^{n}\sum\nolimits_{j_2\neq i_2}^{n}\sigma^8(he_{i_1}e_{j_1})^2(he_{i_2}e_{j_2})^2 \\
&= \sigma^8h^4\sum\nolimits_{t_1=1}^{T}\sum\nolimits_{s_1\neq t_1}^{T}\sum\nolimits_{i_1=1}^{n}\sum\nolimits_{j_1\neq i_1}^{n}\sum\nolimits_{t_2=1}^{T}\sum\nolimits_{s_2\neq t_2}^{T}\sum\nolimits_{i_2=1}^{n}\sum\nolimits_{j_2\neq i_2}^{n}e_{i_1}^2e_{j_1}^2e_{i_2}^2e_{j_2}^2 \\
&\leq \sigma^8h^4T^4=O(T^4).
\end{aligned}
$$

$$
\begin{aligned}
E(4,3) &= \sum\nolimits_{t_1=1}^{T}\sum\nolimits_{s_1\neq t_1}^{T}\sum\nolimits_{i_1=1}^{n}\sum\nolimits_{j_1\neq i_1}^{n}\sum\nolimits_{t_2=1}^{T}\sum\nolimits_{s_2\neq t_2}^{T}\sum\nolimits_{j_2\neq j_1}^{n}\sigma^8(he_{i_1}e_{j_1})^2(he_{j_1}e_{j_2})^2 \\
&= \sigma^8h^4\sum\nolimits_{t_1=1}^{T}\sum\nolimits_{s_1\neq t_1}^{T}\sum\nolimits_{i_1=1}^{n}\sum\nolimits_{j_1\neq i_1}^{n}\sum\nolimits_{t_2=1}^{T}\sum\nolimits_{s_2\neq t_2}^{T}\sum\nolimits_{j_2\neq j_1}^{n}e_{i_1}^2e_{j_1}^4e_{j_2}^2 \\
&\leq \sigma^8h^4T^4=O(T^4).
\end{aligned}
$$

$$
\begin{aligned}
E(4,2) &= \sum\nolimits_{t_1=1}^{T}\sum\nolimits_{s_1\neq t_1}^{T}\sum\nolimits_{i_1=1}^{n}\sum\nolimits_{j_1\neq i_1}^{n}\sum\nolimits_{t_2=1}^{T}\sum\nolimits_{s_2\neq t_2}^{T}\sigma^8(he_{i_1}e_{j_1})^4 \\
&= \sigma^8h^4\sum\nolimits_{t_1=1}^{T}\sum\nolimits_{s_1\neq t_1}^{T}\sum\nolimits_{i_1=1}^{n}\sum\nolimits_{j_1\neq i_1}^{n}\sum\nolimits_{t_2=1}^{T}\sum\nolimits_{s_2\neq t_2}^{T}e_{i_1}^4e_{j_1}^4 \\
&\leq \sigma^8h^4T^4=O(T^4).
\end{aligned}
$$

$$
\begin{aligned}
E(3,4) &= \sum_{t_1=1}^{T}\sum_{s_1\neq t_1}^{T}\sum_{i_1=1}^{n}\sum_{j_1\neq i_1}^{n}\sum_{t_1\neq t_2}^{T}\sum_{i_2=1}^{n}\sum_{j_2\neq i_2}^{n} E(\nu_{i_1 s_1}\\
&\quad \nu_{j_1 s_1})E(\nu_{i_1 t_1}\nu_{j_1 t_1}\nu_{i_2 t_1}\nu_{j_2 t_1})E(\nu_{i_2 t_2}\nu_{j_2 t_2})\\
&= \sum_{t_1=1}^{T}\sum_{s_1\neq t_1}^{T}\sum_{i_1=1}^{n}\sum_{j_1\neq i_1}^{n}\sum_{t_1\neq t_2}^{T}\sum_{i_2=1}^{n}\sum_{j_2\neq i_2}^{n} (\sigma^2 h e_{i_1} e_{j_1})[(\sigma^2 h e_{i_1} e_{j_1})\\
&\quad (\sigma^2 h e_{i_2} e_{j_2}) + (\sigma^2 h e_{i_1} e_{i_2})(\sigma^2 h e_{j_1} e_{j_2}) + (\sigma^2 h e_{i_1} e_{j_2})(\sigma^2 h e_{i_2} e_{j_1})](\sigma^2 h e_{i_2} e_{j_2})\\
&= 3\sigma^8 h^4 \sum_{t_1=1}^{T}\sum_{s_1\neq t_1}^{T}\sum_{i_1=1}^{n}\sum_{j_1\neq i_1}^{n}\sum_{t_1\neq t_2}^{T}\sum_{i_2=1}^{n}\sum_{j_2\neq i_2}^{n} e_{i_1}^2 e_{j_1}^2 e_{i_2}^2 e_{j_2}^2\\
&\leq 3\sigma^8 h^4 T^3 = O(T^3),
\end{aligned}
$$

with $E\nu_r\nu_s\nu_h\nu_q = a_{sr}a_{hq} + a_{sq}a_{hr} + a_{sh}a_{rq}$.

$$
\begin{aligned}
E(3,3) &= \sum_{t_1=1}^{T}\sum_{s_1\neq t_1}^{T}\sum_{i_1=1}^{n}\sum_{j_1\neq i_1}^{n}\sum_{t_1\neq t_2}^{T}\sum_{j_2\neq j_1}^{n} E\nu_{i_1 s_1}\nu_{j_1 s_1}E\nu_{i_1 t_1}\nu_{j_1 t_1}^2\nu_{j_2 t_1}E\nu_{j_1 t_2}\nu_{j_2 t_2}\\
&= \sum_{t_1=1}^{T}\sum_{s_1\neq t_1}^{T}\sum_{i_1=1}^{n}\sum_{j_1\neq i_1}^{n}\sum_{t_1\neq t_2}^{T}\sum_{j_2\neq j_1}^{n} (\sigma^2 h e_{i_1} e_{j_1})[(\sigma^2 + \sigma^2 h e_{j_1}^2)(\sigma^2 h e_{i_1} e_{j_2})\\
&\quad +2(\sigma^2 h e_{i_1} e_{j_1})(\sigma^2 h e_{j_1} e_{j_2})](\sigma^2 h e_{j_1} e_{j_2})\\
&= \sum_{t_1=1}^{T}\sum_{s_1\neq t_1}^{T}\sum_{i_1=1}^{n}\sum_{j_1\neq i_1}^{n}\sum_{t_1\neq t_2}^{T}\sum_{j_2\neq j_1}^{n} (3\sigma^8 h^4 e_{i_1}^2 e_{j_1}^4 e_{j_2}^2 + \sigma^8 h^3 e_{i_1} e_{j_1}^3 e_{j_2}^2),
\end{aligned}
$$

with $E\nu_r^2\nu_s\nu_h = 2a_{sr}a_{hr} + a_{rr}a_{sh}$. Hence,

$$
\begin{aligned}
|E(3,3)| &\leq \sum_{t_1=1}^{T}\sum_{s_1\neq t_1}^{T}\sum_{i_1=1}^{n}\sum_{j_1\neq i_1}^{n}\sum_{t_1\neq t_2}^{T}\sum_{j_2\neq j_1}^{n} 3\sigma^8 h^4 e_{i_1}^2 e_{j_1}^4 e_{j_2}^2\\
&\quad +\sum_{t_1=1}^{T}\sum_{s_1\neq t_1}^{T}\sum_{i_1=1}^{n}\sum_{j_1\neq i_1}^{n}\sum_{t_1\neq t_2}^{T}\sum_{j_2\neq j_1}^{n} \sigma^8 h^3 e_{i_1}\left|e_{j_1}^3\right| e_{j_2}^2\\
&\leq 3\sigma^8 h^4 T^3 + \sigma^8 h^3 T^3\sqrt{n} = O(T^3\sqrt{n}).
\end{aligned}
$$

$$
\begin{aligned}
E(3,2) &= \sum_{t_1=1}^{T}\sum_{s_1\neq t_1}^{T}\sum_{t_1\neq t_2}^{T}\sum_{i=1}^{n}\sum_{j\neq i}^{n} E\nu_{i s_1}\nu_{j s_1}E\nu_{i t_1}^2\nu_{j t_1}^2 E\nu_{i t_2}\nu_{j t_2}\\
&= \sum_{t_1=1}^{T}\sum_{s_1\neq t_1}^{T}\sum_{t_1\neq t_2}^{T}\sum_{i=1}^{n}\sum_{j\neq i}^{n}(\sigma^2 h e_i e_j)^2[(\sigma^2 + \sigma^2 h e_i^2)(\sigma^2 + \sigma^2 h e_j^2)\\
&\quad +2(\sigma^2 h e_i e_j)^2]\\
&= \sum_{t_1=1}^{T}\sum_{s_1\neq t_1}^{T}\sum_{t_1\neq t_2}^{T}\sum_{i=1}^{n}\sum_{j\neq i}^{n}(\sigma^8 h^2 e_i^2 e_j^2 + \sigma^8 h^3 e_i^4 e_j^2 + \sigma^8 h^3 e_i^2 e_j^4 + 3\sigma^8 h^4 e_i^4 e_j^4)\\
&\leq \sigma^8 h^2 T^3 + 2\sigma^8 h^3 T^3 + 3\sigma^8 h^4 T^3 = O(T^3),
\end{aligned}
$$

with $E\nu_s^2\nu_h^2 = a_{ss}a_{hh} + 2a_{sh}^2$.

$$
\begin{aligned}
E(2,4) &= 2\sum_{t=1}^{T}\sum_{s\neq t}^{T}\sum_{i_1=1}^{n}\sum_{j_1\neq i_1}^{n}\sum_{i_2=1}^{n}\sum_{j_2\neq i_2}^{n} E\nu_{i_1 s}\nu_{j_1 s}\nu_{i_2 s}\nu_{j_2 s}E\nu_{i_1 t}\nu_{j_1 t}\nu_{i_2 t}\nu_{j_2 t}\\
&= 2\sum_{t=1}^{T}\sum_{s\neq t}^{T}\sum_{i_1=1}^{n}\sum_{j_1\neq i_1}^{n}\sum_{i_2=1}^{n}\sum_{j_2\neq i_2}^{n}[(\sigma^2 h e_{i_1} e_{j_1})(\sigma^2 h e_{i_2} e_{j_2})\\
&\quad +(\sigma^2 h e_{i_1} e_{i_2})(\sigma^2 h e_{j_1} e_{j_2}) + (\sigma^2 h e_{i_1} e_{j_2})(\sigma^2 h e_{i_2} e_{j_1})]^2\\
&= 18\sigma^8 h^4 \sum_{t=1}^{T}\sum_{s\neq t}^{T}\sum_{i_1=1}^{n}\sum_{j_1\neq i_1}^{n}\sum_{i_2=1}^{n}\sum_{j_2\neq i_2}^{n} e_{i_1}^2 e_{j_1}^2 e_{i_2}^2 e_{j_2}^2\\
&\leq 18\sigma^8 h^4 T^2 = O(T^2),
\end{aligned}
$$

with $E\nu_r\nu_s\nu_h\nu_q = a_{sr}a_{hq} + a_{sq}a_{hr} + a_{sh}a_{rq}$.

$$\begin{aligned}
E(2,3) &= 2\sum_{t=1}^{T}\sum_{s\neq t}^{T}\sum_{i_1=1}^{n}\sum_{j_1\neq i_1}^{n}\sum_{j_2\neq j_1}^{n} E\nu_{i_1t}\nu_{j_1t}^2\nu_{j_2t}E\nu_{i_1s}\nu_{j_1s}^2\nu_{j_2s} \\
&= 2\sum_{t=1}^{T}\sum_{s\neq t}^{T}\sum_{i_1=1}^{n}\sum_{j_1\neq i_1}^{n}\sum_{j_2\neq j_1}^{n}[(\sigma^2+\sigma^2he_{j_1}^2)(\sigma^2he_{i_1}e_{j_2}) \\
&\quad +2(\sigma^2he_{i_1}e_{j_1})(\sigma^2he_{j_1}e_{j_2})]^2 \\
&= 2\sum_{t=1}^{T}\sum_{s\neq t}^{T}\sum_{i_1=1}^{n}\sum_{j_1\neq i_1}^{n}\sum_{j_2\neq j_1}^{n}(\sigma^8h^2e_{i_1}^2e_{j_2}^2+6\sigma^8h^3e_{i_1}^2e_{j_1}^2e_{j_2}^2+9\sigma^8h^4e_{i_1}^2e_{j_1}^4e_{j_2}^2) \\
&\leq 2\sigma^8h^2T^2n+12\sigma^8h^3T^2+18\sigma^8h^4T^2 = O(T^2n),
\end{aligned}$$

with $E\nu_r^2\nu_s\nu_h = 2a_{sr}a_{hr} + a_{rr}a_{sh}$,

$$\begin{aligned}
E(2,2) &= 2\sum_{t=1}^{T}\sum_{s\neq t}^{T}\sum_{i=1}^{n}\sum_{j\neq i}^{n} E\nu_{it}^2\nu_{jt}^2E\nu_{is}^2\nu_{js}^2 \\
&= 2\sum_{t=1}^{T}\sum_{s\neq t}^{T}\sum_{i=1}^{n}\sum_{j\neq i}^{n}[(\sigma^2+\sigma^2he_i^2)(\sigma^2+\sigma^2he_j^2)+2(\sigma^2he_ie_j)^2]^2 \\
&= 2\sum_{t=1}^{T}\sum_{s\neq t}^{T}\sum_{i=1}^{n}\sum_{j\neq i}^{n}(\sigma^8+\sigma^8h^2e_i^4+\sigma^8h^2e_j^4+9\sigma^8h^4e_i^4e_j^4 \\
&\quad +2\sigma^8he_i^2+2\sigma^8he_j^2+8\sigma^8h^2e_i^2e_j^2+6\sigma^8h^3e_i^4e_j^2+6\sigma^8h^3e_i^2e_j^4) \\
&\leq 2\sigma^8T^2n^2+4\sigma^8h^2T^2n+18\sigma^8h^4T^2+8\sigma^8hT^2n+16\sigma^8h^2T^2+24\sigma^8h^3T^2 \\
&= O(T^2n^2),
\end{aligned}$$

with $E\nu_s^2\nu_h^2 = a_{ss}a_{hh} + 2a_{sh}^2$. ■

Proof of part (c). Recall that $\tilde{y}_{it} = \tilde{x}_{it}'\beta + \tilde{\nu}_{it}$, $\hat{\nu}_{it} = \tilde{y}_{it} - \tilde{x}_{it}'\tilde{\beta} = \tilde{\nu}_{it} - \tilde{x}_{it}'(\tilde{\beta}-\beta)$, $\hat{\nu}_t = \tilde{\nu}_t - \tilde{x}_t(\tilde{\beta}-\beta)$, $\tilde{\nu}_t = \nu_t - \bar{\nu}_\cdot$, $\hat{S} = \frac{1}{T}\sum_{t=1}^{T}\hat{\nu}_t\hat{\nu}_t'$, and $S = \frac{1}{T}\sum_{t=1}^{T}\nu_t\nu_t'$. Hence,

$$\begin{aligned}
&\frac{1}{n}tr\hat{S} - \frac{1}{n}trS \\
&= \frac{1}{n}tr(\frac{1}{T}\sum_{t=1}^{T}\hat{\nu}_t\hat{\nu}_t' - \frac{1}{T}\sum_{t=1}^{T}\nu_t\nu_t') \\
&= \frac{1}{n}tr[\frac{1}{T}\sum_{t=1}^{T}\tilde{\nu}_t\tilde{\nu}_t' - \frac{1}{T}\sum_{t=1}^{T}\nu_t\nu_t' - \frac{1}{T}\sum_{t=1}^{T}\tilde{x}_t(\tilde{\beta}-\beta)\tilde{\nu}_t' \\
&\quad -\frac{1}{T}\sum_{t=1}^{T}\tilde{\nu}_t(\tilde{\beta}-\beta)'\tilde{x}_t' + \frac{1}{T}\sum_{t=1}^{T}\tilde{x}_t(\tilde{\beta}-\beta)(\tilde{\beta}-\beta)'\tilde{x}_t'] \\
&= -\frac{1}{T}\sigma^2 - \frac{h}{nT} + O_p(\frac{1}{T\sqrt{nT}}) + O_p(\frac{1}{T\sqrt{n}}) + O_p(\frac{1}{nT}) \\
&= -\frac{\sigma^2}{T} + O_p(\frac{1}{T\sqrt{n}}),
\end{aligned}$$

since

$$-\frac{1}{nT}tr[\sum_{t=1}^{T}\tilde{x}_t(\tilde{\beta}-\beta)\tilde{\nu}_t'] = O_p(\frac{1}{nT}),$$
$$-\frac{1}{nT}tr[\sum_{t=1}^{T}\tilde{\nu}_t(\tilde{\beta}-\beta)'\tilde{x}_t'] = O_p(\frac{1}{nT}),$$

$$\frac{1}{nT} tr[\sum_{t=1}^{T} \tilde{x}_t(\tilde{\beta}-\beta)(\tilde{\beta}-\beta)'\tilde{x}_t'] = O_p(\frac{1}{nT}),$$

$$\begin{aligned}
& \frac{1}{nT} tr[\sum_{t=1}^{T} \tilde{\nu}_t\tilde{\nu}_t' - \frac{1}{T}\sum_{t=1}^{T} \nu_t\nu_t'] \\
= & -\frac{1}{n} tr(\bar{\nu}.\bar{\nu}_.') = -\frac{1}{n}\bar{\nu}_.'\bar{\nu}. = -\frac{1}{n}\sum_{i=1}^{n} \bar{\nu}_{i\cdot}^2 = -\frac{1}{n}\sum_{i=1}^{n}(\frac{1}{T}\sum_{t=1}^{T} \nu_{it})^2 \\
= & -\frac{1}{nT^2}\sum_{t=1}^{T}\sum_{i=1}^{n} \nu_{it}^2 - \frac{1}{nT^2}\sum_{t=1}^{T}\sum_{s\neq t}^{T}\sum_{i=1}^{n} \nu_{is}\nu_{it} \\
= & -\frac{1}{nT^2}\sum_{t=1}^{T}\sum_{i=1}^{n}(\sigma^2+\sigma^2 he_i^2) - \frac{1}{nT^2}\sum_{t=1}^{T}\sum_{i=1}^{n}(\nu_{it}^2 - E\nu_{it}^2) \\
& -\frac{1}{nT^2}\sum_{t=1}^{T}\sum_{s\neq t}^{T}\sum_{i=1}^{n} \nu_{is}\nu_{it} \\
= & -\frac{1}{T}\sigma^2 - \frac{\sigma^2 h}{nT} + O_p(\frac{1}{T\sqrt{nT}}) + O_p(\frac{1}{T\sqrt{n}}).
\end{aligned}$$

In establishing the above results, we have used:

$$\sum_{t=1}^{T}\sum_{i=1}^{n} \tilde{x}_{it}\tilde{x}_{it}' = O_p(nT),$$

$$\sum_{t=1}^{T}\sum_{i=1}^{n} \tilde{x}_{it}\tilde{\nu}_{it} = O_p(\sqrt{nT}),$$

$$\tilde{\beta}-\beta = (\sum_{t=1}^{T}\sum_{i=1}^{n} \tilde{x}_{it}\tilde{x}_{it}')^{-1}\sum_{t=1}^{T}\sum_{i=1}^{n} \tilde{x}_{it}\tilde{\nu}_{it} = O_p\left(\frac{1}{\sqrt{nT}}\right),$$

$$\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}(\nu_{it}^2 - E\nu_{it}^2) = O_p\left(\frac{1}{\sqrt{nT}}\right)\Bigg($$

and

$$\begin{aligned}
& E(-\frac{1}{nT^2}\sum_{t=1}^{T}\sum_{s\neq t}^{T}\sum_{i=1}^{n} \nu_{is}\nu_{it})^2 \\
= & \frac{2}{n^2T^4}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{t=1}^{T}\sum_{s\neq t}^{T} E\nu_{is}\nu_{js}\nu_{it}\nu_{jt} \\
= & \frac{2}{n^2T^4}\sum_{i=1}^{n}\sum_{j=1}^{n} T(T-1)E^2\nu_{is}\nu_{js} \\
= & \frac{2}{n^2T^4}T(T-1)[\sum_{i=1}^{n}\sum_{j\neq i}^{n} \sigma^4 h^2 e_i^2 e_j^2 + \sum_{i=1}^{n}(\sigma^2+\sigma^2 he_i^2)^2] \\
= & \frac{2}{n^2T^4}T(T-1)(n\sigma^4 + 2\sigma^4 h + \sigma^4 h^2) \\
= & \frac{2}{n^2T^4}O_p(nT^2) = O_p(\frac{1}{nT^2}).
\end{aligned}$$

■

Proof of part (d). Note that

$$
\begin{aligned}
& \frac{1}{n}tr\hat{S}^2 - \frac{1}{n}trS^2 \\
= \; & \frac{2}{n}tr[S(\hat{S}-S)] - \frac{1}{n}tr(\hat{S}-S)^2 \\
= \; & \frac{2}{n}tr[S(-A_0-A_1-A_2+A_3)] + \frac{1}{n}tr(-A_0-A_1-A_2+A_3)^2 \\
= \; & -\frac{4}{n}tr(SA_1) + \frac{2}{n}tr(SA_3) + \frac{2}{n}tr(A_1^2) + \frac{2}{n}tr(A_1A_2) - \frac{4}{n}tr(A_1A_3) \\
& + \frac{1}{n}tr(A_3^2) - \frac{2}{n}tr(SA_0) + \frac{1}{n}tr(A_0^2) + \frac{4}{n}tr(A_0A_1) - \frac{2}{n}tr(A_0A_3),
\end{aligned}
$$

since

$$tr(A_0A_1) = tr(A_1A_0) = tr(A_0A_2) = tr(A_2A_0),$$

$$tr(A_1A_2) = tr(A_2A_1),$$

$$tr(A_3A_1) = tr(A_1A_3) = tr(A_3A_2) = tr(A_2A_3),$$

$$tr(A_1^2) = tr(A_2^2),$$

$$tr(SA_2) = tr(SA_1).$$

Using Lemma 3, we have

$$
\begin{aligned}
& \frac{1}{n}tr\hat{S}^2 - \frac{1}{n}trS^2 \\
= \; & -2[\frac{1}{T}\sigma^4 + \frac{n}{T^2}\sigma^4 + O_p(\frac{1}{T\sqrt{T}})] + [\frac{n}{T^2}\sigma^4 + O_p(\frac{\sqrt{n}}{T^2})] \\
& + O_p(\frac{1}{nT}) + O_p(\frac{1}{T^2}) + O_p(\frac{1}{T\sqrt{nT}}) \\
= \; & -\frac{2}{T}\sigma^4 - \frac{n}{T^2}\sigma^4 + O_p(\frac{1}{T\sqrt{T}}).
\end{aligned}
$$

Here we used $\frac{n}{T} \to c \in (0,\infty)$ implicitly. ■

C Proof of Theorem 2

Proof. Now

$$\widehat{J}_{LW} - J_{LW} = \frac{TW_2(\frac{1}{n}trS)^2 - 2TW_1\frac{1}{n}trS\frac{1}{n}trS^2 - TW_1^2\frac{1}{n}trS^2}{2(\frac{1}{n}trS + W_1)^2(\frac{1}{n}trS)^2}.$$

For the numerator,

$$\begin{aligned}
& TW_2(\frac{1}{n}trS)^2 - 2TW_1\frac{1}{n}trS\frac{1}{n}trS^2 - TW_1^2\frac{1}{n}trS^2 \\
= \;& T[-\frac{2}{T}\sigma^4 - \frac{n}{T^2}\sigma^4 + O_p(\frac{1}{T\sqrt{T}})][\sigma^2 + O_p(\frac{1}{\sqrt{nT}})]^2 \\
& -2T[-\frac{\sigma^2}{T} + O_p(\frac{1}{T\sqrt{n}})][\sigma^2 + O_p(\frac{1}{\sqrt{nT}})][(\frac{n}{T}+1)\sigma^4 + O_p(\frac{1}{\sqrt{T}})] \\
& -T[-\frac{\sigma^2}{T} + O_p(\frac{1}{T\sqrt{n}})]^2[(\frac{n}{T}+1)\sigma^4 + O_p(\frac{1}{\sqrt{T}})] \\
= \;& [-2\sigma^4 - \frac{n}{T}\sigma^4 + O_p(\frac{1}{\sqrt{T}})][\sigma^4 + O_p(\frac{1}{\sqrt{nT}})] \\
& +[2\sigma^2 + O_p(\frac{1}{\sqrt{n}})][\sigma^2 + O_p(\frac{1}{\sqrt{nT}})][(\frac{n}{T}+1)\sigma^4 + O_p(\frac{1}{\sqrt{T}})] \\
& +[-\frac{\sigma^4}{T} + O_p(\frac{1}{T\sqrt{n}})][(\frac{n}{T}+1)\sigma^4 + O_p(\frac{1}{\sqrt{T}})] \\
= \;& -2\sigma^8 - \frac{n}{T}\sigma^8 + O_p(\frac{1}{\sqrt{T}}) + O_p(\frac{1}{\sqrt{nT}}) \\
& +2(\frac{n}{T}+1)\sigma^8 + O_p(\frac{1}{\sqrt{n}}) + O_p(\frac{1}{\sqrt{T}}) \\
= \;& \frac{n}{T}\sigma^8 - \frac{n}{T^2}\sigma^8 + O_p(\frac{1}{\sqrt{T}}).
\end{aligned}$$

For the denominator,

$$\begin{aligned}
& 2(\frac{1}{n}trS + W_1)^2(\frac{1}{n}trS)^2 \\
= \;& 2[\sigma^2 + O_p(\frac{1}{\sqrt{nT}}) - \frac{\sigma^2}{T} + O_p(\frac{1}{T\sqrt{n}})]^2[\sigma^2 + O_p(\frac{1}{\sqrt{nT}})]^2 \\
= \;& 2[\frac{(T-1)^2}{T^2}\sigma^4 + O_p(\frac{1}{\sqrt{nT}})][\sigma^4 + O_p(\frac{1}{\sqrt{nT}})] \\
= \;& \frac{2(T-1)^2}{T^2}\sigma^8 + O_p(\frac{1}{\sqrt{nT}}).
\end{aligned}$$

Hence $\widehat{J}_{LW} - J_{LW} - \frac{n}{2(T-1)} = \frac{\frac{n}{T}\sigma^8 - \frac{n}{T^2}\sigma^8 + O_p(\frac{1}{\sqrt{T}})}{\frac{2(T-1)^2}{T^2}\sigma^8 + O_p(\frac{1}{\sqrt{nT}})} - \frac{n}{2(T-1)} \xrightarrow{p} 0$ as $(n,T) \to \infty$ and $\frac{n}{T} \to c \in (0,\infty)$.
■

D Proof of Theorem 4

Proof. Under the strong factor alternative, the $n \times 1$ vectors $\nu_1, ..., \nu_T$ are *iid* $N(0, \Sigma_n)$, where $\Sigma_n = \sigma^2(I_n + \sum_{j=1}^r h_j e_j e_j')$ and $\frac{h_j}{n} \to d_j \in (0,\infty)$ for $j = 1, ..., r$.

$\Sigma_n = \Gamma_n \Lambda_n \Gamma_n'$, where $\Lambda_n = diag(\lambda_1, ..., \lambda_n)$. $\lambda_1, ..., \lambda_n$ are eigenvalues of Σ_n and $\lambda_j = \sigma^2(1+h_j)$ for $j = 1, ..., r$, $\lambda_j = \sigma^2$ for $j = r+1, ..., n$. $\Gamma_n = (e_1, ...e_r, g_1, ...g_{n-r})$ and $g_1, ...g_{n-r}$ are constructed such that Γ_n is orthogonal.

Let $w_t = \Lambda_n^{-\frac{1}{2}}\Gamma_n'\nu_t$, then w_t is *iid* $N(0, I_n)$. Let $V = (\nu_1, ..., \nu_T)$ and $W = (w_1, ..., w_T)$, then $W = \Lambda_n^{-\frac{1}{2}}\Gamma_n'V$. Let $W' = (\omega_1, ...\omega_n)$, then ω_i is *iid* $N(0, I_T)$, since we assume there is no time dependence.

$$\begin{aligned} trS &= \frac{1}{T}trVV' = \frac{1}{T}trV'V = \frac{1}{T}tr(V'\Gamma_n\Lambda_n^{-\frac{1}{2}})\Lambda_n(\Lambda_n^{-\frac{1}{2}}\Gamma_n'V) = \frac{1}{T}trW'\Lambda_nW \\ &= \frac{1}{T}tr(\sum\nolimits_{i=1}^n \lambda_i\omega_i\omega_i') = \frac{1}{T}tr(\sum\nolimits_{i=1}^n \lambda_i\omega_i'\omega_i) = \frac{1}{T}\sum\nolimits_{i=1}^n \lambda_i\alpha_{ii}. \end{aligned}$$

Here $\alpha_{ii} = \omega_i'\omega_i$ has a chi-squared distribution of with T degrees of freedom. Note that

$$\begin{aligned} & E(\frac{1}{n}trS) = \frac{1}{nT}E(\sum\nolimits_{i=1}^n \lambda_i\alpha_{ii}) = \frac{1}{n}\sum\nolimits_{i=1}^n \lambda_i = \sigma^2(1 + \sum\nolimits_{j=1}^r h_j/n) \\ \to\ & \sigma^2(1 + \sum\nolimits_{j=1}^r d_j) \end{aligned}$$

and

$$\begin{aligned} & Var(\frac{1}{n}trS) = E(\frac{1}{n}trS)^2 - E^2(\frac{1}{n}trS) = \frac{1}{n^2T^2}E(\sum\nolimits_{i=1}^n \lambda_i\alpha_{ii})^2 - (\frac{1}{n}\sum\nolimits_{i=1}^n \lambda_i)^2 \\ = & \frac{1}{n^2T^2}E(\sum\nolimits_{i=1}^n \lambda_i\alpha_{ii}^2 + 2\sum\nolimits_{i<j} \lambda_i\lambda_j\alpha_{ii}\alpha_{jj}) - (\frac{1}{n}\sum\nolimits_{i=1}^n \lambda_i)^2, \end{aligned}$$

with

$$E(\alpha_{ii}^2) = T^2 + 2T$$

$$E(\alpha_{ii}\alpha_{jj}) = E(\alpha_{ii})E(\alpha_{jj}) = T^2.$$

We have

$$\begin{aligned} & Var(\frac{1}{n}trS) = \frac{1}{n^2T^2}(2T\sum\nolimits_{i=1}^n \lambda_i^2 + T^2(\sum\nolimits_{i=1}^n \lambda_i)^2) - (\frac{1}{n}\sum\nolimits_{i=1}^n \lambda_i)^2 \\ = & \frac{2}{n^2T}\sum\nolimits_{i=1}^n \lambda_i^2 = \frac{2}{n^2T}\sigma^4(\sum\nolimits_{j=1}^r h_j^2 + 2\sum\nolimits_{j=1}^r h_j + n) \\ = & \frac{2}{T}\sigma^4(\sum\nolimits_{j=1}^r (\frac{h_j}{n})^2 + 2\sum\nolimits_{j=1}^r \frac{h_j}{n^2} + \frac{1}{n}) \to 0. \end{aligned}$$

Therefore $\frac{1}{n}trS \xrightarrow{p} \sigma^2(1 + \sum_{j=1}^r d_j)$. Note that

$$\begin{aligned} \frac{1}{n}trS^2 &= \frac{1}{nT^2}tr(VV'VV') = \frac{1}{nT^2}tr(V'VV'V) = \frac{1}{nT^2}tr(W'\Lambda_nWW'\Lambda_nW) \\ &= \frac{1}{nT^2}tr(\sum\nolimits_{i=1}^n \lambda_i\omega_i\omega_i')(\sum\nolimits_{J=1}^n \lambda_j\omega_j\omega_j') \\ &= \frac{1}{nT^2}[\sum\nolimits_{i=1}^n \lambda_i^2(\omega_i'\omega_i)^2 + 2\sum\nolimits_{i<j} \lambda_i\lambda_j(\omega_i'\omega_j)^2] \\ &= \frac{1}{nT^2}(\sum\nolimits_{i=1}^n \lambda_i^2\alpha_{ii}^2 + 2\sum\nolimits_{i<j} \lambda_i\lambda_j\alpha_{ij}^2) \end{aligned}$$

with $\alpha_{ij} = \omega_i'\omega_j$. $h_j \geq 0$ for $j = 1, ..., r$, so $\lambda_j = \sigma^2(1+h_j) \geq \sigma^2$ for all j. Hence

$$\frac{1}{n}trS^2 \geq \frac{1}{nT^2}(\lambda_1^2 - \sigma^4)\alpha_{11}^2 = \sigma^4 \frac{h_1^2 + 2h_1}{n}\frac{\alpha_{11}^2}{T^2}.$$

Note that α_{11} follows a Chi-square distribution with T degree of freedom. Hence $\frac{\alpha_{11}-T}{\sqrt{2T}} \xrightarrow{d} N(0,1)$, and

$$\alpha_{11} = T + \sqrt{2T}(\frac{\alpha_{11}-T}{\sqrt{2T}}) = T + O_p(\sqrt{T}).$$

Consequently,

$$\begin{aligned}
\frac{1}{n}trS^2 &\geq n\sigma^4\frac{h_1^2+2h_1}{n^2}\frac{\alpha_{11}^2}{T^2} = n\sigma^4\frac{h_1^2+2h_1}{n^2}\frac{T^2+O_p(T\sqrt{T})}{T^2} \\
&= n\sigma^4\frac{h_1^2+2h_1}{n^2} + n\sigma^4\frac{h_1^2+2h_1}{n^2}O_p(\frac{1}{\sqrt{T}}) = n\sigma^4 d_1^2 + O_p(\frac{n}{\sqrt{T}}).
\end{aligned}$$

This implies $\frac{1}{n}trS^2 \xrightarrow{p} \infty$ at least as fast as n. On the other hand,

$$\begin{aligned}
\frac{1}{n}trS^2 &= \frac{1}{nT^2}(\sum\nolimits_{i=1}^n \lambda_i^2\alpha_{ii}^2 + \sum\nolimits_{i\neq j}\lambda_i\lambda_j\alpha_{ij}^2) \\
&= \frac{1}{nT^2}\sum\nolimits_{i=1}^r(h_i^2+2h_i)\alpha_{ii}^2 + \frac{1}{nT^2}\sum\nolimits_{i=1}^n\alpha_{ii}^2 + \frac{1}{nT^2}\sum\nolimits_{i=1}^r\sum\nolimits_{j=1,j\neq i}^r(1+h_i)(1+h_j)\alpha_{ij}^2 \\
&\quad + \frac{1}{nT^2}\sum\nolimits_{i=1}^r\sum\nolimits_{j=r+1}^n(1+h_i)\alpha_{ij}^2 + \frac{1}{nT^2}\sum\nolimits_{i=r+1}^n\sum\nolimits_{j=1}^r(1+h_j)\alpha_{ij}^2 \\
&\quad + \frac{1}{nT^2}\sum\nolimits_{i=r+1}^n\sum\nolimits_{j=r+1,j\neq i}^n\alpha_{ij}^2 \\
&= \frac{1}{nT^2}\sum\nolimits_{i=1}^r(h_i^2+2h_i)\alpha_{ii}^2 + \frac{1}{nT^2}\sum\nolimits_{i=1}^n\alpha_{ii}^2 + \frac{1}{nT^2}\sum\nolimits_{i=1}^r\sum\nolimits_{j=1,j\neq i}^r h_ih_j\alpha_{ij}^2 \\
&\quad + \frac{1}{nT^2}\sum\nolimits_{i=1}^r\sum\nolimits_{j=1,j\neq i}^r(h_i+h_j)\alpha_{ij}^2 + \frac{1}{nT^2}\sum\nolimits_{i=1}^r\sum\nolimits_{j=1,j\neq i}^r\alpha_{ij}^2 \\
&\quad + \frac{1}{nT^2}\sum\nolimits_{i=1}^r\sum\nolimits_{j=r+1}^n h_i\alpha_{ij}^2 + \frac{1}{nT^2}\sum\nolimits_{i=1}^r\sum\nolimits_{j=r+1}^n\alpha_{ij}^2 + \frac{1}{nT^2}\sum\nolimits_{i=r+1}^n\sum\nolimits_{j=1}^r h_j\alpha_{ij}^2 \\
&\quad + \frac{1}{nT^2}\sum\nolimits_{i=r+1}^n\sum\nolimits_{j=1}^r\alpha_{ij}^2 + \frac{1}{nT^2}\sum\nolimits_{i=r+1}^n\sum\nolimits_{j=r+1,j\neq i}^n\alpha_{ij}^2 \\
&= \frac{1}{nT^2}\sum\nolimits_{i=1}^r(h_i^2+2h_i)\alpha_{ii}^2 + \frac{1}{nT^2}\sum\nolimits_{i=1}^n\alpha_{ii}^2 + \frac{1}{nT^2}\sum\nolimits_{i=1}^r\sum\nolimits_{j=1,j\neq i}^r h_ih_j\alpha_{ij}^2 \\
&\quad + \frac{1}{nT^2}\sum\nolimits_{i=1}^r\sum\nolimits_{j=1,j\neq i}^n h_i\alpha_{ij}^2 + \frac{1}{nT^2}\sum\nolimits_{i=1,i\neq j}^n\sum\nolimits_{j=1}^r h_j\alpha_{ij}^2 + \frac{1}{nT^2}\sum\nolimits_{i=1}^n\sum\nolimits_{j=1,j\neq i}^n\alpha_{ij}^2 \\
&= \frac{1}{nT^2}\sum\nolimits_{i=1}^r(h_i^2+2h_i)\alpha_{ii}^2 + \frac{1}{nT^2}\sum\nolimits_{i=1}^r\sum\nolimits_{j=1,j\neq i}^r h_ih_j\alpha_{ij}^2 \\
&\quad + 2\frac{1}{nT^2}\sum\nolimits_{i=1}^r\sum\nolimits_{j=1,j\neq i}^n h_i\alpha_{ij}^2 + \frac{1}{nT^2}\sum\nolimits_{i=1}^n\sum\nolimits_{j=1}^n\alpha_{ij}^2 \\
&= O_p(n) + O_p(\frac{n}{T}) + O_p(\frac{n}{T}) + O_p(\frac{n}{T}) = O_p(n).
\end{aligned}$$

This is because

$$\begin{aligned}
\alpha_{ij} &= O_p(\sqrt{T}), \\
\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_{ij}^2 &= \frac{1}{n}trWW' = (\frac{n}{T}+1)\sigma^4 + O_p(\frac{1}{\sqrt{T}}), \\
\frac{1}{nT^2}\sum_{i=1}^{r}\sum_{j=1,j\neq i}^{n} h_i\alpha_{ij}^2 &= O_p(\frac{n}{T}).
\end{aligned}$$

The last equation follows from

$$\begin{aligned}
E(\frac{1}{nT^2}\sum_{j=1,j\neq i}^{n}\alpha_{ij}^2)^2 &= \frac{1}{n^2T^4}\sum_{j=1,j\neq i}^{n}E\alpha_{ij}^4 + \frac{1}{n^2T^4}\sum_{j=1,j\neq i}^{n}\sum_{k=1,k\neq i,k\neq j}^{n}E\alpha_{ij}^2\alpha_{ik}^2 \\
&= \frac{1}{n^2T^4}(n-1)[3T(T+2)] + \frac{1}{n^2T^4}(n-1)(n-2)[T(T+2)] \\
&= \frac{1}{n^2T^4}(n^2-1)T(T+2) = O(\frac{1}{T^2}),
\end{aligned}$$

for any $i = 1, ..., r$. Therefore, $\frac{1}{n}trS^2 = O_p(n)$ exactly, i.e. $\frac{1}{n}trS^2 \neq o_p(n)$. Hence

$$U = (\frac{1}{n}trS)^{-2}(\frac{1}{n}trS^2) - 1 = O_p(n),$$

and

$$J_{LW} = \frac{TU-n-1}{2} = O_p(nT).$$

■

E Proof of Proposition 2

Proof of part (a). Note that

$$\begin{aligned}
\frac{1}{n}trS &= \frac{1}{n}tr[\frac{1}{T}\sum_{t=1}^{T}\nu_t\nu_t'] = \frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\nu_{it}^2 \\
&= \sigma^2(1+\frac{\sum_{j=1}^{r}h_j}{n}) + \frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}(\nu_{it}^2 - E\nu_{it}^2) \\
&= \sigma^2(1+\sum_{j=1}^{r}d_j) + O_p(\frac{1}{\sqrt{T}}),
\end{aligned}$$

since

$$\begin{aligned}
&E[\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}(\nu_{it}^2 - E\nu_{it}^2)]^2 = \frac{1}{n^2T^2}\sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{n}E(\nu_{it}^2 - E\nu_{it}^2)(\nu_{jt}^2 - E\nu_{jt}^2) \\
= \; &\frac{1}{n^2T^2}O(n^2T) = O(\frac{1}{T}).
\end{aligned}$$

■

Proof of part (b). As shown in part (b) in Proposition 1,

$$
\begin{aligned}
&\frac{1}{n}trS^2 = \frac{1}{nT^2}\sum\nolimits_{t=1}^{T}\sum\nolimits_{i=1}^{n}\nu_{it}^4 + \frac{1}{nT^2}\sum\nolimits_{t=1}^{T}\sum\nolimits_{j\neq i}^{n}\sum\nolimits_{i=1}^{n}\nu_{it}^2\nu_{jt}^2 \\
&+\frac{1}{nT^2}\sum\nolimits_{t=1}^{T}\sum\nolimits_{s\neq t}^{T}\sum\nolimits_{i=1}^{n}\nu_{it}^2\nu_{is}^2 + \frac{1}{nT^2}\sum\nolimits_{t=1}^{T}\sum\nolimits_{s\neq t}^{T}\sum\nolimits_{i=1}^{n}\sum\nolimits_{j\neq i}^{n}\nu_{it}\nu_{is}\nu_{js}\nu_{jt} \\
= \;& O_p(\frac{1}{T}) + O_p(1) + O_p(1) + \{\frac{n(T-1)}{T}\sigma^4[\sum\nolimits_{j=1}^{r} d_j^2 - \sum\nolimits_{i=1}^{n}(\sum\nolimits_{j=1}^{r} d_j e_{i,j}^2)^2] + O_p(\sqrt{n})\} \\
= \;& \frac{n(T-1)}{T}\sigma^4[\sum\nolimits_{j=1}^{r} d_j^2 - \sum\nolimits_{i=1}^{n}(\sum\nolimits_{j=1}^{r} d_j e_{i,j}^2)^2] + O_p(\sqrt{n}).
\end{aligned}
$$

Here we have used the following four results:

(1)

$$
\frac{1}{nT^2}\sum\nolimits_{t=1}^{T}\sum\nolimits_{i=1}^{n}\nu_{it}^4 = \frac{1}{T}E\nu_{it}^4 + \frac{1}{nT^2}\sum\nolimits_{t=1}^{T}\sum\nolimits_{i=1}^{n}(\nu_{it}^4 - E\nu_{it}^4) = O_p(\frac{1}{T}) + O_p(\frac{1}{T}) = O_p(\frac{1}{T}).
$$

(2) If $\frac{n}{T} \to c \in (0,\infty)$,

$$
\frac{1}{nT^2}\sum\nolimits_{t=1}^{T}\sum\nolimits_{j\neq i}^{n}\sum\nolimits_{i=1}^{n}\nu_{it}^2\nu_{jt}^2 = O_p(1).
$$

(3)

$$
\frac{1}{nT^2}\sum\nolimits_{t=1}^{T}\sum\nolimits_{s\neq t}^{T}\sum\nolimits_{i=1}^{n}\nu_{it}^2\nu_{is}^2 = O_p(1).
$$

(4)

$$
\begin{aligned}
&\frac{1}{nT^2}\sum\nolimits_{t=1}^{T}\sum\nolimits_{s\neq t}^{T}\sum\nolimits_{i=1}^{n}\sum\nolimits_{j\neq i}^{n}\nu_{it}\nu_{is}\nu_{js}\nu_{jt} \\
= \;& \frac{n(T-1)}{T}\sigma^4[\sum\nolimits_{j=1}^{r} d_j^2 - \sum\nolimits_{i=1}^{n}(\sum\nolimits_{j=1}^{r} d_j e_{i,j}^2)^2] + O_p(\sqrt{n}).
\end{aligned}
$$

This is because:

$$
\begin{aligned}
&E[\frac{1}{nT^2}\sum\nolimits_{t=1}^{T}\sum\nolimits_{s\neq t}^{T}\sum\nolimits_{i=1}^{n}\sum\nolimits_{j\neq i}^{n}\nu_{it}\nu_{is}\nu_{js}\nu_{jt}]^2 \\
= \;& \frac{1}{n^2T^4}\sum\nolimits_{t_1=1}^{T}\sum\nolimits_{s_1\neq t_1}^{T}\sum\nolimits_{i_1=1}^{n}\sum\nolimits_{j_1\neq i_1}^{n}\sum\nolimits_{t_2=1}^{T}\sum\nolimits_{s_2\neq t_2}^{T}\sum\nolimits_{i_2=1}^{n}\sum\nolimits_{j_2\neq i_2}^{n}E(\nu_{i_1t_1} \\
&\nu_{i_1s_1}\nu_{j_1s_1}\nu_{j_1t_1}\nu_{i_2t_2}\nu_{i_2s_2}\nu_{j_2s_2}\nu_{j_2t_2}) \\
= \;& \frac{1}{n^2T^4}[E(4,4) + E(4,3) + E(4,2) + E(3,4) \\
&+E(3,3) + E(3,2) + E(2,4) + E(2,3) + E(2,2)] \\
= \;& \frac{1}{n^2T^4}E(4,4) + O(n) = \frac{(T-1)^2}{n^2T^2}\sigma^8[\sum\nolimits_{j=1}^{r} h_j^2 - \sum\nolimits_{i_1=1}^{n}(\sum\nolimits_{j=1}^{r} h_j e_{i_1,j}^2)^2]^2 + O(n) \\
= \;& E^2[\frac{1}{nT^2}\sum\nolimits_{t=1}^{T}\sum\nolimits_{s\neq t}^{T}\sum\nolimits_{i=1}^{n}\sum\nolimits_{j\neq i}^{n}\nu_{it}\nu_{is}\nu_{js}\nu_{jt}] + O(n).
\end{aligned}
$$

With $\sum_{j_1}^n e_{j_1,j}^2 = 1$ for each j and $\sum_{j_1}^n e_{j_1,j}e_{j_1,k} = 0$, it can be shown that

$$\begin{aligned}
& E(4,4) = \sum_{t_1=1}^{T}\sum_{s_1\neq t_1}^{T}\sum_{i_1=1}^{n}\sum_{j_1\neq i_1}^{n}\sum_{t_2=1}^{T}\sum_{s_2\neq t_2}^{T}\sum_{i_2=1}^{n}\sum_{j_2\neq i_2}^{n}\sigma^8 \\
& (\sum_{j=1}^{r} h_j e_{i_1,j}e_{j_1,j})^2(\sum_{j=1}^{r} h_j e_{i_2,j}e_{j_2,j})^2 \\
= \ & \sigma^8 T^2(T-1)^2[\sum_{i_1=1}^{n}\sum_{j_1\neq i_1}^{n}(\sum_{j=1}^{r} h_j^2 e_{i_1,j}^2 e_{j_1,j}^2 + 2\sum_{j<k} h_j h_k e_{i_1,j}e_{j_1,j}e_{i_1,k}e_{j_1,k})]^2 \\
= \ & \sigma^8 T^2(T-1)^2[\sum_{i_1=1}^{n}(\sum_{j=1}^{r} h_j^2 e_{i_1,j}^2(1-e_{i_1,j}^2) - 2\sum_{j<k} h_j h_k e_{i_1,j}^2 e_{i_1,k}^2]^2 \\
= \ & \sigma^8 T^2(T-1)^2[\sum_{i_1=1}^{n}(\sum_{j=1}^{r} h_j^2 e_{i_1,j}^2 - \sum_{j=1}^{r} h_j^2 e_{i_1,j}^4) - 2\sum_{j<k} h_j h_k e_{i_1,j}^2 e_{i_1,k}^2]^2 \\
= \ & \sigma^8 T^2(T-1)^2[\sum_{j=1}^{r} h_j^2 - \sum_{i_1=1}^{n}(\sum_{j=1}^{r} h_j e_{i_1,j}^2)^2]^2,
\end{aligned}$$

$$\begin{aligned}
& E[\frac{1}{nT^2}\sum_{t=1}^{T}\sum_{s\neq t}^{T}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\nu_{it}\nu_{is}\nu_{js}\nu_{jt}] = \frac{(T-1)}{nT}\sigma^4\sum_{i_1=1}^{n}\sum_{j_1\neq i_1}^{n}(\sum_{j=1}^{r} h_j e_{i_1,j}e_{j_1,j})^2 \\
= \ & \frac{(T-1)}{nT}\sigma^4[\sum_{j=1}^{r} h_j^2 - \sum_{i_1=1}^{n}(\sum_{j=1}^{r} h_j e_{i_1,j}^2)^2].
\end{aligned}$$

■

Proof of part (c). As shown in part (c) of Proposition 1,

$$\begin{aligned}
& \frac{1}{n}tr(\hat{S}-S) = \frac{1}{n}tr[\sum_{i=1}^{n}(\frac{1}{T}\sum_{t=1}^{T}\nu_{it})^2 - \frac{1}{T}\sum_{t=1}^{T}\tilde{x}_t(\tilde{\beta}-\beta)\tilde{\nu}_t' \\
& -\frac{1}{T}\sum_{t=1}^{T}\tilde{\nu}_t(\tilde{\beta}-\beta)'\tilde{x}_t' + \frac{1}{T}\sum_{t=1}^{T}\tilde{x}_t(\tilde{\beta}-\beta)(\tilde{\beta}-\beta)'\tilde{x}_t'].
\end{aligned}$$

With $\frac{h_j}{n} \to d_j \in (0,\infty)$, $\sum_{t=1}^{T}\sum_{i=1}^{n}\tilde{x}_{it}'\tilde{\nu}_{it} = O_p(\sqrt{nT})$. The proof is as follows. $\tilde{\nu}_t = \Gamma_n\Lambda_n^{\frac{1}{2}}\tilde{w}_t$, where $w_t = \Lambda_n^{-\frac{1}{2}}\Gamma_n'\nu_t$ is *iid* $N(0, I_n)$. Hence

$$\begin{aligned}
& \sum_{t=1}^{T}\sum_{i=1}^{n}\tilde{x}_{it}\tilde{\nu}_{it} = \sum_{t=1}^{T}\tilde{x}_t'\Gamma_n\Lambda_n^{\frac{1}{2}}\tilde{w}_t = \sum_{t=1}^{T}\tilde{x}_t'\Gamma_n(\Lambda_n^{\frac{1}{2}} - \sigma I_n)\tilde{w}_t + \sigma\sum_{t=1}^{T}\tilde{x}_t'\Gamma_n I_n\tilde{w}_t \\
= \ & \sigma\sum_{t=1}^{T}\tilde{x}_t'\Gamma_n H\tilde{w}_t + \sigma\sum_{t=1}^{T}\tilde{x}_t'\Gamma_n\tilde{w}_t = \sigma\sum_{t=1}^{T}y_t'H\tilde{w}_t + \sigma\sum_{t=1}^{T}y_t'\tilde{w}_t,
\end{aligned}$$

where $H = diag(\sqrt{1+h_1}-1, ..., \sqrt{1+h_r}-1, 0, ..., 0)$, $y_t = \Gamma_n'\tilde{x}_t$. Hence

$$\begin{aligned}
& \sum_{t=1}^{T}\sum_{i=1}^{n}\tilde{x}_{it}\tilde{\nu}_{it} = \sigma\sum_{j=1}^{r}\sum_{t=1}^{T}(\sqrt{1+h_j}-1)y_{jt}\tilde{w}_{jt} + \sigma\sum_{t=1}^{T}y_t'\tilde{w}_t \\
= \ & \sqrt{n}\sigma\sum_{j=1}^{r}\sum_{t=1}^{T}(\sqrt{\frac{1+h_j}{n}} - \sqrt{\frac{1}{n}})y_{jt}\tilde{w}_{jt} + \sigma\sum_{t=1}^{T}y_t'\tilde{w}_t.
\end{aligned}$$

With some regularity conditions on X and $\frac{h_j}{n} \to d_j \in (0,\infty)$, it is easy to see that

$$\sum_{t=1}^{T}\sum_{i=1}^{n}\tilde{x}_{it}\tilde{\nu}_{it} = O_p(\sqrt{nT}) + O_p(\sqrt{nT}) = O_p(\sqrt{nT}).$$

Consequently,

$$\tilde{\beta}-\beta = (\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\tilde{x}_{it}\tilde{x}_{it}')^{-1}(\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\tilde{x}_{it}\tilde{\nu}_{it}) = O_p(\frac{1}{\sqrt{nT}}),$$

$$-\frac{1}{nT}tr[\sum_{t=1}^{T}\tilde{x}_t(\tilde{\beta}-\beta)\tilde{\nu}_t'] = O_p(\frac{1}{nT}),$$
$$-\frac{1}{nT}tr[\sum_{t=1}^{T}\tilde{\nu}_t(\tilde{\beta}-\beta)'\tilde{x}_t'] = O_p(\frac{1}{nT}),$$

and

$$\frac{1}{nT}tr[\sum_{t=1}^{T}\tilde{x}_t(\tilde{\beta}-\beta)(\tilde{\beta}-\beta)'\tilde{x}_t'] = O_p(\frac{1}{nT}).$$

In addition,

$$\frac{1}{n}\sum_{i=1}^{n}(\frac{1}{T}\sum_{t=1}^{T}\nu_{it})^2 = \frac{1}{nT}\sum_{i=1}^{n}(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\nu_{it})^2 = O_p(\frac{1}{T}).$$

Therefore,

$$\frac{1}{n}tr(\hat{S}-S) = O_p(\frac{1}{T}) + O_p(\frac{1}{nT}) + O_p(\frac{1}{nT}) + O_p(\frac{1}{nT}) = O_p(\frac{1}{T}).$$

■

Proof of part (d). As in part (d) of Proposition 1,

$$\frac{1}{n}tr\hat{S}^2 - \frac{1}{n}trS^2 = -\frac{4}{n}tr(SA_1) + \frac{2}{n}tr(SA_3) + \frac{2}{n}tr(A_1^2) + \frac{2}{n}tr(A_1A_2) - \frac{4}{n}tr(A_1A_3)$$
$$+\frac{1}{n}tr(A_3^2) - \frac{2}{n}tr(SA_0) + \frac{1}{n}tr(A_0^2) + \frac{4}{n}tr(A_0A_1) - \frac{2}{n}tr(A_0A_3).$$

Using Lemma 3,

$$\frac{1}{n}tr\hat{S}^2 - \frac{1}{n}trS^2 = O_p(\frac{\sqrt{n}}{T}) + O_p(\frac{1}{T}) + O_p(\frac{1}{T^2}) + O_p(\frac{1}{T^2}) + O_p(\frac{1}{\sqrt{n}T^2})$$
$$+O_p(\frac{1}{nT^2}) + O_p(\frac{n}{T}) + O_p(\frac{n}{T^2}) + O_p(\frac{\sqrt{n}}{T^2}) + O_p(\frac{1}{T^2}) = O_p(\frac{n}{T}).$$

Here we used $\frac{n}{T} \to c \in (0,\infty)$ implicitly. ■

F Proof of Theorem 5

Proof. Recall that $\widehat{J}_{LW} - J_{LW} = \frac{TW_2(\frac{1}{n}trS)^2 - 2TW_1\frac{1}{n}trS\frac{1}{n}trS^2 - TW_1^2\frac{1}{n}trS^2}{2(\frac{1}{n}trS+W_1)^2(\frac{1}{n}trS)^2}$.

For the numerator,

$$\begin{aligned}
& TW_2(\frac{1}{n}trS)^2 - 2TW_1\frac{1}{n}trS\frac{1}{n}trS^2 - TW_1^2\frac{1}{n}trS^2 \\
= \quad & TO_p(\frac{n}{T})[\sigma^2(1+\sum_{j=1}^{r}d_j) + O_p(\frac{1}{\sqrt{T}})]^2 \\
& -2TO_p(\frac{1}{T})[\sigma^2(1+\sum_{j=1}^{r}d_j) + O_p(\frac{1}{\sqrt{T}})][\frac{n(T-1)}{T}\sigma^4[\sum_{j=1}^{r}d_j^2 - \\
& \sum_{i=1}^{n}(\sum_{j=1}^{r}d_je_{i,j}^2)^2] + O_p(\sqrt{n})] \\
& -T[O_p(\frac{1}{T})]^2[\frac{n(T-1)}{T}\sigma^4[\sum_{j=1}^{r}d_j^2 - \sum_{i=1}^{n}(\sum_{j=1}^{r}d_je_{i,j}^2)^2] + O_p(\sqrt{n})] \\
= \quad & O_p(n) + O_p(n) + O_p(\frac{n}{T}) = O_p(n).
\end{aligned}$$

For the denominator,

$$\begin{aligned}
& 2(\frac{1}{n}trS+W_1)^2(\frac{1}{n}trS)^2 \\
= & \ 2[\sigma^2(1+\sum\nolimits_{j=1}^{r} d_j)+O_p(\frac{1}{\sqrt{T}})+O_p(\frac{1}{T})]^2[\sigma^2(1+\sum\nolimits_{j=1}^{r} d_j)+O_p(\frac{1}{\sqrt{T}})]^2 \\
= & \ 2\sigma^8(1+\sum\nolimits_{j=1}^{r} d_j)^4+O_p(\frac{1}{\sqrt{T}}).
\end{aligned}$$

Therefore, $\widehat{J}_{LW}-J_{LW}=\frac{O_p(n)}{2\sigma^8(1+\sum_{j=1}^{r} d_j)^4+O_p(\frac{1}{\sqrt{T}})}=O_p(n)$. Here we used $\frac{n}{T}\to c\in(0,\infty)$ implicitly. ■

G Proof of Theorem 6

Proof. From the proof of Theorem 4, we know that $\frac{1}{n}trS \xrightarrow{p} 1+\sum_{j=1}^{r} d_j$ and $\frac{1}{n}trS^2 \geq \sigma^4\frac{h_1^2+2h_1}{n}\frac{\alpha_{11}^2}{T^2}$. For any $M>0$,

$$\begin{aligned}
P(\widehat{J}_{LW} &> M)=P(\widehat{J}_{LW}-J_{LW}+J_{LW}>M) \\
&= P(O_p(n)+J_{LW}>M)=P(\frac{O_p(n)+J_{LW}}{nT}>\frac{M}{nT}) \\
&= P(\frac{O_p(n)+[T\frac{\frac{1}{n}trS^2}{(\frac{1}{n}trS)^2}-T-n-1]}{nT}>\frac{M}{nT}) \\
&= P(\frac{1}{n^2}trS^2>(\frac{1}{n}trS)^2(\frac{M}{nT}+\frac{1}{T}+\frac{T+1}{nT}+O_p(\frac{1}{T}))) \\
&\geq P(\sigma^4\frac{h_1^2+2h_1}{n^2}\frac{\alpha_{11}^2}{T^2}>(\frac{1}{n}trS)^2(\frac{M}{nT}+\frac{1}{T}+\frac{T+1}{nT}+O_p(\frac{1}{T}))) \\
&= P(\frac{\alpha_{11}^2}{T^2}>\frac{1}{\sigma^4 d_1^2}(\frac{1}{n}trS)^2(\frac{M}{nT}+\frac{1}{T}+\frac{T+1}{nT}+O_p(\frac{1}{T})))\geq P(\frac{\alpha_{11}^2}{T^2}>\frac{c}{\sqrt{T}})
\end{aligned}$$

for some $c>0$. This holds since $\frac{1}{\sigma^4 d_1^2}(\frac{1}{n}trS)^2(\frac{M}{nT}+\frac{1}{T}+\frac{T+1}{nT}+O_p(\frac{1}{T}))<\frac{c}{\sqrt{T}}$ for a large enough T. Hence

$$P(\widehat{J}_{LW}>M)\geq P(\frac{\alpha_{11}^2}{T^2}>\frac{c}{\sqrt{T}})=P(\frac{\alpha_{11}}{T}>\frac{\sqrt{c}}{T^{\frac{1}{4}}})=P(\frac{\alpha_{11}-T}{\sqrt{2T}}>\sqrt{\frac{T}{2}}\frac{\sqrt{c}}{T^{\frac{1}{4}}}-\sqrt{\frac{T}{2}})\to 1,$$

since $\frac{\alpha_{11}-T}{\sqrt{2T}} \xrightarrow{d} N(0,1)$ and $\sqrt{\frac{T}{2}}\frac{\sqrt{c}}{T^{\frac{1}{4}}}-\sqrt{\frac{T}{2}}\to-\infty$. ■