July 2016

Three Essays on Testing for Cross-Sectional Dependence and Specification in Large Panel Data Models

Bin Peng

Syracuse University

Three Essays on Testing for Cross-Sectional Dependence and Specification in Large Panel Data Models

Bin Peng

B.A., Huazhong University of Science & Technology, 2007

M.A., Huazhong University of Science & Technology, 2010

DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Economics in the Graduate School of Syracuse University.

July 2016

Acknowledgments

I am using this opportunity to express my gratitude to everyone who helped me complete this dissertation at Syracuse University. I am thankful for their aspiring guidance, invaluable constructive criticism, and friendly advice. I am sincerely grateful to them for sharing their honest and illuminating views.

Particularly, I would like to express my sincere gratitude to my advisors Prof. Chihwa Kao, Prof. Badi H. Baltagi and Prof. Yoonseok Lee for the continuous support, their patience, motivation, and immense knowledge. Their guidance helped me in all the time of writing this dissertation. I would also like to thank the rest of my thesis committee: Prof. William Horrace and Prof. Jan Ondrich, for their insightful comments and encouragement. Besides, I am especially grateful to Prof. Sarah Hamersma for chairing my oral defense.

My sincere thanks also go to all the other faculty members, staffs of the economics department, for their input, guidance, accessibility, and helpfulness. I also would like to thank my classmates for the stimulating discussions, for the sleepless nights we were working together, and for all the fun we have had in the last five years.

Last but not the least, I would like to thank my wife, my parents, my parents in law and my sister for supporting me spiritually throughout writing this thesis and my life in general. Their unconditional support, continuous encouragement, quiet patience, and unwavering love are the wealth of all my life.

Table of Contents

List of Tables

Page

Essay I: On Testing for Sphericity with Non-normality in a Fixed Effects Panel Data Model

Essay II: Testing Cross-Sectional Dependence in Large Panel Data Models with Serial Correlation

Essay III: Tests of Specification for Large Dynamic Panel Data Models

Essay I: On Testing for Sphericity with Non-normality in a Fixed Effects Panel Data Model

1 Introduction

This paper proposes testing the null of sphericity of the variance-covariance matrix in a fixed effects panel data model which does not require the normality assumption on the disturbances. This builds on the paper by Chen et al.(2010) who use U-statistics to test for sphericity of the variance-covariance matrix in statistics. The null of sphericity means that the variance-covariance matrix is proportional to the identity matrix. Rejecting the null means having cross-sectional dependence among the individual units of observation or heteroskedasticity or both. In empirical economic studies, individuals are affected by common shocks. For example, investors' decisions may be influenced by the way they interact with each other and also by common macro-economic shocks or public policies. These potentially cause cross-sectional dependence among the units.

In statistics, the $n \times n$ sample covariance matrix S_n is widely used for tests of sphericity since it is a consistent estimator for the variance-covariance matrix Σ_n. One could either use the likelihood ratio test, see Anderson (2003), or test the Frobenius norm of the difference between S_n and Σ_n, see John (1971,1972). However, with panel data sets where n the number of individuals is larger than the time series dimension of the data T, the sample covariance matrix becomes singular. This causes problems for the likelihood ratio test which is based on the inverse of S_n. Even when n is smaller than T, the sample covariance matrix S_n is ill-conditioned as shown in the Random Matrix Theory (RMT) literature. In fact, the eigenvalues of the sample covariance matrix S_n are no longer consistent for their population counterpart, see Johnstone (2001). Ledoit and Wolf (2004) show that the scaled Frobenius norm of S_n does not converge to that of Σ_n with $n/T \to c \in (0, \infty)$. As a result, John's test, see John (1971,1972), is no longer applicable. Hence, Ledoit and Wolf (2002) propose a new test for the null of sphericity which could be applied even when n is relatively as large as T. However, these statistical tests for raw data are not directly applicable to testing sphericity in panel data regressions since the disturbances are unobservable. Baltagi et al. (2011) extend the Ledoit and Wolf (2002)'s John test to the fixed effects panel data model and correct for the bias due to substituting within residuals for the actual disturbances. However, their test relies on the normality assumption and their simulation results show that the test has size distortion under non-normality of the disturbances.

To account for the possible "non-normality" of the disturbances as well as the "large n, small T" issues in testing the null of sphericity, Chen et al.(2010) propose a modified John test by constructing U-statistics of observable samples for estimating $\mathrm{tr}\Sigma_n$ and $\mathrm{tr}\Sigma_n^2$. Based on their work, this paper proposes a new test for the null of sphericity of the disturbances in a fixed effects regression panel data model. This test does not require the assumption of normality of the disturbances, and can be applied to the case where n is larger than T. The limiting distribution of this test statistic under the null is derived. Also, its finite sample properties are studied using Monte Carlo simulations.

The paper is organized as follows. Section 2 specifies the fixed effects panel data regression model and the assumptions required. Section 3 introduces the test statistic. Section 4 derives the limiting distribution of this test statistic under the null and discusses its power properties. Section 5 reports the results of Monte Carlo simulations, while Section 6 concludes. All the proofs and technical details can be found in an Appendix available upon request from the authors.

Notation: $||B|| = (\mathrm{tr}(B'B))^{1/2}$ is the Frobenius norm of a matrix B or the Euclidean norm of a vector B, and $\mathrm{tr}(B)$ is the trace of B. $\xrightarrow{d}$ denotes convergence in distribution and $\xrightarrow{p}$ denotes convergence in probability. For two matrices $B = (b_{ij})$ and $C = (c_{ij})$, we define $B \circ C = (b_{ij}c_{ij})$.

2 The Model and Assumptions

Consider the following fixed effects panel data regression model

$$y_{it} = \alpha + x_{it}'\beta + \mu_i + v_{it}, \text{ for } i = 1, 2, \ldots, n; \ t = 1, 2, \ldots, T, \tag{2.1}$$

where i indexes the cross-sectional dimension and t indexes the time series dimension. y_{it} is the dependent variable, x_{it} denotes the $k \times 1$ vector of exogenous regressors, and β is the corresponding $k \times 1$ vector of parameters. μ_i denotes the time-invariant individual effects which can be fixed or random and could be correlated with the regressors. Define the vector of disturbances $v_t = (v_{1t}, \ldots, v_{nt})'$ and its corresponding variance-covariance matrix Σ_n. The null hypothesis of interest is sphericity:

$$H_0 : \Sigma_n = \sigma_v^2 I_n \quad \text{vs} \quad H_1 : \Sigma_n \neq \sigma_v^2 I_n. \tag{2.2}$$

The alternative hypothesis allows cross-sectional dependence or heteroskedasticity or both.

For the panel data regression model, v_{it} is unobserved, and the test statistic is based upon consistent estimates of variance-covariance matrix, denoted by S_n or its correlation coefficients matrix counterpart, see Breusch and Pagan (1980). Baltagi et al. (2011) extend the Ledoit and Wolf (2002) test to a fixed effects panel data model with large n and large T. They show that the noise resulting from using within residuals rather than the actual disturbances accumulates and causes bias for the proposed test statistic. However, their simulations show that their test is oversized under non-normality of the disturbances. This paper extends Chen et al. (2010) to test the null of sphericity of the variance-covariance matrix of the disturbances in a fixed effects panel data regression model without assuming normality of the disturbances. We use the within residuals which are given by

$$\hat{v}_{it} = \tilde{y}_{it} - \tilde{x}_{it}'\tilde{\beta} = v_{it} - \bar{v}_{i.} - \tilde{x}_{it}'\left(\tilde{\beta} - \beta\right), \tag{2.3}$$

where $\tilde{x}_{it} = x_{it} - \bar{x}_{i.}$ and $\bar{x}_{i.} = \frac{1}{T}\sum_{t=1}^{T} x_{it}$. Similarly, $\tilde{y}_{it} = y_{it} - \bar{y}_{i.}$, $\bar{y}_{i.} = \frac{1}{T}\sum_{t=1}^{T} y_{it}$, and $\bar{v}_{i.} = \frac{1}{T}\sum_{t=1}^{T} v_{it}$. The within estimator of β is given by $\tilde{\beta} = \left(\sum_{t=1}^{T}\sum_{i=1}^{n} \tilde{x}_{it}\tilde{x}_{it}'\right)^{-1}$ $\left(\sum_{t=1}^{T}\sum_{i=1}^{n} \tilde{x}_{it}\tilde{y}_{it}\right)$. Let $\tilde{y}_t = (\tilde{y}_{1t}, \ldots, \tilde{y}_{nt})'$, $\hat{v}_t = (\hat{v}_{1t}, \ldots, \hat{v}_{nt})'$, $\bar{v}_{.} = (\bar{v}_{1.}, \ldots, \bar{v}_{n.})'$, and $\tilde{x}_t = (\tilde{x}_{1t}, \ldots, \tilde{x}_{nt})$. The within residuals can be rewritten in matrix form as $\hat{v}_t = v_t - \bar{v}_{.} - \tilde{x}_t'\left(\tilde{\beta} - \beta\right)$. To facilitate our analysis, we require the following assumptions:

Assumption 1 *The $n \times 1$ vectors $v_1, v_2, \ldots, v_T$ are independent and identically distributed (i.i.d.) with mean vector 0 and covariance matrix $\Sigma_n = \Gamma\Gamma'$, where Γ is an $n \times m$ $(m \leq \infty)$ matrix, v_t can be written as $v_t = \Gamma Z_t$, where $Z_t = (z_{t1}, \ldots, z_{tm})$ are i.i.d.random vectors with mean vector 0 and covariance matrix I_m. We also assume that each v_{it}, for $i = 1, \ldots, n$ has uniformly bounded 8th moment and there exists a finite constant Δ such that $E(z_{1l}^4) = 3 + \Delta$, for $l = 1, \ldots, m$.*

Assumption 2 *The regressors x_{it}, $i = 1, ..., n$, $t = 1, ..., T$ are independent of the idiosyncratic disturbances v_{it}, $i = 1, \ldots, n$, $t = 1, \ldots, T$. The regressors x_{it} have finite fourth moments: $E[||x_{it}||^4] \leq K < \infty$, where K is a positive constant.*

Assumption 3 *As $(n, T) \to \infty$, $tr(\Sigma_n^2) \to \infty$, $tr(\Sigma_n^4)/tr^2(\Sigma_n^2) \to 0$.*

The asymptotics follow the framework employed by Chen et al. (2010). Assumption 3 requires $\text{tr}(\Sigma_n^4)$ to grow at a slower rate than $\text{tr}^2(\Sigma_n^2)$. This assumption is flexible. In fact, if all the eigenvalues of Σ_n are bounded away from zero and infinity, $\text{tr}(\Sigma_n^4)/\text{tr}^2(\Sigma_n^2) \to 0$, is always true for any n as $n \to \infty$. Moreover, this assumption allows n to be much larger than T, which is more suitable for micro-panel data.

3 J_u Test

For testing the null hypothesis (2.2), the test statistic is based on the scaled distance measure between $\sigma_v^{-2}\Sigma_n$ and I_n:

$$U_0 = \frac{1}{n}\text{tr}\left[S_n\left(\frac{1}{n}\text{tr}S_n\right)^{-1} - I_n\right]^2 = \left(\frac{1}{n}\text{tr}S_n\right)^{-2}\left(\frac{1}{n}\text{tr}(S_n^2)\right) - 1, \tag{3.1}$$

where S_n is the $n \times n$ sample covariance matrix and I_n is an $n \times n$ identity matrix. With the normality assumption, John (1972) shows that for fixed n, and as $T \to \infty$:

$$\frac{nT}{2}U_0 \xrightarrow{d} \chi^2_{n(n+1)/2-1}. \tag{3.2}$$

But when n goes to infinity, the test statistic diverges. Ledoit and Wolf (2002) propose a modified test statistic under the null, as $(n,T)\to\infty$ and $n/T\to c\in(0,\infty)$:

$$TU_0 - n \xrightarrow{d} N(1,4). \tag{3.3}$$

Define $J_0 = \frac{TU_0-n}{2} - \frac{1}{2}$, then under the null $J_0 \xrightarrow{d} N(0,1)$. However, this test cannot be used directly in a fixed effects panel data regression model. The raw data sample covariance matrix S_n is replaced by its counterpart $\hat{S}_n = \frac{1}{T}\sum_{t=1}^{T}\hat{v}_t\hat{v}_t'$, where $\hat{v}_t$ is the within residual given by (2.3). The residual-based $\hat{U}_0$ is defined as $\hat{U}_0 = \left(\frac{1}{n}\mathrm{tr}\hat{S}_n\right)^{-2}\frac{1}{n}\mathrm{tr}\left(\hat{S}_n^2\right) - 1$ and the corresponding residual-based J_0 test is $\hat{J}_0 = \frac{T\hat{U}_0-n}{2} - \frac{1}{2}$. Baltagi et al. (2011) propose a bias correction:

$$J_{BFK} = \hat{J}_0 - \frac{n}{2(T-1)}. \tag{3.4}$$

They show that in a fixed effects panel data regression, as $(n,T)\to\infty$ and $n/T\to c$, $J_{BFK}\xrightarrow{d} N(0,1)$ under the null. However, their result relies on the normality assumption of v_t. Without the normality assumption, the bias-corrected John test is not robust, see the simulations in Baltagi et al. (2011).

Chen et al. (2010) propose a new test statistic for the sphericity of the variance-covariance matrix of the disturbances without the normality assumption and under much relaxed conditions where n could be much larger than T. They construct the U-statistics for estimating $\mathrm{tr}\Sigma_n$ and $\mathrm{tr}\Sigma_n^2$. Following their framework, we propose a residual-based test statistic for testing the null of sphericity described in (2.2) in a fixed effects panel data model. Define

$$\hat{M}_{1,T} = \frac{1}{T}\sum_{t=1}^{T}\hat{v}_t'\hat{v}_t; \quad \hat{M}_{2,T} = \frac{1}{C_T^2}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\hat{v}_t'\hat{v}_s; \quad \hat{M}_{3,T} = \frac{1}{C_T^2}\sum_{t\neq s}^{T}\sum_{s=1}^{T}(\hat{v}_t'\hat{v}_s)^2;$$

$$\hat{M}_{4,T} = \frac{1}{C_T^3}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\hat{v}_t'\hat{v}_s\hat{v}_s'\hat{v}_\tau; \quad \hat{M}_{5,T} = \frac{1}{C_T^4}\sum_{t\neq s\neq\tau\neq\eta}^{T}\sum_{s\neq\tau\neq\eta}^{T}\sum_{\tau\neq\eta}^{T}\sum_{\eta=1}^{T}\hat{v}_t'\hat{v}_s\hat{v}_\tau'\hat{v}_\eta,$$

where $C_T^i = T!/(T-i)!$. Also, let $\hat{R}_1 = \hat{M}_{1,T} - \hat{M}_{2,T}$ and $\hat{R}_2 = \hat{M}_{3,T} - 2\hat{M}_{4,T} + \hat{M}_{5,T}$.

If we observe the true v_t, then R_1, R_2 and $M_{j,T}$, for $j = 1,2,3,4,5$ are obtained similarly by replacing $(\hat{v}_t, \hat{v}_s, \hat{v}_\tau, \hat{v}_\eta)$ with $(v_t, v_s, v_\tau, v_\eta)$. R_1 and R_2 are unbiased estimators for $\mathrm{tr}\Sigma_n$ and $\mathrm{tr}\Sigma_n^2$, respectively. The scaled distance measure between $\sigma_v^{-2}\Sigma_n$ and I_n is given by $U_T = \frac{nR_2}{R_1^2} - 1$. Define $A = \Gamma'\Gamma$ and $\psi^2 = \frac{4}{T^2} + \frac{8}{T}\mathrm{tr}\left[\left(\frac{\Sigma_n^2}{\mathrm{tr}(\Sigma_n^2)} - \frac{\Sigma_n}{\mathrm{tr}(\Sigma_n)}\right)^2\right] + \frac{4\Delta}{T}\mathrm{tr}\left[\left(\frac{A^2}{\mathrm{tr}(\Sigma_n^2)} - \frac{A}{\mathrm{tr}(\Sigma_n)}\right) \circ \left(\frac{A^2}{\mathrm{tr}(\Sigma_n^2)} - \frac{A}{\mathrm{tr}(\Sigma_n)}\right)\right]$. Chen et al. (2010) show that as $(n,T) \to \infty$:

$$\psi^{-1}\left[\left(\frac{U_T+1}{n}\right)\left(\frac{\mathrm{tr}^2(\Sigma_n)}{\mathrm{tr}\,(\Sigma_n^2)}\right) - 1\right] \xrightarrow{d} N(0,1). \tag{3.5}$$

Let $J_{CZZ} = \frac{TU_T}{2}$, then under the null $J_{CZZ} \xrightarrow{d} N(0,1)$. Following this framework, we propose the following test statistic:

$$J_u = \frac{T}{2}\hat{U}_T = \frac{T}{2}\left(n\frac{\hat{R}_2}{\hat{R}_1^2} - 1\right). \tag{3.6}$$

J_u is the residual-based statistic corresponding to J_{CZZ}. There are two important issues to be considered. First, whether the residual-based $\hat{R}_1$ and $\hat{R}_2$ are consistent estimates for $\mathrm{tr}\Sigma_n$ and $\mathrm{tr}\Sigma_n^2$ under the null, respectively. Second, the asymptotics of the proposed test need to be derived. Both concerns are tackled in the next Section.

4 Asymptotics of the J_u Test

In this Section, we prove that, under the null, $\frac{1}{n}\hat{R}_1$ and $\frac{1}{n}\hat{R}_2$ are consistent estimators for $\frac{1}{n}\mathrm{tr}\Sigma_n = \sigma_v^2$ and $\frac{1}{n}\mathrm{tr}\Sigma_n^2 = \sigma_v^4$, respectively. Next, we show J_u converges to $N(0,1)$ under the null and we discuss its power properties. To examine the asymptotics of J_u, we rewrite it as

$$J_u = J_{CZZ} + (J_u - J_{CZZ}) = J_{CZZ} + \frac{T\left(\hat{U}_T - U_T\right)}{2}. \tag{4.1}$$

The first term J_{CZZ} is asymptotically standard normal under the null. The second term $J_u - J_{CZZ}$ is the scaled difference between the residual-based $\hat{U}_T$ and the true U_T. From

Section 3, this difference can be rewritten as follows:

$$J_u - J_{CZZ} = \frac{T}{2}\left[\left(\frac{1}{n}\hat{R}_2\right)\left(\frac{1}{n}R_1\right)^2 - \left(\frac{1}{n}R_2\right)\left(\frac{1}{n}\hat{R}_1\right)^2\right]\left(\frac{1}{n}\hat{R}_1\right)^{-2}\left(\frac{1}{n}R_1\right)^{-2}. \quad (4.2)$$

From equation (4.2), it is clear that this term depends upon the two differences: $\frac{1}{n}\hat{R}_1 - \frac{1}{n}R_1$ and $\frac{1}{n}\hat{R}_2 - \frac{1}{n}R_2$. Their asymptotic behavior is given in the following propositions:

Proposition 1 *Under Assumptions 1-2 and the null, (1)* $\frac{1}{n}\hat{M}_{1,T} = \frac{1}{n}M_{1,T} - \frac{\sigma_v^2}{T} + O_p\left(\frac{1}{T\sqrt{n}}\right)$*; (2)* $\frac{1}{n}\hat{M}_{2,T} = \frac{1}{n}M_{2,T} - \frac{\sigma_v^2}{T} + O_p\left(\frac{1}{T\sqrt{n}}\right)$*; (3)* $\frac{1}{n}\hat{R}_1 - \frac{1}{n}R_1 = O_p\left(\frac{1}{nT}\right)$.

Proposition 2 *Under Assumptions 1-2 and the null, (1)* $\frac{1}{n}\hat{M}_{3,T} = \frac{1}{n}M_{3,T} + \frac{n-2T-6}{T^2}\sigma_v^4 + O_p\left(\frac{\sqrt{n}}{T^2}\right) + O_p\left(\frac{1}{T\sqrt{T}}\right) + O_p\left(\frac{1}{T\sqrt{n}}\right)$*; (2)* $\frac{1}{n}\hat{M}_{4,T} = \frac{1}{n}M_{4,T} + \frac{n-T-5}{T^2}\sigma_v^4 + O_p\left(\frac{1}{T\sqrt{n}}\right) + O_p\left(\frac{\sqrt{n}}{T^2}\right) + O_p\left(\frac{1}{T\sqrt{T}}\right)$*; (3)* $\frac{1}{n}\hat{M}_{5,T} = \frac{1}{n}M_{5,T} + \frac{n-4}{T^2}\sigma_v^4 + O_p\left(\frac{\sqrt{n}}{T^2}\right)$*; (4)* $\frac{1}{n}\hat{R}_2 - \frac{1}{n}R_2 = O_p\left(\frac{1}{T^2}\right) + O_p\left(\frac{1}{nT}\right) + O_p\left(\frac{1}{T\sqrt{nT}}\right)$.

Propositions 1 and 2 show that the differences $\frac{1}{n}\hat{R}_1 - \frac{1}{n}R_1$ and $\frac{1}{n}\hat{R}_2 - \frac{1}{n}R_2$ vanish as $(n,T) \to \infty$. Therefore, since $\frac{1}{n}R_1 \xrightarrow{p} \sigma_v^2$ and $\frac{1}{n}R_2 \xrightarrow{p} \sigma_v^4$, we conclude that $\frac{1}{n}\hat{R}_1$ and $\frac{1}{n}\hat{R}_2$ are consistent estimates for σ_v^2 and σ_v^4 respectively. The following corollary gives these conclusions:

Corollary 1 *Under Assumptions 1-2 and the null, as* $(n,T) \to \infty$*, (1)* $\frac{1}{n}\hat{R}_1 \xrightarrow{p} \sigma_v^2$*; (2)* $\frac{1}{n}\hat{R}_2 \xrightarrow{p} \sigma_v^4$.

Note that $\frac{1}{n}\hat{R}_2$ is a consistent estimator of σ_v^4 under the null with large n and large T. However, $\frac{1}{n}\mathrm{tr}\hat{S}_n^2$ is not consistent, see Baltagi et al. (2011).

Proposition 3 *Under Assumptions 1-2 and the null,* $\frac{T\left(\hat{U}_T - U_T\right)}{2} = O_p\left(\frac{1}{n}\right) + O_p\left(\frac{1}{T}\right) + O_p\left(\frac{1}{\sqrt{nT}}\right)$.

Propositions 1, 2 and 3 give the asymptotics of the bias term $J_u - J_{CZZ}$. Compared with the statistic based on raw data, the test statistic based on the within residuals, defined below equation (2.3), can be expressed by $\hat{v}_t = v_t - \bar{v}_{.} - \tilde{x}_t'\left(\tilde{\beta} - \beta\right)$. This has the additional terms $\bar{v}_{\cdot}$ and $\tilde{x}_t'\left(\tilde{\beta}-\beta\right)$. These two terms can be regarded as extra noise resulting from the regression. Based on equations (4.1) and (4.2), the extra noise $\frac{T\left(\hat{U}_T-U_T\right)}{2}$ depends upon $\frac{1}{n}\hat{R}_1-\frac{1}{n}R_1=\frac{1}{n}(\hat{M}_{1,T}-M_{1,T})-\frac{1}{n}(\hat{M}_{2,T}-M_{2,T})$ and $\frac{1}{n}\hat{R}_2-\frac{1}{n}R_2=\frac{1}{n}(\hat{M}_{3,T}-M_{3,T})-\frac{2}{n}(\hat{M}_{4,T}-M_{4,T})+\frac{1}{n}(\hat{M}_{5,T}-M_{5,T})$. Hence the magnitude of $\frac{T\left(\hat{U}_T-U_T\right)}{2}$ depends upon how $\bar{v}_{\cdot}$ and $\tilde{x}_t'(\tilde{\beta}-\beta)$ accumulate in $\left(\frac{1}{n}\hat{M}_{j,T}-\frac{1}{n}M_{j,T}\right)$, for $j=1,2,3,4,5$. Note that $\bar{v}_{\cdot}$ is an n dimensional vector, although each element of $\bar{v}_{\cdot}$ is $O_p(\frac{1}{\sqrt{T}})$, $\bar{v}_{\cdot}$ may still accumulate in the above five terms as $(n,T)\to\infty$, depending upon the relative speed of n and T. $\tilde{x}_t'(\tilde{\beta}-\beta)$ is $O_p\left(\frac{1}{\sqrt{nT}}\right)$ which is related to both n and T. We may expect its convergence speed $\frac{1}{\sqrt{nT}}$ to be fast enough so that $\tilde{x}_t'(\tilde{\beta}-\beta)$ vanishes as $(n,T)\to\infty$. More specifically, Proposition 2 shows the leading terms of $\frac{1}{n}\hat{M}_{j,T}-\frac{1}{n}M_{j,T}$, for $j=3,4,5$ will not vanish if $\frac{n}{T^2}$ does not converge to zero. These terms are caused by the accumulation of $\bar{v}_{\cdot}$. However, due to the subtraction formulation of the test statistic, the leading terms cancel each other in both $\frac{1}{n}\hat{R}_1-\frac{1}{n}R_1$ and $\frac{1}{n}\hat{R}_2-\frac{1}{n}R_2$. Similar cancellations occur for other terms which are $O_p\left(\frac{\sqrt{n}}{T^2}\right)$, $O_p\left(\frac{1}{T\sqrt{n}}\right)$ and $O_p\left(\frac{1}{T\sqrt{T}}\right)$ since their expressions are exactly the same. These cancellations lead us to $\frac{1}{n}\hat{R}_1-\frac{1}{n}R_1=O_p\left(\frac{1}{nT}\right)$ and $\frac{1}{n}\hat{R}_2-\frac{1}{n}R_2=O_p\left(\frac{1}{T^2}\right)+O_p\left(\frac{1}{nT}\right)+O_p\left(\frac{1}{T\sqrt{nT}}\right)$, and consequently $\frac{T\left(\hat{U}_T-U_T\right)}{2}=O_p(\frac{1}{n})+O_p(\frac{1}{T})+O_p\left(\frac{1}{\sqrt{nT}}\right)$. Therefore, $J_u-J_{CZZ}\xrightarrow{p}0$ as $(n,T)\to\infty$ and we do not need to correct the bias in the fixed effects panel data regression model. This result is based on our detailed calculation of how $\bar{v}_{\cdot}$ and $\tilde{x}_t'(\tilde{\beta}-\beta)$ are accumulating in $\left(\frac{1}{n}\hat{M}_{j,T}-\frac{1}{n}M_{j,T}\right)$, for $j=1,2,3,4,5$ and the special formulation of J_{CZZ}. As discussed above, the convergence of J_u is given by the following theorem:

Theorem 4 *Under Assumptions 1-3 and the null, in the fixed effects panel data regression model (2.1), as* $(n,T)\to\infty$

$$J_u\xrightarrow{d}N(0,1). \tag{4.3}$$

Under the alternative, the limiting distribution of J_u is the same as (3.5) if $\frac{T\left(\hat{U}_T-U_T\right)}{2}$ vanishes as $(n,T)\to 0$. Similar to Chen et al. (2010), we consider an alternative: H_1 : $\Sigma_n=\left(\sigma_i\sigma_j\rho^{|j-i|}\right)_{n\times n}$, where $\rho\in(-1,1)$ and $\rho\neq 0$. $\sigma_l^2=\text{var}(v_{lt})$, which is uniformly bounded away from infinity and zero, for $l = 1, \ldots, n$. Under this alternative, we can show that $J_u - J_{CZZ} = o_p(1)$, which in turn implies that J_u and J_{CZZ} have the same power properties. Define $\delta_{1,T} = 1 - \frac{\text{tr}^2(\Sigma_n)}{n\text{tr}(\Sigma_n^2)}$ and $\delta_{2,T} = \text{tr}\left[\left(\frac{\Sigma_n^2}{\text{tr}(\Sigma_n^2)} - \frac{\Sigma_n}{\text{tr}(\Sigma_n)}\right)^2\right]$. One can show that $T\delta_{1,T} \to \infty$ and $\delta_{2,T}/(T\delta_{1,T}^2) \to 0$ as $(n,T) \to \infty$. This satisfies the conditions of Theorem 4 in Chen et al. (2010). By using this Theorem, the corresponding power function $P(J_u \geq z_\alpha | \Sigma_n = (\sigma_i \sigma_j \rho^{|j-i|})_{n \times n}) \to 1$, as $(n,T) \to \infty$, where z_α is the upper quantile of $N(0,1)$. Let us consider a special case under this alternative. More specifically, assume that $\Delta = 0$, $\sigma_i = \sigma_j = \sigma_v$ for any (i,j) and $T/n \to 0$ as $(n,T) \to \infty$. It follows that $\psi^{-1} \to \frac{T}{2}$ and $(1-\rho^2)\, J_u - T\rho^2/2 \xrightarrow{d} N(0,1)$.

5 Monte Carlo Simulations

We conduct Monte Carlo experiments to assess the empirical size and power of the J_u test proposed in this paper. We follow the design of Baltagi et al. (2011) and assume homoskedasticity on the remainder error term. In this case, the J_u test becomes a test for cross-sectional dependence. We also report the performance of J_{BFK} proposed by Baltagi et al. (2011) for comparison purposes.

5.1 Experiment Design

Consider the following data-generating process:

$$y_{it} = \alpha + \beta x_{it} + \mu_i + v_{it}, \quad i = 1, \ldots, n; \; t = 1, \ldots, T, \tag{5.1}$$

$$x_{it} = \lambda x_{i,t-1} + \mu_i + \eta_{it}, \tag{5.2}$$

where μ_i is the fixed effects and v_{it} is the idiosyncratic error, $\eta_{it} \sim i.i.d.\ N(\phi_\eta, \sigma_\eta^2)$. The regressor x_{it} is allowed to be correlated with the μ_i's. This follows the design by Im et al. (1999).

To study the power of the tests, we consider two different types of cross-sectional dependence models: a factor model and a spatial model. For the factor model, see Pesaran (2004), Pesaran and Tosetti (2011), Baltagi et al. (2011), we assume:

$$v_{it} = \gamma_i f_t + \epsilon_{it}, \tag{5.3}$$

where f_t $(t = 1, \ldots, T)$ are the factors and γ_i $(i = 1, \ldots, n)$ are the loadings. For the spatial model, we consider a first-order spatial autocorrelation model SAR(1), see Anselin and Bera (1998) and Baltagi et al. (2003), given by:

$$v_{it} = \delta(0.5v_{i-1,t} + 0.5v_{i+1,t}) + \epsilon_{it}. \tag{5.4}$$

The ϵ_{it} in (5.3) and (5.4) are assumed to be $i.i.d.(0, \sigma_\epsilon^2)$ across individuals and over time. Under the null, we have $\gamma_i = 0$ and $\delta = 0$.

Under the null, the v_{it} comes from some $i.i.d.$ distribution across individuals and over time with mean zero and variance σ_v^2. These are not necessarily normally distributed. For models (5.1) and (5.2), we set $\alpha = 1$ and $\beta = 2$; μ_i is drawn from $i.i.d.$ $N(\phi_\mu, \sigma_\mu^2)$ with $\phi_\mu = 0$ and $\sigma_\mu^2 = 0.25$. We also set $\lambda = 0.7$, $\phi_\eta = 0$ and $\sigma_\eta^2 = 1$. For models (5.3) and (5.4), $\gamma_i \sim i.i.d.$ $U(-0.5, 0.55)$; f_t is set to be $i.i.d.$ $N(0, 1)$ and $\delta = 0.4$. Various distributions are considered in generating the model errors, v_{it} in (5.1) and ϵ_{it} in (5.3) and (5.4) are assumed be normal, lognormal, gamma, chi-squared with mean zero and variance 0.5.

The Monte Carlo experiments are conducted for $n = 20, 40, 60, 80, 100, 200, 400$ and $T = 20, 40, 60, 80$. We perform $1,000$ replications to compute the J_u and J_{BFK} test statistics. We conduct the tests at the positive one-sided 5% nominal significance level to obtain the empirical size.

5.2 Results

Table 1 gives the empirical size of the J_u and J_{BFK} tests allowing v_{it} to be generated from different distributions. When the disturbances are normally distributed, the size of J_u and J_{BFK} are both close to 5%, which is consistent with the theoretical results. The rest of Table 1 shows the results with v_{it} coming from alternative non-normal distributions. The size of J_u is close to 5% when n and T are large; for small n or small T, it is slightly oversized. However, J_{BFK} is no longer robust to non-normality and suffers from size distortions.

Table 2 presents the size adjusted power of the tests under the alternative specification of a factor model. Both tests have size adjusted power that is almost 1 when n and T are large with v_{it} normally distributed. For small n and small T, the size adjusted power of J_u works as well as J_{BFK}. Note that the size adjusted power of J_{BFK} is quite good even when n is a lot larger than T for the normal distribution scenario. However, for non-normal distributions, the size adjusted power of J_u is 1 as n and T become large; and it is larger than the size adjusted power of J_{BFK} for all (n, T) combinations.

Table 3 reports the size adjusted power of both tests under the alternative specification of SAR(1). The results are similar to the factor model. J_u works as well as J_{BFK} for the normal distribution scenario, but better for all combinations of n and T for non-normal distribution scenarios.

6 Conclusion

Though the John test proposed by Baltagi et al. (2011) has been shown to perform well for a large panel data regression model with fixed effects, it relies heavily on the normality assumption. This paper proposes a new test, J_u, for the null of sphericity of the disturbances which does not rely on the normality assumption. Instead of $n/T \to c$, we allow n to be a larger order of T which is consistent with micro-panel data sets with "large n and small T"
.

Acknowledgments

Essay I is based on the paper of Baltagi, Kao and Peng (2015).

References

[1] Anderson, T. W., 2003. An Introduction to Multivariate Statistical Analysis, Wiley.

[2] Anselin, L., and Bera, A. K., 1998. Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics. Handbook of Applied Economic Statistics, Amman Ullah and David E. A. Giles, eds., Marcel Dekker, New York.

[3] Bai, Z.D., and Saranadasa, H., 1996. Effect of High Dimension: By an Example of a Two Sample Problem. Statistical Sinica. 6, 311-329.

[4] Baltagi, B. H., Song, S. H., and Koh, W., 2003. Testing Panel Data Regression Models with Spatial Error Correlation. Journal of Econometrics. 117, 123-150.

[5] Baltagi, B. H., Feng, Q., and Kao, C., 2011. Testing for Sphericity in a Fixed Effects Panel Data Model. The Econometrics Journal. 14, 25-47.

[6] Baltagi, B. H., Kao, C., and Peng, B., 2015. On Testing for Sphericity with Non-normality in a Fixed Effects Panel Data Model. Statistics & Probability Letters. 98, 123-130.

[7] Breusch, T. S., and Pagan, A. R., 1980. The Lagrange Multiplier Test and Its Application to Model Specification in Econometrics. Review of Economic Studies. 47, 239-254.

[8] Chen, S., Zhang, L., and Zhong, P., 2010. Tests for High-dimensional Covariance Matrices. Journal of the American Statistical Association. 105, 810-819.

[9] Im, K. S., Ahn, S. C., Schmidt, P., and Wooldridge, J. M., 1999. Efficient Estimation of Panel Data Models with Strictly Exogenous Explanatory Variables. Journal of Econometrics. 93, 177-201.

[10] John, S., 1971. Some Optimal Multivariate Test. Biometrika. 58, 123-127.

[11] John, S., 1972. The Distribution of a Statistic Used for Testing Sphericity of Normal Distributions. Biometrika. 59, 169–173.

[12] Johnstone, I., 2001. On the Distribution of the Largest Principal Component. Annals of Statistics. 29, 295-327.

[13] Ledoit, O., and 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.

[14] Ledoit, O., and Wolf, M., 2004. A Well-conditioned Estimator for Large-Dimensional Covariance Matrices. Journal of Multivariate Analysis. 88, 365-411.

[15] Pesaran, M. H., 2004. General Diagnostic Test for Cross Section Dependence in Panels. CESifo Working Paper, No.1229, University of Cambridge.

[16] Pesaran, M. H., and Tosetti, E., 2011. Large Panels with Common Factors and Spatial Correlations. Journal of Econometrics. 161, 182-202.

Table 1: Size of Tests

		n						
Normal Errors	T	20	40	60	80	100	200	400
J_u	20	6.4	7.1	6.5	8.0	7.8	7.2	6.3
	40	5.6	7.0	4.9	4.7	5.8	5.9	6.0
	60	6.7	7.1	6.5	5.7	5.4	5.5	5.9
	80	5.2	5.6	4.9	7.1	5.8	5.1	4.4
J_{BFK}	20	6.4	6.7	5.6	6.6	6.9	6.1	5.8
	40	5.8	6.7	5.0	4.9	5.6	6.5	5.1
	60	6.5	6.6	6.7	5.9	4.8	5.0	5.9
	80	5.0	5.3	4.6	6.7	6.1	4.7	4.7
Gamma Errors								
J_u	20	7.2	6.8	8.0	7.3	5.5	7.7	6.9
	40	7.1	5.6	7.0	6.8	5.1	5.3	5.1
	60	7.4	7.0	5.2	6.0	6.0	5.4	5.1
	80	6.2	5.9	5.6	5.2	6.5	5.3	5.6
J_{BFK}	20	16.0	17.8	19.4	20.1	18.6	21.3	18.3
	40	17.4	17.5	21.0	19.3	17.0	19.7	18.2
	60	19.5	19.9	16.4	19.3	18.0	18.6	18.5
	80	18.5	18.8	17.7	18.5	19.3	18.2	18.8
Lognormal Errors								
J_u	20	9.3	7.9	6.8	7.6	7.2	6.2	7.1
	40	8.0	8.0	5.7	6.3	6.9	6.7	6.4
	60	8.3	6.4	6.6	6.5	5.9	5.3	5.4
	80	7.0	6.3	7.1	6.0	5.8	5.0	6.0
J_{BFK}	20	27.1	26.9	27.9	27.8	28.5	28.0	29.0
	40	26.5	30.2	27.0	29.0	28.3	29.7	28.7
	60	25.4	27.1	29.9	29.7	30	30.3	30.9
	80	26.2	26.7	29.0	28.1	28.4	32.0	30.1
Chi-squared Errors								
J_u	20	8.4	8.2	7.6	7.9	7.7	6.9	7.6
	40	8.6	6.8	6.6	6.3	5.0	7.3	6.2
	60	8.4	8.1	7.6	5.6	4.6	5.5	5.4
	80	8.0	6.4	7.1	7.3	4.9	6.5	6.0
J_{BFK}	20	26.6	26.2	28.7	29.8	30.7	29.6	31.9
	40	27.9	29.3	31.5	31.9	31.0	32.4	33.2
	60	30.6	33.5	33.6	31.6	32.3	32.5	28.9
	80	30.7	30.1	35.0	34.1	32.9	32.4	31.8

Notes: This table reports the size of J_u and J_{BFK} with different error distribution specification in a fixed effects panel data model without cross-sectional dependence among the errors. The tests are one-sided and are conducted at the 5% nominal significance level. We conduct the simulation with four distributions: normal, gamma, lognormal and chi-squared with mean 0, and variance 0.5.

Table 2: Size adjusted power of tests: factor model

		n						
Normal Errors	T	20	40	60	80	100	200	400
Ju	20	73.1	94.0	98.3	99.5	99.7	99.8	100
	40	95.6	99.8	99.9	100	100	100	100
	60	99.3	100	100	100	100	100	100
	80	99.8	100	100	100	100	100	100
J_{BFK}	20	73.4	94.7	98.3	99.5	99.9	99.9	100
	40	95.8	99.8	99.9	100	100	100	100
	60	99.4	100	100	100	100	100	100
	80	99.8	100	100	100	100	100	100
Gamma Errors								
J_u	20	68.2	93.3	97.1	99.1	99.6	100	100
	40	94.6	99.7	100	100	100	100	100
	60	99.1	100	100	100	100	100	100
	80	99.6	100	100	100	100	100	100
J_{BFK}	20	60.1	89.0	96.1	98.5	99.2	99.9	100
	40	91.5	99.5	99.9	100	100	100	100
	60	98.1	100	100	100	100	100	100
	80	99.2	100	100	100	100	100	100
Lognormal Errors								
Ju	20	68.1	91.6	98.2	99.2	99.4	99.9	100
	40	95.3	99.7	100	100	100	100	100
	60	99.6	100	100	100	100	100	100
	80	99.9	100	100	100	100	100	100
J_{BFK}	20	48.7	85.2	95.5	97.6	98.3	99.9	100
	40	88.3	99	100	100	100	100	100
	60	97.9	100	100	100	100	100	100
	80	99.5	100	100	100	100	100	100
Chi-squared Errors								
J_u	20	70.1	90.3	98	98.9	99.4	100	100
	40	94.1	100	100	100	100	100	100
	60	98.6	100	100	100	100	100	100
	80	99.6	100	100	100	100	100	100
J_{BFK}	20	53.8	80.4	93.5	97.8	98.3	100	100
	40	84.8	99.3	100	100	100	100	100
	60	96.1	100	100	100	100	100	100
	80	98.6	100	100	100	100	100	100

Notes: This table computes the size adjusted power for a factor structure model that allows for cross-sectional dependence in the error. We conduct the simulation with four distributions: normal, gamma, lognormal and chi-squared with mean 0, and variance 0.5.

Table 3: Size adjusted power of tests: SAR(1) model

		n						
Normal Errors	T	20	40	60	80	100	200	400
J_u	20	81.4	83.7	86.0	82.1	83.3	84.8	88.0
	40	99.9	99.9	100	100	100	100	100
	60	100	100	100	100	100	100	100
	80	100	100	100	100	100	100	100
J_{BFK}	20	82	87.1	89.7	87.5	86.4	88.3	90.6
	40	100	99.9	100	100	100	100	100
	60	100	100	100	100	100	100	100
	80	100	100	100	100	100	100	100
Gamma Errors								
J_u	20	75.1	81.6	84.3	84.1	87.3	85.0	86.2
	40	99.9	100	100	100	100	100	100
	60	100	100	100	100	100	100	100
	80	100	100	100	100	100	100	100
J_{BFK}	20	74	78.3	82.2	83.3	83.8	84.1	84.9
	40	99.9	100	99.9	100	100	100	100
	60	100	100	100	100	100	100	100
	80	100	100	100	100	100	100	100
Lognormal Errors								
J_u	20	71.4	80.2	83.5	83.5	85.2	87.8	88.6
	40	99.9	100	100	100	100	100	100
	60	100	100	100	100	100	100	100
	80	100	100	100	100	100	100	100
J_{BFK}	20	61.4	72.3	79.8	80.4	80.1	86.1	84.8
	40	99.4	100	99.9	99.9	100	99.9	100
	60	100	100	100	100	100	100	100
	80	100	100	100	100	100	100	100
Chi-squared Errors								
J_u	20	74.2	79.5	83.6	84.2	84.8	86.3	84.5
	40	99.7	100	99.7	100	100	100	100
	60	100	100	100	100	100	100	100
	80	100	100	100	100	100	100	100
J_{BFK}	20	65.6	70.3	79.5	77.9	76.7	82.6	78.4
	40	99.4	99.7	99.7	100	100	100	100
	60	100	100	100	100	100	100	100
	80	100	100	100	100	100	100	100

Notes: This table computes the size adjusted power for a SAR(1) structure model that allows for cross-sectional dependence in the error. We conduct the simulation with four distributions: normal, gamma, lognormal and chi-squared with mean 0, and variance 0.5.

Appendix

This appendix includes all the proofs of the Propositions and Theorems in the text, and it also presents some useful Lemmas, which are frequently used in the proofs of the Propositions and Theorems.

In the fixed effects panel data regression model:

$$y_{it} = x'_{it}\beta + \mu_i + v_{it} \ (i = 1, \ldots, n, t = 1, \ldots, T).$$

$\tilde{\beta}$ is the within estimator and the within residual is $\hat{v}_{it} = \tilde{y}_{it} - \tilde{x}'_{it}\tilde{\beta}$, where $\tilde{y}_{it} = y_{it} - \bar{y}_{i.}$ and $\tilde{x}_{it} = x_{it} - \bar{x}_{i.}$. Define $\tilde{v}_{it} = v_{it} - \bar{v}_{i.}$, then the residuals $\hat{v}_{it} = \tilde{v}_{it} - \tilde{x}'_{it}(\tilde{\beta} - \beta)$ and in vector form we have $\hat{v}_t = \tilde{v}_t - \tilde{x}_t(\tilde{\beta} - \beta)$. Before we go to the proofs, we introduce the following lemmas which will be used frequently later.

Lemma 1 *For a random sequence Z_n, if $E(Z_n^2) = O(n^v)$, where v is a constant, then $Z_n = O_p(n^{v/2})$.*

It is Lemma 1 in Baltagi, Feng and Kao(2011), we put it here since we use it frequently.

Lemma 2 *Under Assumption 1, 2 and the null,*

(a) $\frac{1}{nT^2}\sum_{t=1}^{T}\sum_{i=1}^{n} v_{it}^2 = \frac{\sigma_v^2}{T} + O_p(\frac{1}{T\sqrt{nT}})$;

(b) $\frac{1}{nT^3}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n} v_{it}^2 v_{is}^2 = \frac{T-1}{T^2}\sigma_v^4 + O_p(\frac{1}{T^2\sqrt{n}})$;

(c) $\frac{1}{nT^3}\sum_{t=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it}^2 v_{jt}^2 = \frac{n-1}{T^2}\sigma_v^4 + O_p(\frac{1}{T^2\sqrt{T}})$;

(d) $n(\frac{1}{nT^2}\sum_{t=1}^{T}\sum_{i=1}^{n} v_{it}^2)^2 = \frac{n}{T^2}\sigma_v^4 + O_p(\frac{\sqrt{n}}{T^2\sqrt{T}})$;

(e) $\frac{1}{nT^3(T-1)}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{j\neq i}^{n}\sum_{i=1}^{n} v_{it}^2 v_{js}^2 = \frac{n-1}{T^2}\sigma_v^4 + O_p(\frac{1}{T^3})$.

Proof. Consider (a),

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

by using the fact $\frac{1}{\sqrt{nT}}(v_{it}^2-\sigma_v^2)=O_p(1)$.

For (b)

$$\frac{1}{nT^3}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n}v_{it}^2v_{is}^2=\frac{T-1}{T^2}\sigma_v^4+\frac{1}{nT^3}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n}(v_{it}^2v_{is}^2-\sigma_v^4)=\frac{T-1}{T^2}\sigma_v^4+O_p(\frac{1}{T^2\sqrt{n}}).$$

by using the fact $\frac{1}{\sqrt{nT}}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n}(v_{it}^2v_{is}^2-\sigma_v^4)=O_p(1)$.

For (c)

$$\frac{1}{nT^3}\sum_{t=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}v_{it}^2v_{jt}^2=\frac{n-1}{T^2}\sigma_v^4+\frac{1}{nT^3}\sum_{t=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}(v_{it}^2v_{jt}^2-\sigma_v^4)=\frac{n-1}{T^2}\sigma_v^4+O_p(\frac{1}{T^2\sqrt{T}}).$$

using $\frac{1}{n\sqrt{T}}\sum_{t=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}(v_{it}^2v_{jt}^2-\sigma_v^4)=O_p(1)$.

For (d)

$$\begin{aligned}n(\frac{1}{nT^2}\sum_{t=1}^{T}\sum_{i=1}^{n}v_{it}^2)^2=&[\frac{\sqrt{n}}{T}\sum_{t=1}^{T}\sum_{i=1}^{n}v_{it}^2]^2=[\frac{\sigma_v^2}{T^2\sqrt{n}}+\frac{1}{T^2\sqrt{n}}\sum_{t=1}^{T}\sum_{i=1}^{n}(v_{it}^2-\sigma_v^2)]^2\\=&[\frac{\sqrt{n}}{T}\sigma_v^2+O_p(\frac{1}{T\sqrt{T}})]^2=\frac{n}{T^2}\sigma_v^4+O_p(\frac{\sqrt{n}}{T^2\sqrt{T}}).\end{aligned}$$

For (e)

$$\begin{aligned}\frac{1}{nT^3(T-1)}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{j\neq i}^{n}\sum_{i=1}^{n}v_{it}^2v_{js}^2=&\frac{n-1}{T^2}\sigma_v^4+\frac{1}{nT^3(T-1)}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{j\neq i}^{n}\sum_{i=1}^{n}(v_{it}^2v_{js}^2-\sigma_v^4)\\=&\frac{n-1}{T^2}\sigma_v^4+O_p(\frac{1}{T^3}),\end{aligned}$$

by using the fact $\frac{1}{nT}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{j\neq i}^{n}\sum_{i=1}^{n}(v_{it}^2v_{js}^2-\sigma_v^4)=O_p(1)$. ■

The following lemma verifies the consistency propositions of $\frac{1}{n}R_1$ and $\frac{1}{n}R_2$.

Lemma 3 *Under Assumption 1, 2, and the null,*

(a) $\frac{1}{n}M_{1,T} = \sigma_v^2 + O_p(\frac{1}{\sqrt{nT}})$;

(b) $\frac{1}{n}M_{2,T} = O_p(\frac{1}{T\sqrt{n}})$;

(c) $\frac{1}{n}M_{3,T} = \sigma_v^4 + O_p(\frac{1}{T})$;

(d) $\frac{1}{n}M_{4,T} = O_p(\frac{1}{T\sqrt{n}}) + O_p(\frac{1}{T\sqrt{T}})$;

(e) $\frac{1}{n}M_{5,T} = O_p(\frac{1}{T^2})$;

(f) $\frac{1}{n}R_1 = \sigma_v^2 + O_p(\frac{1}{\sqrt{nT}})$;

(g) $\frac{1}{n}R_2 = \sigma_v^4 + O_p(\frac{1}{T})$.

Proof. Before the proofs, to simplify the notation, let's define

$$T_4 = T(T-1)(T-2)(T-3), \quad \sum_{t,s,\tau,\eta}^{T} = \sum_{t\neq s\neq \tau\neq \eta}^{T} \sum_{s\neq \tau\neq \eta}^{T} \sum_{\tau\neq \eta}^{T} \sum_{\eta=1}^{T}$$

First consider (a),

$$\begin{aligned}\frac{1}{n}M_{1,T} =& \frac{1}{n} \times \frac{1}{T}\sum_{t=1}^{T} v_t' v_t = \frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n} v_{it}^2 \\ =& \sigma_v^2 + \frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}(v_{it}^2 - \sigma_v^2) \\ =& \sigma_v^2 + O_p(\frac{1}{\sqrt{nT}}).\end{aligned}$$

Here we get the results by part (a) of Lemma 2.

For (b),

$$\begin{aligned}\frac{1}{n}M_{2,T} =& \frac{1}{n} \times \frac{1}{T(T-1)} \sum_{t \neq s}^{T} \sum_{s=1}^{T} v_t' v_s \\ =& \frac{1}{nT(T-1)} \sum_{t \neq s}^{T} \sum_{s=1}^{T} \sum_{i=1}^{n} v_{it} v_{is} \\ =& O_p(\frac{1}{T\sqrt{n}}).\end{aligned}$$

Since $E[(\frac{1}{nT(T-1)} \sum_{t \neq s}^{T} \sum_{s=1}^{T} \sum_{i=1}^{n} v_{it} v_{is})^2] = E[\frac{1}{n^2 T^2 (T-1)^2} \sum_{t \neq s}^{T} \sum_{s=1}^{T} \sum_{i=1}^{n} v_{it}^2 v_{is}^2] = O_p(\frac{1}{nT^2})$, we prove it.

For (c),

$$\begin{aligned}\frac{1}{n}M_{3,T} =& \frac{1}{n} \times \frac{1}{T(T-1)} \sum_{t \neq s}^{T} \sum_{s=1}^{T} (v_t' v_s)^2 \\ =& \frac{1}{nT(T-1)} \sum_{t \neq s}^{T} \sum_{s=1}^{T} \sum_{i=1}^{n} \sum_{j=1}^{n} v_{it} v_{is} v_{jt} v_{js} \\ =& \frac{1}{nT(T-1)} \sum_{t \neq s}^{T} \sum_{s=1}^{T} \sum_{i=1}^{n} v_{it}^2 v_{is}^2 + \frac{1}{nT(T-1)} \sum_{t \neq s}^{T} \sum_{s=1}^{T} \sum_{i \neq j}^{n} \sum_{j=1}^{n} v_{it} v_{is} v_{jt} v_{js} \\ =& \sigma_v^4 + O_p(\frac{1}{T\sqrt{n}}) + O_p(\frac{1}{T}).\end{aligned}$$

using part (b) of Lemma 2 and $\frac{1}{nT(T-1)} \sum_{t \neq s}^{T} \sum_{s=1}^{T} \sum_{i \neq j}^{n} \sum_{j=1}^{n} v_{it} v_{is} v_{jt} v_{js} = O_p(\frac{1}{T})$, hence

$$\frac{1}{n}M_{3,T} = \sigma_v^4 + O_p(\frac{1}{T\sqrt{n}}) + O_p(\frac{1}{T}) = \sigma_v^4 + O_p(\frac{1}{T}).$$

For (d),

$$
\begin{aligned}
\frac{1}{n}M_{4,T} =& \frac{1}{n} \times \frac{1}{T(T-1)(T-2)} \sum_{t\neq s\neq \tau}^{T} \sum_{s\neq \tau}^{T} \sum_{\tau=1}^{T} v_t' v_s v_s' v_\tau \\
=& \frac{1}{nT(T-1)(T-2)} \sum_{t\neq s\neq \tau}^{T} \sum_{s\neq \tau}^{T} \sum_{\tau=1}^{T} \sum_{i=1}^{n} \sum_{j=1}^{n} v_{it} v_{is} v_{js} v_{j\tau} \\
=& \frac{1}{nT(T-1)(T-2)} \sum_{t\neq s\neq \tau}^{T} \sum_{s\neq \tau}^{T} \sum_{\tau=1}^{T} \sum_{i=1}^{n} v_{it} v_{is}^2 v_{i\tau} \\
&+ \frac{1}{nT(T-1)(T-2)} \sum_{t\neq s\neq \tau}^{T} \sum_{s\neq \tau}^{T} \sum_{\tau=1}^{T} \sum_{i\neq j}^{n} \sum_{j=1}^{n} v_{it} v_{is} v_{js} v_{j\tau} \\
=& O_p(\frac{1}{T\sqrt{n}}) + O_p(\frac{1}{T\sqrt{T}}).
\end{aligned}
$$

For (e),

$$
\begin{aligned}
\frac{1}{n}M_{5,T} =& \frac{1}{n} \times \frac{1}{T(T-1)(T-2)(T-3)} \sum_{t,s,\tau,\eta}^{T} v_t' v_s v_\tau' v_\eta \\
=& \frac{1}{nT_4} \sum_{t,s,\tau,\eta}^{T} \sum_{i=1}^{n} v_{it} v_{is} v_{i\tau} v_{i\eta} + \frac{1}{nT_4} \sum_{t,s,\tau,\eta}^{T} \sum_{i\neq j}^{n} \sum_{j=1}^{n} v_{it} v_{is} v_{j\tau} v_{j\eta} \\
=& O_p(\frac{1}{T^2\sqrt{n}}) + O_p(\frac{1}{T^2}) = O_p(\frac{1}{T^2}).
\end{aligned}
$$

For (f) and (g), we can use the results of (a), (b), (c), (d), (e) and show that

$$
\frac{1}{n}R_1 = \frac{1}{n}M_{1,T} - \frac{1}{n}M_{2,T} = \sigma_v^2 + O_p(\frac{1}{\sqrt{nT}}) - O_p(\frac{1}{T\sqrt{n}}) = \sigma_v^2 + O_p(\frac{1}{\sqrt{nT}}).
$$

$$
\begin{aligned}
\frac{1}{n}R_2 = \frac{1}{n}M_{3,T} - \frac{2}{n}M_{4,T} + \frac{1}{n}M_{5,T} =& \sigma_v^4 + O_p(\frac{1}{T}) - 2(O_p(\frac{1}{T\sqrt{n}}) + O_p(\frac{1}{T\sqrt{T}})) + O_p(\frac{1}{T^2}) \\
=& \sigma_v^4 + O_p(\frac{1}{T}).
\end{aligned}
$$

■

A Proof of Proposition 1

A.1 Proof of part (1)

Proof. Recall in a panel data model $\tilde{y}_{it} = \tilde{x}_{it}^{'}\beta + \tilde{v}_{it}$, we have

$$\tilde{\beta} - \beta = (\sum_{t=1}^{T}\sum_{i=1}^{n}\tilde{x}_{it}^{'}\tilde{x}_{it})^{-1}(\sum_{t=1}^{T}\sum_{i=1}^{n}\tilde{x}_{it}^{'}\tilde{v}_{it}).$$

It is easy to show that

$$\tilde{\beta} - \beta = O_p(\frac{1}{\sqrt{nT}}).$$

Now $\hat{v}_{it} = \tilde{y}_{it} - \tilde{x}_{it}^{'}\tilde{\beta} = \tilde{x}_{it}^{'}\beta + \tilde{v}_{it} - \tilde{x}_{it}^{'}\tilde{\beta} = \tilde{v}_{it} - \tilde{x}_{it}^{'}(\beta - \tilde{\beta})$. In vector form $\hat{v}_t = \tilde{v}_t - \tilde{x}_t(\tilde{\beta} - \beta)$, where $\tilde{v}_t = v_t - \bar{v}_{\cdot\cdot}$.

Consider (a),

$$\begin{aligned}\frac{1}{n}\hat{M}_{1,T} =& \frac{1}{n} \times \frac{1}{T}\sum_{t=1}^{T}\hat{v}_t'\hat{v}_t = \frac{1}{n}\sum_{t=1}^{T}(\tilde{v}_t - \tilde{x}_t(\tilde{\beta} - \beta))'(\tilde{v}_t - \tilde{x}_t(\tilde{\beta} - \beta)) \\ =& \frac{1}{nT}\sum_{t=1}^{T}(\tilde{v}_t'\tilde{v}_t - \tilde{v}_t'\tilde{x}(\tilde{\beta} - \beta) - (\tilde{\beta} - \beta)'\tilde{x}_t'\tilde{v}_t + (\tilde{\beta} - \beta)'\tilde{x}_t'\tilde{x}_t(\tilde{\beta} - \beta)) \\ =& \sum_{k=1}^{3} A_k,\end{aligned}$$

where $A_1 = \frac{1}{nT}\sum_{t=1}^{T}\tilde{v}_t'\tilde{v}_t$; $A_2 = -\frac{2}{nT}\sum_{t=1}^{T}\tilde{v}_t'\tilde{x}_t(\tilde{\beta} - \beta)$; and $A_3 = \frac{1}{nT}\sum_{t=1}^{T}(\tilde{\beta} - \beta)'\tilde{x}_t'\tilde{x}_t(\tilde{\beta} - \beta)$.

Lemma 4 *Under Assumption 1, 2 and the null,*

(a_1) $A_1 = \frac{1}{n}M_{1,T} - \frac{\sigma_v^2}{T} + O_p(\frac{1}{T\sqrt{n}})$;

(a_2) $A_2 = O_p(\frac{1}{nT})$;

(a_3) $A_3 = O_p(\frac{1}{T^2})$.

Proof. 1) Proof of a_1

$$\begin{aligned}
A_1 =& \frac{1}{nT}\sum_{t=1}^{T}\tilde{v}_t'\tilde{v}_t = \frac{1}{nT}\sum_{t=1}^{T}(v_t-\bar{v}.)'(v_t-\bar{v}.)\\
=& \frac{1}{nT}\sum_{t=1}^{T}v_t'v_t - \frac{2}{nT}\sum_{t=1}^{T}v_t'\bar{v}. + \frac{1}{n}\bar{v}.'\bar{v}.\\
=& \frac{1}{n}M_{1,T} - \frac{2}{nT}\sum_{t=1}^{T}v_t'(\frac{1}{T}\sum_{s=1}^{T}v_s) + \frac{1}{n}(\frac{1}{T}\sum_{t=1}^{T}v_t)'(\frac{1}{T}\sum_{s=1}^{T}v_s)\\
=& \frac{1}{n}M_{1,T} - \frac{2}{nT^2}\sum_{t=1}^{T}\sum_{s=1}^{T}v_t'v_s + \frac{1}{nT^2}\sum_{t=1}^{T}\sum_{s=1}^{T}v_t'v_s\\
=& \frac{1}{n}M_{1,T} - \frac{1}{nT^2}(\sum_{t=1}^{T}\sum_{i=1}^{n}v_{it}^2 + \sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n}v_{it}v_{is}).
\end{aligned}$$

With part (a) of Lemma 3, we can easily get

$$A_1 = \frac{1}{n}M_{1,T} - \frac{\sigma_v^2}{T} + O_p(\frac{1}{T\sqrt{n}}).$$

2) Proof of a_2 Now we consider A_2.

$$\begin{aligned}
A_2 =& -\frac{2}{nT}\sum_{t=1}^{T}(v_t-\bar{v}.)'\tilde{x}_t(\tilde{\beta}-\beta)\\
=& -\frac{2}{nT}\sum_{t=1}^{T}v_t'\tilde{x}_t(\tilde{\beta}-\beta) - \frac{2}{nT}\sum_{t=1}^{T}\bar{v}.'\tilde{x}_t(\tilde{\beta}-\beta)\\
=& -\frac{2}{nT}\sum_{t=1}^{T}v_{it}\tilde{x}_{it}(\tilde{\beta}-\beta) - \frac{2}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}(\frac{2}{T}\sum_{s=1}^{T}v_{is})\tilde{x}_{it}(\tilde{\beta}-\beta)\\
=& -\frac{2}{nT}\sum_{t=1}^{T}v_{it}\tilde{x}_{it}(\tilde{\beta}-\beta) - \frac{2}{nT^2}(\sum_{t=1}^{T}\sum_{i=1}^{n}v_{it}\tilde{x}_{it}(\tilde{\beta}-\beta) + \sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n}v_{is}\tilde{x}_{it}(\tilde{\beta}-\beta))\\
=& O_p(\frac{1}{nT}) + O_p(\frac{1}{nT^2})(O_p(\sqrt{nT}) + O_p(T\sqrt{n}))O_p(\frac{1}{\sqrt{nT}})\\
=& O_p(\frac{1}{nT}) + O_p(\frac{1}{nT\sqrt{T}}) = O_p(\frac{1}{nT}).
\end{aligned}$$

Here we use the fact that $\tilde{\beta} - \beta = O_p(\frac{1}{\sqrt{nT}})$.

3) Proof of a_3

$$
\begin{aligned}
A_4 =& \frac{1}{nT}(\tilde{\beta} - \beta)' \sum_{t=1}^{T} \sum_{i=1}^{n} \tilde{x}_{it}^2 (\tilde{\beta} - \beta) \\
=& O_p(\frac{1}{nT}) O_p(\frac{1}{nT}) O_p(nT) = O_p(\frac{1}{nT}).
\end{aligned}
$$

■

Hence with Lemma 4, we have

$$
\frac{1}{n}\hat{M}_{1,T} = \frac{1}{n}M_{1,T} - \frac{\sigma_v^2}{T} + O_p(\frac{1}{T\sqrt{n}}).
$$

A.2 Proof of part (2)

$$
\begin{aligned}
\frac{1}{n}\hat{M}_{2,T} =& \frac{1}{n} \times \frac{1}{T(T-1)} \sum_{t \neq s}^{T} \sum_{s=1}^{T} \hat{v}_t' \hat{v}_s \\
=& \frac{1}{nT(T-1)} \sum_{t \neq s}^{T} \sum_{s=1}^{T} (\tilde{v}_t - \tilde{x}_t(\tilde{\beta} - \beta))'(\tilde{v}_s - \tilde{x}_s(\tilde{\beta} - \beta)) \\
=& \sum_{k=1}^{3} B_k,
\end{aligned}
$$

where $B_1 = \frac{1}{nT(T-1)} \sum_{t \neq s}^{T} \sum_{s=1}^{T} \tilde{v}_t' \tilde{v}_s$; $B_2 = -\frac{2}{nT(T-1)} \sum_{t \neq s}^{T} \sum_{s=1}^{T} (\tilde{\beta} - \beta)' \tilde{x}_t' \tilde{v}_s$ and $B_3 = \frac{1}{nT(T-1)} \sum_{t \neq s}^{T} \sum_{s=1}^{T} (\tilde{\beta} - \beta)' \tilde{x}_t' \tilde{x}_s (\tilde{\beta} - \beta)$.

Lemma 5 *Under Assumption 1, 2 and the null,*

(b_1) $B_1 = \frac{1}{n}M_{2,T} - \frac{\sigma_v^2}{T} + O_p(\frac{1}{T\sqrt{n}})$;

(b_2) $B_2 = O_p(\frac{1}{nT})$;

(b_3) $B_3 = O_p(\frac{1}{nT})$.

Proof. 1) Proof of b_1

$$\begin{aligned}
B_1 =& \frac{1}{nT(T-1)} \sum_{t \neq s}^{T} \sum_{s=1}^{T} (v_t - \bar{v}.)'(v_s - \bar{v}.) \\
=& \frac{1}{nT(T-1)} \sum_{t \neq s}^{T} \sum_{s=1}^{T} v_t' v_s - \frac{1}{nT} \sum_{t=1}^{T} v_t' \bar{v}. - \frac{1}{nT} \sum_{s=1}^{T} \bar{v}.' v_s + \frac{1}{n} \bar{v}.' \bar{v}. \\
=& \frac{1}{n} M_{2,T} - \frac{1}{n} \bar{v}.' \bar{v}. = \frac{1}{n} M_{2,T} - \frac{1}{nT^2} \left(\sum_{t=1}^{T} \sum_{i=1}^{n} v_{it}^2 + \sum_{t \neq s}^{T} \sum_{s=1}^{T} \sum_{i=1}^{n} v_{it} v_{is} \right) \\
=& \frac{1}{n} M_{2,T} - \frac{\sigma_v^2}{T} + O_p\left(\frac{1}{T\sqrt{n}}\right).
\end{aligned}$$

2) Proof of b_2

$$\begin{aligned}
B_2 =& -\frac{2}{nT(T-1)} \sum_{t \neq s}^{T} \sum_{s=1}^{T} (\tilde{\beta} - \beta)' \tilde{x}_t' (v_s - \bar{v}.) \\
=& -\frac{2}{nT(T-1)} \sum_{t \neq s}^{T} \sum_{s=1}^{T} (\tilde{\beta} - \beta)' \tilde{x}_t' v_s + \frac{2}{nT} \sum_{t=1}^{T} (\tilde{\beta} - \beta)' \tilde{x}_t' \bar{v}. \\
=& -2 \frac{1}{nT(T-1)} \sum_{t \neq s}^{T} \sum_{s=1}^{T} \sum_{i=1}^{n} (\tilde{\beta} - \beta)' \tilde{x}_{it}' v_{is} + \frac{2}{nT} \sum_{t=1}^{T} \sum_{i=1}^{n} (\tilde{\beta} - \beta)' \tilde{x}_{it}' \left(\frac{1}{T} \sum_{s=1}^{T} v_{is} \right) \\
=& -\frac{2}{nT^2(T-1)} \sum_{t \neq s}^{T} \sum_{s=1}^{T} \sum_{i=1}^{n} (\tilde{\beta} - \beta)' \tilde{x}_{it}' v_{is} + \frac{2}{nT^2} \sum_{t=1}^{T} \sum_{i=1}^{n} (\tilde{\beta} - \beta)' \tilde{x}_{it}' v_{it} \\
=& O_p\left(\frac{1}{nT^2\sqrt{T}}\right) + O_p\left(\frac{1}{nT^2}\right) = O_p\left(\frac{1}{nT^2}\right).
\end{aligned}$$

3) Proof of b_3

$$\begin{aligned}
B_3 =& \frac{1}{nT(T-1)} (\tilde{\beta} - \beta)' \sum_{t \neq s}^{T} \sum_{s=1}^{T} \sum_{i=1}^{n} \tilde{x}_{it} \tilde{x}_{is} (\tilde{\beta} - \beta) \\
=& O_p\left(\frac{1}{nT^2}\right) O_p\left(\frac{1}{nT}\right) O_p(nT^2) = O_p\left(\frac{1}{nT}\right).
\end{aligned}$$

■

Then Lemma 5 is proved and we could prove part (2) of Proposition 1, which is

$$\frac{1}{n}\hat{M}_{2,T} = \frac{1}{n}M_{2,T} - \frac{\sigma_v^2}{T} + O_p(\frac{1}{T\sqrt{n}}).$$

■

A.3 Proof of part (3)

With the results of part (1) and (2), we obtain

$$\begin{aligned}\frac{1}{n}\hat{R}_1 =& \frac{1}{n}\hat{M}_{1,T} - \frac{1}{n}\hat{M}_{2,T} \\ =& (\frac{1}{n}M_{1,T} - \frac{\sigma_v^2}{T} + O_p(\frac{1}{T\sqrt{n}})) - (\frac{1}{n}M_{2,T} - \frac{\sigma_v^2}{T} + O_p(\frac{1}{T\sqrt{n}})) \\ =& \frac{1}{n}M_{1,T} - \frac{1}{n}M_{2,T} + O_p(\frac{1}{T\sqrt{n}}) = \frac{1}{n}R_1 + O_p(\frac{1}{T\sqrt{n}}).\end{aligned}$$

B Proof of Proposition 2

Proof. Due to space limit, we put the proofs for part (1), (2), (3) in supplementary appendix. With part (1), (2), (3) of Proposition 2, we prove part (4) as the following

$$\begin{aligned}\frac{1}{n}\hat{R}_2 =& \frac{1}{n}\hat{M}_{3,T} - \frac{2}{n}\hat{M}_{4,T} + \frac{1}{n}\hat{M}_{5,T} \\ =& \frac{1}{n}M_{3,T} - \frac{2}{n}M_{4,T} + \frac{1}{n}M_{5,T} + O_p(\frac{\sqrt{n}}{T^2}) + O_p(\frac{1}{T\sqrt{T}}) + O_p(\frac{1}{T\sqrt{n}}) \\ =& \frac{1}{n}R_2 + O_p(\frac{\sqrt{n}}{T^2}) + O_p(\frac{1}{T\sqrt{T}}) + O_p(\frac{1}{T\sqrt{n}}).\end{aligned}$$

■

C Proof of Corollary 1

Proof. With part (3) of Proposition 1 and Lemma 2, as $(n, T) \to \infty$, we have

$$\begin{aligned}\frac{1}{n}\hat{R}_1 =& \frac{1}{n}\hat{M}_{1,T} - \frac{1}{n}\hat{M}_{2,T} = \frac{1}{n}R_1 + O_p(\frac{1}{T\sqrt{n}}) \\ =& \sigma_v^2 + O_p(\frac{1}{\sqrt{nT}}) \xrightarrow{p} \sigma_v^2.\end{aligned}$$

■

D Proof of Corollary 2

Proof. With part (4) of Proposition 2 and part (g) of Lemma 2, as $(n, T) \to \infty$ and $n/T^2 \to 0$, we have

$$\begin{aligned}\frac{1}{n}\hat{R}_2 =& \frac{1}{n}\hat{M}_{3,T} - \frac{2}{n}\hat{M}_{4,T} + \frac{1}{n}\hat{M}_{5,T} \\ =& \frac{1}{n}R_2 + O_p(\frac{\sqrt{n}}{T^2}) + O_p(\frac{1}{T\sqrt{n}}) + O_p(\frac{1}{T\sqrt{T}}) \xrightarrow{p} \sigma_v^4.\end{aligned}$$

■

E Proof of Proposition 3

Proof. With Proposition 1 and 2, we have

$$\begin{aligned}J_u - J_{czz} =& \frac{T(\hat{U}_n - U_n)}{2} = \frac{T}{2}\frac{(\frac{1}{n}\hat{R}_2)(\frac{1}{n}R_1)^2 - (\frac{1}{n}R_2)(\frac{1}{n}\hat{R}_1)^2}{(\frac{1}{n}\hat{R}_1)^2(\frac{1}{n}R_1)^2} \\ =& \frac{T[(\frac{1}{n}R_2 + O_p(\frac{\sqrt{n}}{T^2}) + O_p(\frac{1}{T\sqrt{n}}) + O_p(\frac{1}{T\sqrt{T}}))(\frac{1}{n}R_1)^2 - (\frac{1}{n}R_2)(\frac{1}{n}R_1 + O_p(\frac{1}{T\sqrt{n}}))^2]}{2(\frac{1}{n}R_1)^2(\frac{1}{n}R_1 + O_p(\frac{1}{T\sqrt{n}}))^2}.\end{aligned}$$

Let's first consider the numerator, since $\frac{1}{n}R_1 = O_p(1)$ and $\frac{1}{n}R_2 = O_p(1)$, as $(n,T) \to 0$ and $n/T^2 \xrightarrow{p} 0$, the numerator becomes:

$$
\begin{aligned}
&T[(\frac{1}{n}R_2 + O_p(\frac{\sqrt{n}}{T^2}) + O_p(\frac{1}{T\sqrt{n}}) + O_p(\frac{1}{T\sqrt{T}}))(\frac{1}{n}R_1)^2 - (\frac{1}{n}R_2)(\frac{1}{n}R_1 + O_p(\frac{1}{T\sqrt{n}}))^2] \\
=&T[(\frac{1}{n}R_2)(\frac{1}{n}R_1)^2 + (O_p(\frac{\sqrt{n}}{T^2}) + O_p(\frac{1}{T\sqrt{n}}) + O_p(\frac{1}{T\sqrt{T}}))(\frac{1}{n}R_1)^2 \\
&- (\frac{1}{n}R_2)(\frac{1}{n}R_1)^2 - 2O_p(\frac{1}{T\sqrt{n}})(\frac{1}{n}R_2) - (O_p(\frac{1}{T\sqrt{n}}))^2] \\
=&T[O_p(\frac{\sqrt{n}}{T^2}) + O_p(\frac{1}{T\sqrt{n}}) + O_p(\frac{1}{T\sqrt{T}}) + O_p(\frac{1}{nT^2})] \\
=&O_p(\frac{\sqrt{n}}{T}) + O_p(\frac{1}{\sqrt{n}}) + O_p(\frac{1}{\sqrt{T}})) \xrightarrow{p} 0.
\end{aligned}
$$

Next we consider the denominator, we expand it as the follows:

$$
\begin{aligned}
&2(\frac{1}{n}R_1)^2(\frac{1}{R_1} + O_p(\frac{1}{T\sqrt{n}}))^2 \\
=&2(\frac{1}{n}R_1)^4 + 4(\frac{1}{R_1})^3 O_p(\frac{1}{T\sqrt{n}}) + 2(\frac{1}{n}R_1)^2 O_p(\frac{1}{nT^2}) \\
=&O_p(1) + O_p(\frac{1}{T\sqrt{n}}) + O_p(\frac{1}{nT^2}) \\
=&O_p(1).
\end{aligned}
$$

From the results above, we can easily show that

$$\frac{T(\hat{U}_n - U_n)}{2} = o_p(1) \xrightarrow{p} 0.$$

■

F Proof of Theorem 1

Proof. Since $J_u = J_{czz} + \frac{T(\hat{U}_n - U_n)}{2}$, $J_{czz} \xrightarrow{d} N(0,1)$ and with Proposition 3, we obtain

$$J_u \xrightarrow{d} N(0,1).$$

■

The following Sections prove part (1), (2), (3) of Proposition 2.

G Proof of part (1)

Proof. Since $\frac{1}{n}\hat{M}_{3,T}$ is a scaler, we could expand it as the follows:

$$\begin{aligned}\frac{1}{n}\hat{M}_{3,T} &= \frac{1}{nT(T-1)}\sum_{t\neq s}^{T}\sum_{s=1}^{T}(\hat{v}_t'\hat{v}_s)^2 \\ &= \frac{1}{nT(T-1)}\sum_{t\neq s}^{T}\sum_{s=1}^{T}((\tilde{v}_t-\tilde{x}_t(\tilde{\beta}-\beta))'(\tilde{v}_s-\tilde{x}_s(\tilde{\beta}-\beta)))^2 \\ &= tr\{\frac{1}{nT(T-1)}\sum_{t\neq s}^{T}\sum_{s=1}^{T}(\tilde{v}_t'\tilde{v}_s-(\tilde{\beta}-\beta)'\tilde{x}_t'\tilde{v}_s-\tilde{v}_t'\tilde{x}_s(\tilde{\beta}-\beta)+(\tilde{\beta}-\beta)'\tilde{x}_t'\tilde{x}_s(\tilde{\beta}-\beta))^2\} \\ &= \sum_{i=1}^{7}C_i.\end{aligned}$$

where $C_1 = \frac{1}{nT(T-1)}\sum_{t\neq s}^{T}\sum_{s=1}^{T}(\tilde{v}_t'\tilde{v}_s\tilde{v}_t'\tilde{v}_s)$; $C_2 = -\frac{4}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}(\tilde{v}_t'\tilde{v}_s\tilde{x}_t'\tilde{v}_s)$; $C_3 = \frac{2}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\tilde{v}_t'\tilde{v}_s\tilde{x}_t'\tilde{x}_s(\tilde{\beta}-\beta)$; $C_4 = \frac{2}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}(\tilde{x}_t'\tilde{v}_s\tilde{v}_s'\tilde{x}_t)(\tilde{\beta}-\beta)$; $C_5 = \frac{2}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}(\tilde{x}_t'\tilde{v}_s\tilde{v}_t'\tilde{x}_s)(\tilde{\beta}-\beta)$; $C_6 = -\frac{4}{nT(T-1)}(\tilde{\beta}-\beta)'(\sum_{t\neq s}^{T}\sum_{s=1}^{T}\tilde{x}_t'\tilde{v}_s(\tilde{\beta}-\beta)\tilde{x}_t'\tilde{x}_s)(\tilde{\beta}-\beta)$; $C_7 = \frac{1}{nT(T-1)}(\tilde{\beta}-\beta)'(\sum_{t\neq s}^{T}\sum_{s=1}^{T}\tilde{x}_t'\tilde{x}_s(\tilde{\beta}-\beta)(\tilde{\beta}-\beta)'\tilde{x}_t'\tilde{x}_s)(\tilde{\beta}-\beta)$.

Lemma 6 *Under Assumption 1, 2, 3 and the null,*

(c_1) $C_1 = \frac{1}{n}M_{3,T} - \frac{4\sigma_v^4}{T} + 2(\frac{T-1}{T^2} + 2\frac{n-1}{T^2})\sigma_v^4 - \frac{3n}{T^2}\sigma_v^4 + O_p(\frac{\sqrt{n}}{T^2}) + O_p(\frac{1}{T\sqrt{T}}) + O_p(\frac{1}{T\sqrt{n}})$;

(c_2) $C_2 = O_p(\frac{1}{nT}) + O_p(\frac{1}{T^2}) + O_p(\frac{1}{T\sqrt{nT}})$;

(c_3) $C_3 = O_p(\frac{1}{T^2})$;

(c_4) $C_4 = O_p(\frac{1}{nT})$;

(c_5) $C_5 = O_p(\frac{1}{nT^2})$;

(c_6) $C_6 = O_p(\frac{1}{nT^2})$;

(c_7) $C_7 = O_p(\frac{1}{nT^2})$.

G.1 Proof of part (c_1)

$$
\begin{aligned}
C_1 =& \frac{1}{nT(T-1)} \sum_{t\neq s}^{T} \sum_{s=1}^{T} (\tilde{v}_t' \tilde{v}_s \tilde{v}_t' \tilde{v}_s) \\
=& \frac{1}{nT(T-1)} \sum_{t\neq s}^{T} \sum_{s=1}^{T} (v_t - \bar{v}.)'(v_s - \bar{v}.)(v_t - \bar{v}.)'(v_s - \bar{v}.) \\
=& \frac{1}{nT(T-1)} \sum_{t\neq s}^{T} \sum_{s=1}^{T} (v_t' v_s - \bar{v}.' v_s - v_t' \bar{v}. + \bar{v}.' \bar{v}.)^2 \\
=& \frac{1}{n} M_{3,T} + \sum_{k=1}^{5} C_1^k.
\end{aligned}
$$

Where $C_1^1 = -\frac{4}{nT(T-1)} \sum_{t\neq s}^{T} \sum_{s=1}^{T} v_t' v_s v_t' \bar{v}.$; $C_1^2 = \frac{2}{nT(T-1)} \sum_{t\neq s}^{T} \sum_{s=1}^{T} v_t' v_s \bar{v}.' \bar{v}.$; $C_1^3 = \frac{2}{nT} \sum_{t=1}^{T} v_t' \bar{v}. v_t' \bar{v}.$; $C_1^4 = \frac{2}{nT(T-1)} \sum_{t\neq s}^{T} \sum_{s=1}^{T} v_t' \bar{v}. \bar{v}.' v_s$ and $C_1^5 = -\frac{3}{n} \bar{v}.' \bar{v}. \bar{v}.' \bar{v}.$.

Lemma 7 *Under Assumption 1, 2, 3 and the null,*

(1) $C_1^1 = -\frac{4\sigma_v^4}{T} + O_p(\frac{1}{T\sqrt{n}}) + O_p(\frac{\sqrt{n}}{T^2})$;

(2) $C_1^2 = O_p(\frac{\sqrt{n}}{T^2})$;

(3) $C_1^3 = 2(\frac{T-1}{T^2} + \frac{n-1}{T^2})\sigma_v^4 + O_p(\frac{\sqrt{n}}{T^2}) + O_p(\frac{1}{T\sqrt{T}}) + O_p(\frac{1}{T\sqrt{n}})$;

(4) $C_1^4 = 2\frac{n-1}{T^2}\sigma_v^4 + O_p(\frac{\sqrt{n}}{T^2})$;

(5) $C_1^5 = -\frac{3n}{T^2}\sigma_v^4 + O_p(\frac{\sqrt{n}}{T^2})$.

Proof. 1) Proof of part (1)

$$
\begin{aligned}
C_1^1 &= -\frac{4}{nT(T-1)} \sum_{t\neq s}^{T} \sum_{s=1}^{T} v_t' v_s v_t' (\frac{1}{T} \sum_{\tau=1}^{T} v_\tau) \\
&= -\frac{4}{nT^2(T-1)} \sum_{t\neq s}^{T} \sum_{s=1}^{T} \sum_{\tau=1}^{T} v_t' v_s v_t' v_\tau \\
&= -\frac{4}{nT^2(T-1)} \sum_{t\neq s}^{T} \sum_{s=1}^{T} \sum_{\tau=1}^{T} \sum_{i=1}^{n} \sum_{j=1}^{n} v_{it} v_{is} v_{jt} v_{j\tau} \\
&= -\frac{4}{nT^2(T-1)} \sum_{t\neq s}^{T} \sum_{s=1}^{T} \sum_{\tau=1}^{T} \sum_{i=1}^{n} v_{it}^2 v_{is} v_{i\tau} - \frac{4}{nT^2(T-1)} \sum_{t\neq s}^{T} \sum_{s=1}^{T} \sum_{\tau=1}^{T} \sum_{i\neq j}^{n} \sum_{j=1}^{n} v_{it} v_{is} v_{jt} v_{j\tau}.
\end{aligned}
$$

Here we need to discuss three cases of (s, t, τ): (1) $t \neq s; t = \tau$; (2) $t \neq s; s = \tau$; (3) $t \neq s \neq \tau$. Then the first term becomes the following:

$$
\begin{aligned}
& -\frac{4}{nT^2(T-1)} \sum_{t\neq s}^{T} \sum_{s=1}^{T} \sum_{\tau=1}^{T} \sum_{i=1}^{n} v_{it}^2 v_{is} v_{i\tau} \\
= & -\frac{4}{nT^2(T-1)} \sum_{t\neq s}^{T} \sum_{s=1}^{T} \sum_{i=1}^{n} v_{it}^2 v_{is}^2 - \frac{4}{nT^2(T-1)} \sum_{t\neq s}^{T} \sum_{s=1}^{T} \sum_{i=1}^{n} v_{it}^3 v_{is} \\
& -\frac{4}{nT^2(T-1)} \sum_{t\neq s,\tau}^{T} \sum_{s\neq \tau}^{T} \sum_{\tau=1}^{T} \sum_{i=1}^{n} v_{it}^2 v_{is} v_{i\tau} \\
= & -\frac{4\sigma_v^4}{T} + O_p(\frac{1}{T\sqrt{n}}) + O_P(\frac{1}{T^2\sqrt{n}}) + O_p(\frac{1}{T\sqrt{n}}) \\
= & -\frac{4\sigma_v^4}{T} + O_p(\frac{1}{T\sqrt{n}}).
\end{aligned}
$$

Next we consider the second term

$$
\begin{aligned}
&-\frac{4}{nT^2(T-1)}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}v_{it}v_{is}v_{jt}v_{j\tau}\\
=&-\frac{4}{nT^2(T-1)}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}v_{it}v_{is}v_{jt}^2-\frac{4}{nT^2(T-1)}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}v_{it}v_{is}v_{jt}v_{js}\\
&-\frac{4}{nT^2(T-1)}\sum_{t\neq s,\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}v_{it}v_{is}v_{jt}v_{j\tau}\\
=&O_p(\frac{\sqrt{n}}{T^2})+O_p(\frac{1}{T^2})+O_p(\frac{1}{T\sqrt{T}})=O_p(\frac{\sqrt{n}}{T^2}).
\end{aligned}
$$

Therefore, we obtain that

$$
C_1^1=-\frac{4\sigma_v^4}{T}+O_p(\frac{1}{T\sqrt{n}})+O_p(\frac{\sqrt{n}}{T^2}).
$$

2) Proof of part (2)

$$
\begin{aligned}
C_1^2=&\frac{2}{nT(T-1)}\sum_{t\neq s}^{T}\sum_{s=1}^{T}v_t'v_s(\frac{1}{T}\sum_{\tau=1}^{T}v_\tau)'(\frac{1}{T}\sum_{\eta=1}^{T}v_\eta)\\
=&\frac{2}{nT^3(T-1)}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{\eta=1}^{T}v_t'v_sv_\tau'v_\eta\\
=&\frac{2}{nT^3(T-1)}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{\eta=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{n}v_{it}v_{is}v_{j\tau}v_{j\eta}\\
=&\frac{2}{nT^3(T-1)}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{\eta=1}^{T}\sum_{i=1}^{n}v_{it}v_{is}v_{i\tau}v_{i\eta}\\
&+\frac{2}{nT^3(T-1)}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{\eta=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}v_{it}v_{is}v_{j\tau}v_{j\eta}.
\end{aligned}
$$

To consider the order of the two terms of C_1^2. First, we distinct the cases of the first term, for $(t\neq s,\tau,\eta)$, we have ten cases:

I) Two " = "s: (1) $(s=\tau=\eta)\neq t$; (2) $(t=\tau=\eta)\neq s$; (3) $(s=\tau)\neq(t=\eta)$; (4) $(s=\eta)\neq(t=\tau)$;

II) One " = ": (5) $(t=\tau)\neq s\neq\eta$; (6) $(t=\eta)\neq s\neq\tau$; (7) $(s=\tau)\neq t\neq\eta$; (8) $(s=\eta)\neq t\neq\tau$; (9) $(\tau=\eta)\neq t\neq s$;

III) No " = ": (10) $t\neq s\neq\tau\neq\eta$.

Then the first term of C_1^2 can be calculated respectively as the following:

(1) $(s=\tau=\eta)\neq t: \frac{1}{nT^3(T-1)}\sum_{t\neq s}^T\sum_{s=1}^T\sum_{i=1}^n v_{it}v_{is}^3 = O_p(\frac{1}{T^3\sqrt{n}})$;

(2) $(t=\tau=\eta)\neq s: \frac{1}{nT^3(T-1)}\sum_{t\neq s}^T\sum_{s=1}^T\sum_{i=1}^n v_{it}^3v_{is} = O_p(\frac{1}{T^3\sqrt{n}})$;

(3) $(s=\tau)\neq(t=\eta): \frac{1}{nT^3(T-1)}\sum_{t\neq s}^T\sum_{s=1}^T\sum_{i=1}^n v_{it}^2v_{is}^2 = O_p(\frac{1}{T^2})$;

(4) $(s=\eta)\neq(t=\tau): \frac{1}{nT^3(T-1)}\sum_{t\neq s}^T\sum_{s=1}^T\sum_{i=1}^n v_{it}^2v_{is}^2 = O_p(\frac{1}{T^2})$;

(5) $(t=\tau)\neq s\neq\eta: \frac{1}{nT^3(T-1)}\sum_{t\neq s,\eta}^T\sum_{s\neq\eta}^T\sum_{\eta=1}^T\sum_{i=1}^n v_{it}^2v_{is}v_{i\eta} = O_p(\frac{1}{T^2\sqrt{n}})$;

(6) $(t=\eta)\neq s\neq\tau: \frac{1}{nT^3(T-1)}\sum_{t\neq s,\tau}^T\sum_{s\neq\tau}^T\sum_{\tau=1}^T\sum_{i=1}^n v_{it}^2v_{is}v_{i\tau} = O_p(\frac{1}{T^2\sqrt{n}})$;

(7) $(s=\tau)\neq t\neq\eta: \frac{1}{nT^3(T-1)}\sum_{t\neq s,\eta}^T\sum_{s\neq\eta}^T\sum_{\eta=1}^T\sum_{i=1}^n v_{it}v_{is}^2v_{i\eta} = O_p(\frac{1}{T^2\sqrt{n}})$;

(8) $(s=\eta)\neq s\neq\tau: \frac{1}{nT^3(T-1)}\sum_{t\neq s,\tau}^T\sum_{s\neq\tau}^T\sum_{\tau=1}^T\sum_{i=1}^n v_{it}v_{is}^2v_{i\tau} = O_p(\frac{1}{T^2\sqrt{n}})$;

(9) $(\tau=\eta)\neq t\neq s: \frac{1}{nT^3(T-1)}\sum_{t\neq s,\eta}^T\sum_{s\neq\eta}^T\sum_{\eta=1}^T\sum_{i=1}^n v_{it}v_{is}v_{i\eta}^2 = O_p(\frac{1}{T^2\sqrt{n}})$;

(10)$t\neq s\neq\tau\neq\eta: \frac{1}{nT^3(T-1)}\sum_{t\neq s,\tau,\eta}^T\sum_{s\neq\tau,\eta}^T\sum_{\tau\neq\eta}^T\sum_{\eta=1}^T\sum_{i=1}^n v_{it}v_{is}v_{i\tau}v_{i\eta} = O_p(\frac{1}{T^2\sqrt{n}})$.

Similarly, we can also expand the second term as the following ten cases:

(1) $(s=\tau=\eta)\neq t: \frac{1}{nT^3(T-1)}\sum_{t\neq s}^T\sum_{s=1}^T\sum_{i\neq j}^n\sum_{j=1}^n v_{it}v_{is}v_{js}^2 = O_p(\frac{\sqrt{n}}{T^3})$;

(2) $(t=\tau=\eta)\neq s: \frac{1}{nT^3(T-1)}\sum_{t\neq s}^T\sum_{s=1}^T\sum_{i\neq j}^n\sum_{j=1}^n v_{it}v_{is}v_{jt}^2 = O_p(\frac{\sqrt{n}}{T^3})$;

(3) $(s=\tau)\neq(t=\eta): \frac{1}{nT^3(T-1)}\sum_{t\neq s}^T\sum_{s=1}^T\sum_{i\neq j}^n\sum_{j=1}^n v_{it}v_{is}v_{jt}v_{js} = O_p(\frac{1}{T^3})$;

(4) $(s=\eta)\neq(t=\tau): \frac{1}{nT^3(T-1)}\sum_{t\neq s}^T\sum_{s=1}^T\sum_{i\neq j}^n\sum_{j=1}^n v_{it}v_{is}v_{js}v_{jt} = O_p(\frac{1}{T^3})$;

(5) $(t=\tau)\neq s\neq\eta: \frac{1}{nT^3(T-1)}\sum_{t\neq s,\eta}^T\sum_{s\neq\eta}^T\sum_{\eta=1}^T\sum_{i\neq j}^n\sum_{j=1}^n v_{it}v_{is}v_{jt}v_{j\eta} = O_p(\frac{1}{T^2\sqrt{T}})$;

(6) $(t=\eta)\neq s\neq\tau: \frac{1}{nT^3(T-1)}\sum_{t\neq s,\eta}^T\sum_{s\neq\eta}^T\sum_{\eta=1}^T\sum_{i\neq j}^n\sum_{j=1}^n v_{it}v_{is}v_{j\tau}v_{jt} = O_p(\frac{1}{T^2\sqrt{T}})$;

(7) $(s=\tau)\neq t\neq\eta: \frac{1}{nT^3(T-1)}\sum_{t\neq s,\eta}^T\sum_{s\neq\eta}^T\sum_{\eta=1}^T\sum_{i\neq j}^n\sum_{j=1}^n v_{it}v_{is}v_{js}v_{j\eta} = O_p(\frac{1}{T^2\sqrt{T}})$;

(8) $(s=\eta)\neq s\neq\tau: \frac{1}{nT^3(T-1)}\sum_{t\neq s,\eta}^T\sum_{s\neq\eta}^T\sum_{\eta=1}^T\sum_{i\neq j}^n\sum_{j=1}^n v_{it}v_{is}v_{j\tau}v_{js} = O_p(\frac{1}{T^2\sqrt{T}})$;

(9) $(\tau=\eta)\neq t\neq s: \frac{1}{nT^3(T-1)}\sum_{t\neq s,\eta}^T\sum_{s\neq\eta}^T\sum_{\eta=1}^T\sum_{i\neq j}^n\sum_{j=1}^n v_{it}v_{is}v_{j\tau}^2 = O_p(\frac{\sqrt{n}}{T^2})$;

(10)$t\neq s\neq\tau\neq\eta: \frac{1}{nT^3(T-1)}\sum_{t\neq s,\tau,\eta}^T\sum_{s\neq\tau,\eta}^T\sum_{\tau\neq\eta}^T\sum_{\eta=1}^T\sum_{i=1}^n v_{it}v_{is}v_{j\tau}v_{j\eta} = O_p(\frac{1}{T^2})$.

Hence

$$C_1^2 = O_p(\frac{\sqrt{n}}{T^2}).$$

3) Proof of part (3)

$$
\begin{aligned}
C_1^3 =& \frac{2}{nT}\sum_{t=1}^{T} v_t'(\frac{1}{T}\sum_{s=1}^{T} v_s)v_t'(\frac{1}{T}\sum_{\tau=1}^{T} v_\tau) \\
=& \frac{2}{nT^3}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T} v_t'v_s v_t' v_\tau \\
=& \frac{2}{nT^3}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{n} v_{it}v_{is}v_{jt}v_{j\tau} \\
=& \frac{2}{nT^3}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n} v_{it}^2 v_{is}v_{i\tau} + \frac{2}{nT^3}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it}v_{is}v_{jt}v_{j\tau}.
\end{aligned}
$$

There are five cases for the first term, they are $t = s = \tau$; $t = s \neq \tau$, $t = \tau \neq s$, $s = \tau \neq t$, $t \neq s \neq \tau$. Then the first term can be expressed by the sum of the following five cases:

(1) $t = s = \tau : \frac{2}{nT^3}\sum_{t=1}^{T}\sum_{i=1}^{n} v_{it}^4 = O_p(\frac{1}{T^2})$;

(2) $t = s \neq \tau : \frac{2}{nT^3}\sum_{t\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n} v_{it}^3 v_{i\tau} = O_p(\frac{1}{T^2\sqrt{n}})$;

(3) $t = \tau \neq s : \frac{2}{nT^3}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n} v_{it}^3 v_{is} = O_p(\frac{1}{T^2\sqrt{n}})$;

(4) $s = \tau \neq t : \frac{2}{nT^3}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n} v_{it}^2 v_{is}^2 = 2\frac{T-1}{T^2}\sigma_v^4 + O_p(\frac{1}{T^2\sqrt{n}})$;

(5) $t \neq s \neq \tau : \frac{2}{nT^3}\sum_{t\neq s,\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n} v_{it}^2 v_{is}v_{i\tau} = O_p(\frac{1}{T\sqrt{n}})$.

Similarly, we can also get the five cases of the second term:

(1) $t = s = \tau : \frac{2}{nT^3}\sum_{t=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it}^2 v_{jt}^2 = 2\frac{n-1}{T^2}\sigma_v^4 + O_p(\frac{1}{T^2\sqrt{T}})$;

(2) $t = s \neq \tau : \frac{2}{nT^3}\sum_{t\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it}^2 v_{jt}v_{j\tau} = O_p(\frac{\sqrt{n}}{T^2})$;

(3) $t = \tau \neq s : \frac{2}{nT^3}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it}v_{is}v_{jt}^2 = O_p(\frac{\sqrt{n}}{T^2})$;

(4) $s = \tau \neq t : \frac{2}{nT^3}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it}v_{is}v_{jt}v_{js} = O_p(\frac{1}{T^2})$;

(5) $t \neq s \neq \tau : \frac{2}{nT^3}\sum_{t\neq s,\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it}v_{is}v_{jt}v_{j\tau} = O_p(\frac{1}{T\sqrt{T}})$.

Add all the terms up, we can easily get

$$
C_1^3 = 2(\frac{T-1}{T^2} + \frac{n-1}{T^2})\sigma_v^4 + O_p(\frac{\sqrt{n}}{T^2}) + O_p(\frac{1}{T\sqrt{T}}) + O_p(\frac{1}{T\sqrt{n}}).
$$

4) Proof of part (4)

$$
\begin{aligned}
C_1^4 =& \frac{2}{nT(T-1)} \sum_{t\neq s}^{T} \sum_{s=1}^{T} v_t' \bar{v}_. \bar{v}_.' v_s \\
=& \frac{2}{nT(T-1)} \sum_{t\neq s}^{T} \sum_{s=1}^{T} v_t' (\frac{1}{T} \sum_{\tau=1}^{T} v_\tau)(\frac{1}{T} \sum_{\eta=1}^{T} v_\eta)' v_s \\
=& \frac{2}{nT^3(T-1)} \sum_{t\neq s}^{T} \sum_{s=1}^{T} \sum_{\tau=1}^{T} \sum_{\eta=1}^{T} v_t' v_\tau v_\eta' v_s \\
=& \frac{2}{nT^3(T-1)} \sum_{t\neq s}^{T} \sum_{s=1}^{T} \sum_{\tau=1}^{T} \sum_{\eta=1}^{T} \sum_{i=1}^{n} v_{it} v_{i\tau} v_{i\eta} v_{is} \\
&+ \frac{2}{nT^3(T-1)} \sum_{t\neq s}^{T} \sum_{s=1}^{T} \sum_{\tau=1}^{T} \sum_{\eta=1}^{T} \sum_{i\neq j}^{n} \sum_{j=1}^{n} v_{it} v_{i\tau} v_{j\eta} v_{js}.
\end{aligned}
$$

To distinguish the two terms, there are 10 cases to discuss:

I) Two "="s: (1) $t = \tau = \eta \neq s$; (2) $t \neq s = \tau = \eta$; (3) $t = \tau \neq s = \eta$; (4) $t = \eta \neq s = \tau$;

II) One "=": (5) $(t = \tau) \neq s \neq \eta$; (6) $(t = \eta) \neq s \neq \tau$; (7) $t \neq (s = \tau) \neq \eta$; (8) $t \neq (s = \eta) \neq \tau$; (9) $t \neq s \neq (\tau = \eta)$;

III) No "=": (10) $t \neq s \neq \tau \neq \eta$.

Then the first term of C_1^4 can be expanded to 10 cases as the following:

(1) $t = \tau = \eta \neq s : \frac{2}{nT^3(T-1)} \sum_{t\neq s}^{T} \sum_{s=1}^{T} \sum_{i=1}^{n} v_{it}^3 v_{is} = O_p(\frac{1}{T^3\sqrt{n}})$;

(2) $t \neq s = \tau = \eta : \frac{2}{nT^3(T-1)} \sum_{t\neq s}^{T} \sum_{s=1}^{T} \sum_{i=1}^{n} v_{it} v_{is}^3 = O_p(\frac{1}{T^3\sqrt{n}})$;

(3) $t = \tau \neq s = \eta : \frac{2}{nT^3(T-1)} \sum_{t\neq s}^{T} \sum_{s=1}^{T} \sum_{i=1}^{n} v_{it}^2 v_{is}^2 = O_p(\frac{1}{T^2})$;

(4) $t = \eta \neq s = \tau : \frac{2}{nT^3(T-1)} \sum_{t\neq s}^{T} \sum_{s=1}^{T} \sum_{i=1}^{n} v_{it}^2 v_{is}^2 = O_p(\frac{1}{T^2})$;

(5) $(t = \tau) \neq s \neq \eta : \frac{2}{nT^3(T-1)} \sum_{t\neq s\neq \eta}^{T} \sum_{s\neq \eta}^{T} \sum_{\eta=1}^{T} \sum_{i=1}^{n} v_{it}^2 v_{i\eta} v_{is} = O_p(\frac{1}{T^2\sqrt{n}})$;

(6) $(t = \eta) \neq s \neq \tau : \frac{2}{nT^3(T-1)} \sum_{t\neq s\neq \tau}^{T} \sum_{s\neq \tau}^{T} \sum_{\tau=1}^{T} \sum_{i=1}^{n} v_{it}^2 v_{i\tau} v_{is} = O_p(\frac{1}{T^2\sqrt{n}})$;

(7) $t \neq (s = \tau) \neq \eta : \frac{2}{nT^3(T-1)} \sum_{t\neq s\neq \eta}^{T} \sum_{s\neq \eta}^{T} \sum_{\eta=1}^{T} \sum_{i=1}^{n} v_{it} v_{i\eta} v_{is}^2 = O_p(\frac{1}{T^2\sqrt{n}})$;

(8) $t \neq (s = \eta) \neq \tau : \frac{2}{nT^3(T-1)} \sum_{t\neq s\neq \tau}^{T} \sum_{s\neq \tau}^{T} \sum_{\tau=1}^{T} \sum_{i=1}^{n} v_{it} v_{i\tau} v_{is}^2 = O_p(\frac{1}{T^2\sqrt{n}})$;

(9) $t \neq s \neq (\tau = \eta) : \frac{2}{nT^3(T-1)} \sum_{t\neq s\neq \tau}^{T} \sum_{s\neq \tau}^{T} \sum_{\tau=1}^{T} \sum_{i=1}^{n} v_{it} v_{i\tau}^2 v_{is} = O_p(\frac{1}{T^2\sqrt{n}})$;

(10) $t \neq s \neq \tau \neq \eta : \frac{2}{nT^3(T-1)} \sum_{t\neq s\neq \tau\neq \eta}^{T} \sum_{s\neq \tau\neq \eta}^{T} \sum_{\tau\neq \eta}^{T} \sum_{\eta=1}^{T} \sum_{i=1}^{n} v_{it} v_{i\tau} v_{i\eta} v_{is} = O_p(\frac{1}{T^2\sqrt{n}})$.

Similarly, we can also get 10 terms for the second term's expansion:

(1) $t = \tau = \eta \neq s : \frac{2}{nT^3(T-1)} \sum_{t\neq s}^T \sum_{s=1}^T \sum_{i\neq j}^n \sum_{j=1}^n v_{it}^2 v_{jt} v_{js} = O_p(\frac{\sqrt{n}}{T^3})$;

(2) $t \neq s = \tau = \eta : \frac{2}{nT^3(T-1)} \sum_{t\neq s}^T \sum_{s=1}^T \sum_{i\neq j}^n \sum_{j=1}^n v_{it} v_{is} v_{js}^2 = O_p(\frac{\sqrt{n}}{T^3})$;

(3) $t = \tau \neq s = \eta : \frac{2}{nT^3(T-1)} \sum_{t\neq s}^T \sum_{s=1}^T \sum_{i\neq j}^n \sum_{j=1}^n v_{it}^2 v_{js}^2 = 2\frac{n-1}{T^2}\sigma_v^4 + O_p(\frac{1}{T^3})$;

(4) $t = \eta \neq s = \tau : \frac{2}{nT^3(T-1)} \sum_{t\neq s}^T \sum_{s=1}^T \sum_{i\neq j}^n \sum_{j=1}^n v_{it} v_{is} v_{jt} v_{js} = O_p(\frac{1}{T^3})$;

(5) $(t = \tau) \neq s \neq \eta : \frac{2}{nT^3(T-1)} \sum_{t\neq s\neq \eta}^T \sum_{s\neq \eta}^T \sum_{\eta=1}^T \sum_{i\neq j}^n \sum_{j=1}^n v_{it}^2 v_{j\eta} v_{js} = O_p(\frac{\sqrt{n}}{T^2})$;

(6) $(t = \eta) \neq s \neq \tau : \frac{2}{nT^3(T-1)} \sum_{t\neq s\neq \tau}^T \sum_{s\neq \tau}^T \sum_{\tau=1}^T \sum_{i\neq j}^n \sum_{j=1}^n v_{it} v_{i\tau} v_{jt} v_{js} = O_p(\frac{1}{T^2\sqrt{T}})$;

(7) $t \neq (s = \tau) \neq \eta : \frac{2}{nT^3(T-1)} \sum_{t\neq s\neq \eta}^T \sum_{s\neq \eta}^T \sum_{\eta=1}^T \sum_{i\neq j}^n \sum_{j=1}^n v_{it} v_{is} v_{j\eta} v_{js} = O_p(\frac{1}{T^2\sqrt{T}})$;

(8) $t \neq (s = \eta) \neq \tau : \frac{2}{nT^3(T-1)} \sum_{t\neq s\neq \tau}^T \sum_{s\neq \tau}^T \sum_{\tau=1}^T \sum_{i\neq j}^n \sum_{j=1}^n v_{it} v_{i\tau} v_{js}^2 = O_p(\frac{\sqrt{n}}{T^2})$;

(9) $t \neq s \neq (\tau = \eta) : \frac{2}{nT^3(T-1)} \sum_{t\neq s\neq \tau}^T \sum_{s\neq \tau}^T \sum_{\tau=1}^T \sum_{i\neq j}^n \sum_{j=1}^n v_{it} v_{i\tau} v_{j\tau} v_{js} = O_p(\frac{1}{T^2\sqrt{T}})$;

(10) $t \neq s \neq \tau \neq \eta : \frac{2}{nT^3(T-1)} \sum_{t\neq s\neq \tau\neq \eta}^T \sum_{s\neq \tau\neq \eta}^T \sum_{\tau\neq \eta}^T \sum_{\eta=1}^T \sum_{i\neq j}^n \sum_{j=1}^n v_{it} v_{i\tau} v_{j\eta} v_{js} = O_p(\frac{1}{T^2})$.

Hence,

$$C_1^4 = 2\frac{n-1}{T^2}\sigma_v^4 + O_p(\frac{\sqrt{n}}{T^2}).$$

5) Proof of part (5)

$$\begin{aligned}
C_1^5 =& -\frac{3}{n}\bar{v}_.'\bar{v}_.\bar{v}_.'\bar{v}_. = -3n(\frac{1}{n}\bar{v}_.'\bar{v}_.)^2 \\
=& -3n(\frac{1}{nT^2}\sum_{t=1}^T\sum_{i=1}^n v_{it}^2 + \frac{1}{nT^2}\sum_{t\neq s}^T\sum_{s=1}^T\sum_{i=1}^n v_{it}v_{is})^2 \\
=& -3n((\frac{1}{nT^2}\sum_{t=1}^T\sum_{i=1}^n v_{it}^2)^2 + 2(\frac{1}{nT^2}\sum_{t=1}^T\sum_{i=1}^n v_{it}^2)(\frac{1}{nT^2}\sum_{t\neq s}^T\sum_{s=1}^T\sum_{i=1}^n v_{it}v_{is}) \\
&+ (\frac{1}{nT^2}\sum_{t\neq s}^T\sum_{s=1}^T\sum_{i=1}^n v_{it}v_{is})^2) \\
=& -3n((\frac{1}{nT^2}\sum_{t=1}^T\sum_{i=1}^n v_{it}^2)^2 + 2(\frac{1}{nT^2}\sum_{t=1}^T\sum_{i=1}^n v_{it}^2)(\frac{1}{nT^2}\sum_{t\neq s}^T\sum_{s=1}^T\sum_{i=1}^n v_{it}v_{is}) + o_p) \\
=& -\frac{3n}{T^2}\sigma_v^4 + O_p(\frac{\sqrt{n}}{T^2\sqrt{T}}) + O_p(\frac{\sqrt{n}}{T^2}) = -\frac{3n}{T^2}\sigma_v^4 + O_p(\frac{\sqrt{n}}{T^2}).
\end{aligned}$$

Hence, with the results above we have

$$C_1 = \frac{1}{n}M_{3,T} - \frac{4\sigma_v^4}{T} + 2(\frac{T-1}{T^2} + \frac{n-1}{T^2})\sigma_v^4 - \frac{3n}{T^2}\sigma_v^4 + O_p(\frac{\sqrt{n}}{T^2}) + O_p(\frac{1}{T\sqrt{T}}) + O_p(\frac{1}{T\sqrt{n}}).$$

■

G.2 Proof of part c_2

$$\begin{aligned} C_2 &= -\frac{4}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}(\tilde{v}_t'\tilde{v}_s\tilde{x}_t'\tilde{v}_s) \\ &= -\frac{4}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}((v_t-\bar{v}.)'(v_s-\bar{v}.)\tilde{x}_t'(v_s-\bar{v}.)) \\ &= \sum_{k=1}^{8} C_2^k. \end{aligned}$$

Where $C_2^1 = -\frac{4}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^T\sum_{s=1}^T v_t'v_s\tilde{x}_t'v_s$; $C_2^2 = \frac{4}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^T\sum_{s=1}^T v_t'v_s\tilde{x}_t'\bar{v}.$; $C_2^3 = \frac{4}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^T\sum_{s=1}^T v_t'\bar{v}.\tilde{x}_t'v_s$; $C_2^4 = -\frac{4}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^T\sum_{s=1}^T v_t'\bar{v}.\tilde{x}_t'\bar{v}.$; $C_2^5 = \frac{4}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^T\sum_{s=1}^T \bar{v}.v_s\tilde{x}_t'v_s$; $C_2^6 = -\frac{4}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^T\sum_{s=1}^T \bar{v}.v_s\tilde{x}_t'\bar{v}.$; $C_2^7 = -\frac{4}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^T\sum_{s=1}^T \bar{v}.\bar{v}.\tilde{x}_t'v_s$; $C_2^8 = \frac{4}{nT}(\tilde{\beta}-\beta)'\sum_{t=1}^T \bar{v}.\bar{v}.\tilde{x}_t'\bar{v}..$

Lemma 8 *Under Assumption 1, 2, 3 and the null,*

(1) $C_2^1 = O_p(\frac{1}{nT}) + O_p(\frac{1}{T\sqrt{nT}})$;

(2) $C_2^2 = O_p(\frac{1}{T^2\sqrt{n}})$;

(3) $C_2^3 = O_p(\frac{1}{T^2})$;

(4) $C_2^4 = O_p(\frac{1}{T^2})$;

(5) $C_2^5 = O_p(\frac{1}{T^2}) + O_p(\frac{1}{nT}) + O_p(\frac{1}{T\sqrt{nT}})$;

(6) $C_2^6 = O_p(\frac{1}{T^2})$;

(7) $C_2^7 = O_p(\frac{1}{T^2})$;

(8) $C_2^8 = O_p(\frac{1}{T^2})$.

Proof. Proof of part (1)

$$\begin{aligned}
C_2^1 =& -\frac{4}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{n} v_{it}v_{is}\tilde{x}_{jt}v_{js} \\
=& -\frac{4}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n} v_{it}v_{is}^2\tilde{x}_{it} - \frac{2}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it}v_{is}\tilde{x}_{jt}v_{js} \\
=& O_p(\frac{1}{nT}) + O_p(\frac{1}{T\sqrt{nT}}).
\end{aligned}$$

Proof of part (2)

$$\begin{aligned}
C_2^2 =& \frac{4}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T} v_t' v_s \tilde{x}_t' \bar{v}. \\
=& \frac{4}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{n} v_{it}v_{is}\tilde{x}_{jt}(\frac{1}{T}\sum_{\tau=1}^{T} v_{j\tau}) \\
=& \frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n} v_{it}v_{is}\tilde{x}_{it}v_{i\tau} \\
&+ \frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it}v_{is}\tilde{x}_{jt}v_{j\tau}.
\end{aligned}$$

To expand the two terms of C_2^2, we consider 3 cases : (1) $t = \tau \neq s$; (2) $t \neq \tau \neq s$; (3) $t \neq s = \tau$. The first term of can be expanded as the following terms:

(1) $t = \tau \neq s : \frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n} v_{it}^2 v_{is}\tilde{x}_{it} = O_p(\frac{1}{nT^2})$;

(2) $t \neq s \neq \tau : \frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s,\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n} v_{it}v_{is}\tilde{x}_{it}v_{i\tau} = O_p(\frac{1}{nT^2})$;

(3) $t \neq s = \tau : \frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s,\tau}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n} v_{it}v_{is}^2\tilde{x}_{it} = O_p(\frac{1}{nT^2})$.

The second term can be expanded as:

(1) $t = \tau \neq s : \frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it}v_{is}\tilde{x}_{jt}v_{jt} = O_p(\frac{1}{T^2\sqrt{nT}})$;

(2) $t \neq s \neq \tau : \frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s,\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it}v_{is}\tilde{x}_{jt}v_{j\tau} = O_p(\frac{1}{T^2\sqrt{n}})$;

(3) $t \neq s = \tau$: $\frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s,\tau}^{T} \sum_{s=1}^{T} \sum_{i\neq j}^{n} \sum_{j=1}^{n} v_{it}v_{is}\tilde{x}_{jt}v_{js} = O_p(\frac{1}{T^2\sqrt{nT}})$.

Then we have

$$C_2^2 = O_p(\frac{1}{T^2\sqrt{n}}).$$

Proof of part (3)

$$\begin{aligned}
C_2^3 =& \frac{4}{nT(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s}^{T} \sum_{s=1}^{T} v_t' \bar{v}.\tilde{x}_t' v_s \\
=& \frac{4}{nT(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s}^{T} \sum_{s=1}^{T} v_t' (\frac{1}{T} \sum_{\eta=1}^{T} v_\tau)\tilde{x}_t' v_s \\
=& \frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s}^{T} \sum_{s=1}^{T} \sum_{\tau=1}^{T} \sum_{i=1}^{n} v_{it}v_{i\tau}\tilde{x}_{it}' v_{is} \\
&+ \frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s}^{T} \sum_{s=1}^{T} \sum_{\tau=1}^{T} \sum_{i\neq j}^{n} \sum_{j=1}^{n} v_{it}v_{i\tau}\tilde{x}_{jt}' v_{js}.
\end{aligned}$$

As C_2^2, the two terms can be expanded as the following three cases, let's first consider the first term:

(1) $t = \tau \neq s$: $\frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s}^{T} \sum_{s=1}^{T} \sum_{i=1}^{n} v_{it}^2\tilde{x}_{it}v_{is} = O_p(\frac{1}{nT^2})$;

(2) $t \neq s = \tau$: $\frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s}^{T} \sum_{s=1}^{T} \sum_{i=1}^{n} v_{it}v_{is}^2\tilde{x}_{it} = O_p(\frac{1}{nT^2})$;

(3) $t \neq s \neq \tau$: $\frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s\neq \eta}^{T} \sum_{s\neq \eta}^{T} \sum_{\tau=1} \sum_{i=1}^{n} v_{it}v_{i\tau}\tilde{x}_{it}v_{is} = O_p(\frac{1}{nT^2})$.

The second term can be expanded as the following:

(1) $t = \tau \neq s$: $\frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s}^{T} \sum_{s=1}^{T} \sum_{i\neq j}^{n} \sum_{j=1}^{n} v_{it}^2\tilde{x}_{jt}v_{js} = O_p(\frac{1}{T^2})$;

(2) $t \neq s = \tau$: $\frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s}^{T} \sum_{s=1}^{T} \sum_{i\neq j}^{n} \sum_{j=1}^{n} v_{it}v_{is}\tilde{x}_{jt}v_{js} = O_p(\frac{1}{T^2\sqrt{nT}})$;

(3) $t \neq s \neq \tau$: $\frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s\neq \tau}^{T} \sum_{s\neq \tau}^{T} \sum_{\tau=1} \sum_{i\neq j}^{n} \sum_{j=1}^{n} v_{it}v_{i\tau}\tilde{x}_{jt}v_{js} = O_p(\frac{1}{T^2\sqrt{n}})$.

Then we have

$$C_2^3 = O_p(\frac{1}{T^2}).$$

Proof of part (4)

$$\begin{aligned}
C_2^4 =& -\frac{4}{nT}(\tilde{\beta}-\beta)'\sum_{t=1}^{T} v_t'\bar{v}.\tilde{x}_t'\bar{v}.\\
=& -\frac{4}{nT}(\tilde{\beta}-\beta)'\sum_{t=1}^{T} v_t'(\frac{1}{T}\sum_{s=1}^{T} v_s)\tilde{x}_t'(\frac{1}{T}\sum_{\tau=1}^{T} v_\tau)\\
=& -\frac{4}{nT^3}(\tilde{\beta}-\beta)'\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n} v_{it}v_{is}\tilde{x}_{it}'v_{i\tau}\\
& -\frac{4}{nT^3}(\tilde{\beta}-\beta)'\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it}v_{is}\tilde{x}_{jt}'v_{j\tau}.
\end{aligned}$$

We distinguish 5 cases for the expansion of the two terms, consider the first term:

(1) $t=s=\tau: \frac{4}{nT^3}(\tilde{\beta}-\beta)'\sum_{t=1}^{T}\sum_{i=1}^{n} v_{it}^3\tilde{x}_{it} = O_p(\frac{1}{nT^3})$;

(2) $t=s\neq\tau: \frac{4}{nT^3}(\tilde{\beta}-\beta)'\sum_{t\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n} v_{it}^2\tilde{x}_{it}v_{i\tau} = O_p(\frac{1}{nT^2})$;

(3) $t=\tau\neq s: \frac{4}{nT^3}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n} v_{it}^2 v_{is}\tilde{x}_{it} = O_p(\frac{1}{nT^2})$;

(4) $t\neq s=\tau: \frac{4}{nT^3}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n} v_{it}v_{is}^2\tilde{x}_{it} = O_p(\frac{1}{nT^2})$;

(5) $t\neq s\neq\tau: \frac{4}{nT^3}(\tilde{\beta}-\beta)'\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{i=1}^{n} v_{it}v_{is}\tilde{x}_{it}v_{i\eta} = O_p(\frac{1}{nT^2})$.

Similarly the second term also can be expanded as the following terms:

(1) $t=s=\tau: \frac{4}{nT^3}(\tilde{\beta}-\beta)'\sum_{t=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it}^2\tilde{x}_{jt}v_{jt} = O_p(\frac{1}{T^3})$;

(2) $t=s\neq\tau: \frac{4}{nT^3}(\tilde{\beta}-\beta)'\sum_{t\neq\tau}^{T}\sum_{\eta=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it}^2\tilde{x}_{jt}v_{j\tau} = O_p(\frac{1}{T^2})$;

(3) $t=\tau\neq s: \frac{4}{nT^3}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it}v_{is}\tilde{x}_{jt}v_{jt} = O_p(\frac{1}{T^2\sqrt{nT}})$;

(4) $t\neq s=\tau: \frac{4}{nT^3}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it}v_{is}\tilde{x}_{jt}v_{js} = O_p(\frac{1}{T^2\sqrt{nT}})$;

(5) $t\neq s\neq\tau: \frac{4}{nT^3}(\tilde{\beta}-\beta)'\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it}v_{is}\tilde{x}_{jt}v_{j\tau} = O_p(\frac{1}{T^2\sqrt{n}})$.

With the results above, we get

$$C_2^4 = O_p(\frac{1}{T^2}).$$

Proof of part (5)

$$\begin{aligned}
C_2^5 =& \frac{4}{nT(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s}^{T}\sum_{s=1}^{T} \bar{v}_{\cdot}' v_s \tilde{x}_t' v_s \\
=& \frac{4}{nT(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s}^{T}\sum_{s=1}^{T} (\frac{1}{T}\sum_{\tau=1}^{n} v_\tau)' v_s \tilde{x}_t' v_s \\
=& \frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n} v_{i\tau} v_{is}^2 \tilde{x}_{it} \\
&+ \frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{i=1}^{n} v_{i\tau} v_{is} \tilde{x}_{jt} v_{js}.
\end{aligned}$$

The first term then can be expressed as the following 3 terms:

(1) $t = \tau \neq s : \frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n} v_{it} v_{is}^2 \tilde{x}_{it} = O_p(\frac{1}{nT^2})$;

(2) $t \neq s = \tau : \frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n} v_{is}^3 \tilde{x}_{it} = O_p(\frac{1}{nT^2})$;

(3) $t \neq s \neq \tau : \frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s\neq \tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n} v_{i\tau} v_{is}^2 \tilde{x}_{it} = O_p(\frac{1}{nT})$.

The second term of C_2^5 can be expressed as:

(1) $t = \tau \neq s : \frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i\neq j}^{n}\sum_{i=1}^{n} v_{it} v_{is} \tilde{x}_{jt} v_{js} = O_p(\frac{1}{T^2\sqrt{nT}})$;

(2) $t \neq s = \tau : \frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i\neq j}^{n}\sum_{i=1}^{n} v_{is}^2 \tilde{x}_{jt} v_{js} = O_p(\frac{1}{T^2})$;

(3) $t \neq s \neq \tau : \frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s\neq \tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{i=1}^{n} v_{i\tau} v_{is} \tilde{x}_{jt} v_{js} = O_p(\frac{1}{T\sqrt{nT}})$.

Then, we have

$$C_2^5 = O_p(\frac{1}{T^2}) + O_p(\frac{1}{nT}) + O_p(\frac{1}{T\sqrt{nT}}).$$

Proof of part (6)

$$
\begin{aligned}
C_2^6 =& -\frac{4}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\bar{v}_{\cdot}'v_s\tilde{x}_t'\bar{v}_{\cdot} \\
=& -\frac{4}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}(\frac{1}{T}\sum_{\tau=1}^{T}v_\tau)'v_s\tilde{x}_t'(\frac{1}{T}\sum_{\eta=1}^{T}v_\eta) \\
=& -\frac{4}{nT^3(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{\eta=1}^{T}\sum_{i=1}^{n}v_{i\tau}v_{is}\tilde{x}_{it}v_{i\eta} \\
&-\frac{4}{nT^3(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{\eta=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}v_{i\tau}v_{is}\tilde{x}_{jt}v_{j\eta}.
\end{aligned}
$$

There are 10 cases for expanding:

I) Two "="s: (1) $t=\tau=\eta\neq s$; (2) $t\neq s=\tau=\eta$; (3) $t=\tau\neq s=\eta$; (4) $t=\eta\neq s=\tau$;

II) One "=": (5) $t=\tau\neq s\neq\eta$; (6) $t=\eta\neq s\neq\tau$;(7) $t\neq\tau\neq s=\eta$; (8) $t\neq\eta\neq s=\tau$; (9) $t\neq s\neq\tau=\eta$;

III) No "=": (10) $t\neq s\neq\tau\neq\eta$.

The first term of C_2^6 can be expanded as the following 10 terms:

(1) $t=\tau=\eta\neq s: -\frac{4}{nT^3(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n}v_{it}^2v_{is}\tilde{x}_{it}=O_p(\frac{1}{nT^3})$;

(2) $t\neq s=\tau=\eta: -\frac{4}{nT^3(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n}v_{is}^3\tilde{x}_{it}=O_p(\frac{1}{nT^3})$;

(3) $t=\tau\neq s=\eta: -\frac{4}{nT^3(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n}v_{it}v_{is}^2\tilde{x}_{it}=O_p(\frac{1}{nT^3})$;

(4) $t=\eta\neq s=\tau: -\frac{4}{nT^3(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n}v_{is}^2\tilde{x}_{it}v_{it}=O_p(\frac{1}{nT^3})$;

(5) $t=\tau\neq s\neq\eta: -\frac{4}{nT^3(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s\neq\eta}^{T}\sum_{s\neq\eta}^{T}\sum_{\eta=1}^{T}\sum_{i=1}^{n}v_{it}v_{is}\tilde{x}_{it}v_{i\eta}=O_p(\frac{1}{nT^3})$;

(6) $t=\eta\neq s\neq\tau: -\frac{4}{nT^3(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n}v_{i\tau}v_{is}\tilde{x}_{it}v_{it}=O_p(\frac{1}{nT^3})$;

(7) $t\neq\tau\neq s=\eta: -\frac{4}{nT^3(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n}v_{i\tau}v_{is}^2\tilde{x}_{it}=O_p(\frac{1}{nT^2})$;

(8) $t\neq\eta\neq s=\tau: -\frac{4}{nT^3(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s\neq\eta}^{T}\sum_{s\neq\eta}^{T}\sum_{\eta=1}^{T}\sum_{i=1}^{n}v_{is}^2\tilde{x}_{it}v_{i\eta}=O_p(\frac{1}{nT^2})$;

(9) $t\neq s\neq\tau=\eta: -\frac{4}{nT^3(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n}v_{i\tau}^2v_{is}\tilde{x}_{it}=O_p(\frac{1}{nT^2})$;

(10) $t\neq s\neq\tau\neq\eta: -\frac{4}{nT^3(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s\neq\tau\neq\eta}^{T}\sum_{s\neq\tau\neq\eta}^{T}\sum_{\tau\neq\eta}^{T}\sum_{\eta=1}^{T}\sum_{i=1}^{n}v_{i\tau}v_{is}\tilde{x}_{it}v_{i\eta}=O_p(\frac{1}{nT^2})$.

Similarly, there are also 10 terms for expanding the second term of C_2^6:

(1) $t=\tau=\eta\neq s: -\frac{4}{nT^3(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^T\sum_{s=1}^T\sum_{i\neq j}^n\sum_{j=1}^n v_{it}v_{is}\tilde{x}_{jt}v_{jt}=O_p(\frac{1}{T^3\sqrt{nT}})$;

(2) $t\neq s=\tau=\eta: -\frac{4}{nT^3(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^T\sum_{s=1}^T\sum_{i\neq j}^n\sum_{j=1}^n v_{is}^2\tilde{x}_{jt}v_{js}=O_p(\frac{1}{T^3})$;

(3) $t=\tau\neq s=\eta: -\frac{4}{nT^3(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^T\sum_{s=1}^T\sum_{i\neq j}^n\sum_{j=1}^n v_{it}v_{is}\tilde{x}_{jt}v_{js}=O_p(\frac{1}{T^3\sqrt{nT}})$;

(4) $t=\eta\neq s=\tau: -\frac{4}{nT^3(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^T\sum_{s=1}^T\sum_{i\neq j}^n\sum_{j=1}^n v_{is}^2\tilde{x}_{jt}v_{jt}=O_p(\frac{1}{T^3})$;

(5) $t=\tau\neq s\neq\eta: -\frac{4}{nT^3(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s\neq\eta}^T\sum_{s\neq\eta}^T\sum_{\eta=1}^T\sum_{i\neq j}^n\sum_{j=1}^n v_{it}v_{is}\tilde{x}_{jt}v_{j\eta}=O_p(\frac{1}{T^3\sqrt{n}})$;

(6) $t=\eta\neq s\neq\tau: -\frac{4}{nT^3(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s\neq\tau}^T\sum_{s\neq\tau}^T\sum_{\tau=1}^T\sum_{i\neq j}^n\sum_{j=1}^n v_{i\tau}v_{is}\tilde{x}_{jt}v_{jt}=O_p(\frac{1}{T^3\sqrt{n}})$;

(7) $t\neq\tau\neq s=\eta: -\frac{4}{nT^3(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s\neq\tau}^T\sum_{s\neq\tau}^T\sum_{\tau=1}^T\sum_{i\neq j}^n\sum_{j=1}^n v_{i\tau}v_{is}\tilde{x}_{jt}v_{js}=O_p(\frac{1}{T^2\sqrt{nT}})$;

(8) $t\neq\eta\neq s=\tau: -\frac{4}{nT^3(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s\neq\eta}^T\sum_{s\neq\eta}^T\sum_{\eta=1}^T\sum_{i\neq j}^n\sum_{j=1}^n v_{is}^2\tilde{x}_{jt}v_{j\eta}=O_p(\frac{1}{T^2})$;

(9) $t\neq s\neq\tau=\eta: -\frac{4}{nT^3(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s\neq\tau}^T\sum_{s\neq\tau}^T\sum_{\tau=1}^T\sum_{i\neq j}^n\sum_{j=1}^n v_{i\tau}v_{is}\tilde{x}_{jt}v_{j\tau}=O_p(\frac{1}{T^2\sqrt{nT}})$;

(10) $t\neq s\neq\tau\neq\eta: -\frac{4}{nT^3(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s\neq\tau\neq\eta}^T\sum_{s\neq\tau\neq\eta}^T\sum_{\tau\neq\eta}^T\sum_{\eta=1}^T\sum_{i\neq j}^n\sum_{j=1}^n v_{i\tau}v_{is}\tilde{x}_{jt}v_{j\eta}=$ $O_p(\frac{1}{T^2\sqrt{n}})$.

Hence,

$$C_2^6=O_p(\frac{1}{T^2}).$$

Proof of part (7)

$$\begin{aligned}
C_2^7=&-\frac{4}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^T\sum_{s=1}^T\bar{v}_{\cdot}'\bar{v}_{\cdot}\tilde{x}_t'v_s\\
=&-\frac{4}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^T\sum_{s=1}^T(\frac{1}{T}\sum_{\tau=1}^T v_\tau)'(\frac{1}{T}\sum_{\eta=1}^T v_\eta)\tilde{x}_t'v_s\\
=&-\frac{4}{nT^3(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^T\sum_{s=1}^T\sum_{\tau=1}^T\sum_{\eta=1}^T\sum_{i=1}^n v_{i\tau}v_{i\eta}\tilde{x}_{it}v_{is}\\
&-\frac{4}{nT^3(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^T\sum_{s=1}^T\sum_{\tau=1}^T\sum_{\eta=1}^T\sum_{i\neq j}^n\sum_{j=1}^n v_{i\tau}v_{i\eta}\tilde{x}_{jt}v_{js}.
\end{aligned}$$

We consider the 10 cases to distinguish C_2^7

I) Two "="s: (1) $t=\tau=\eta\neq s$; (2) $t\neq s=\tau=\eta$; (3) $t=\tau\neq s=\eta$; (4) $t=\eta\neq s=\tau$;

2) One "=": (5) $t=\tau\neq s\neq\eta$; (6) $t=\eta\neq s\neq\tau$; (7) $t\neq\tau\neq s=\eta$; (8) $t\neq\eta\neq s=\tau$; (9) $t\neq s\neq\tau=\eta$;

3) No "=": (10) $t \neq s \neq \tau \neq \eta$.

Obviously the first term is equal to the first term of C_6^2, then

$$-\frac{4}{nT^3(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{\eta=1}^{T}\sum_{i=1}^{n} v_{i\tau}v_{i\eta}\tilde{x}_{it}v_{is} = O_p(\frac{1}{nT^2}).$$

Then consider the second term:

(1) $t = \tau = \eta \neq s : -\frac{4}{nT^3(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it}^2\tilde{x}_{jt}v_{js} = O_p(\frac{1}{T^3})$;

(2) $t \neq s = \tau = \eta : -\frac{4}{nT^3(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{is}^2\tilde{x}_{jt}v_{js} = O_p(\frac{1}{T^3})$;

(3) $t = \tau \neq s = \eta : -\frac{4}{nT^3(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it}v_{is}\tilde{x}_{jt}v_{js} = O_p(\frac{1}{T^3\sqrt{nT}})$;

(4) $t = \eta \neq s = \tau : -\frac{4}{nT^3(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{is}v_{it}\tilde{x}_{jt}v_{js} = O_p(\frac{1}{T^3\sqrt{nT}})$;

(5) $t = \tau \neq s \neq \eta : -\frac{4}{nT^3(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s\neq \eta}^{T}\sum_{s\neq \eta}^{T}\sum_{\eta=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it}v_{is}\tilde{x}_{jt}v_{js} = O_p(\frac{1}{T^3\sqrt{nT}})$;

(6) $t = \eta \neq s \neq \tau : -\frac{4}{nT^3(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s\neq \tau}^{T}\sum_{s\neq \tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{i\tau}v_{it}\tilde{x}_{jt}v_{js} = O_p(\frac{1}{T^3\sqrt{n}})$;

(7) $t \neq \tau \neq s = \eta : -\frac{4}{nT^3(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s\neq \tau}^{T}\sum_{s\neq \tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{i\tau}v_{is}\tilde{x}_{jt}v_{js} = O_p(\frac{1}{T^2\sqrt{nT}})$;

(8) $t \neq \eta \neq s = \tau : -\frac{4}{nT^3(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s\neq \eta}^{T}\sum_{s\neq \eta}^{T}\sum_{\eta=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{is}v_{i\eta}\tilde{x}_{jt}v_{js} = O_p(\frac{1}{T^2\sqrt{nT}})$;

(9) $t \neq s \neq \tau = \eta : -\frac{4}{nT^3(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s\neq \tau}^{T}\sum_{s\neq \tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{i\tau}^2\tilde{x}_{jt}v_{js} = O_p(\frac{1}{T^2})$;

(10) $t \neq s \neq \tau \neq \eta : -\frac{4}{nT^3(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s\neq \tau\neq \eta}^{T}\sum_{s\neq \tau\neq \eta}^{T}\sum_{\tau\neq \eta}^{T}\sum_{\eta=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{i\tau}v_{i\eta}\tilde{x}_{jt}v_{js} = O_p(\frac{1}{T^2\sqrt{n}})$.

With the results above, we obtain

$$C_2^7 = O_p(\frac{1}{T^2}).$$

Proof of part (8)

$$
\begin{aligned}
C_2^8 =& \frac{4}{nT}(\tilde{\beta}-\beta)' \sum_{t=1}^{T} \bar{v}_.' \bar{v}_. \tilde{x}_t' \bar{v}_. \\
=& \frac{4}{nT}(\tilde{\beta}-\beta)' \sum_{t=1}^{T} (\frac{1}{T}\sum_{s=1}^{T} v_s)'(\frac{1}{T}\sum_{\tau=1}^{T} v_\tau)\tilde{x}_t'(\frac{1}{T}\sum_{\eta=1}^{T} v_\eta) \\
=& \frac{4}{nT^4}(\tilde{\beta}-\beta)' \sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{\eta=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{n} v_{is}v_{i\tau}\tilde{x}_{jt}v_{j\eta} \\
=& \frac{4}{nT^4}(\tilde{\beta}-\beta)' \sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{\eta=1}^{T}\sum_{i=1}^{n} v_{is}v_{i\tau}\tilde{x}_{it}v_{i\eta} \\
&+ \frac{4}{nT^4}(\tilde{\beta}-\beta)' \sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{\eta=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{is}v_{i\tau}\tilde{x}_{jt}v_{j\eta}.
\end{aligned}
$$

We consider the following 15 cases for expansion:

I) Three "="s: (1) $t = s = \tau = \eta$;

II) Two "="s: (2) $(t = s) \neq (\tau = \eta)$; (3) $(t = \tau) \neq (s = \eta)$; (4) $(t = \eta) \neq (s = \tau)$; (5) $(t = s = \tau) \neq \eta$; (6) $(t = s = \eta) \neq \tau$; (7) $(t = \tau = \eta) \neq s$; (8) $t \neq (s = \tau = \eta)$;

III) One "=": (9) $t = s \neq \tau \neq \eta$; (10) $t = \tau \neq s \neq \eta$; (11) $t = \eta \neq s \neq \tau$; (12) $s = \tau \neq t \neq \eta$; (13) $s = \eta \neq t \neq \tau$; (14) $\tau = \eta \neq t \neq s$;

IV) No "=":(15) $t \neq s \neq \tau \neq \eta$.

For each case, the first term can be calculated respectively as above:

(1) $t = s = \tau = \eta : \frac{4}{nT^4}(\tilde{\beta}-\beta)' \sum_{t=1}^{T}\sum_{i=1}^{n} v_{it}^3 \tilde{x}_{it} = O_p(\frac{1}{nT^4})$;

(2) $(t = s) \neq (\tau = \eta) : \frac{4}{nT^4}(\tilde{\beta}-\beta)' \sum_{t\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n} v_{it}v_{i\tau}^2 \tilde{x}_{it} = O_p(\frac{1}{nT^3})$;

(3) $(t = \tau) \neq (s = \eta) : \frac{4}{nT^4}(\tilde{\beta}-\beta)' \sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n} v_{is}^2 v_{it} \tilde{x}_{it} = O_p(\frac{1}{nT^3})$;

(4) $(t = \eta) \neq (s = \tau) : \frac{4}{nT^4}(\tilde{\beta}-\beta)' \sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n} v_{is}^2 v_{it} \tilde{x}_{it} = O_p(\frac{1}{nT^3})$;

(5) $(t = s = \tau) \neq \eta : \frac{4}{nT^4}(\tilde{\beta}-\beta)' \sum_{t\neq\eta}^{T}\sum_{\eta=1}^{T}\sum_{i=1}^{n} v_{it}^2 v_{i\eta} \tilde{x}_{it} = O_p(\frac{1}{nT^3})$;

(6) $(t = s = \eta) \neq \tau : \frac{4}{nT^4}(\tilde{\beta}-\beta)' \sum_{t\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n} v_{it}^2 v_{i\tau} \tilde{x}_{it} = O_p(\frac{1}{nT^3})$;

(7) $(t = \tau = \eta) \neq s : \frac{4}{nT^4}(\tilde{\beta}-\beta)' \sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n} v_{it}^2 v_{is} \tilde{x}_{it} = O_p(\frac{1}{nT^3})$;

(8) $t \neq (s = \tau = \eta) : \frac{4}{nT^4}(\tilde{\beta}-\beta)' \sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n} v_{is}^3 \tilde{x}_{it} = O_p(\frac{1}{nT^3})$;

(9) $t = s \neq \tau \neq \eta : \frac{4}{nT^4}(\tilde{\beta} - \beta)' \sum_{t\neq\tau\neq\eta}^T \sum_{\tau\neq\eta}^T \sum_{\tau\neq\eta}^T \sum_{i=1}^n v_{it} v_{i\tau} \tilde{x}_{it} v_{i\eta} = O_p(\frac{1}{nT^3})$;

(10) $t = \tau \neq s \neq \eta : \frac{4}{nT^4}(\tilde{\beta} - \beta)' \sum_{t\neq s\neq\eta}^T \sum_{s\neq\eta}^T \sum_{\eta=1}^T \sum_{i=1}^n v_{is} v_{it} \tilde{x}_{it} v_{i\eta} = O_p(\frac{1}{nT^3})$;

(11) $t = \eta \neq s \neq \tau : \frac{4}{nT^4}(\tilde{\beta} - \beta)' \sum_{t\neq s\neq\tau}^T \sum_{s\neq\tau}^T \sum_{\tau=1}^T \sum_{i=1}^n v_{is} v_{i\tau} \tilde{x}_{it} v_{it} = O_p(\frac{1}{nT^3})$;

(12) $s = \tau \neq t \neq \eta : \frac{4}{nT^4}(\tilde{\beta} - \beta)' \sum_{t\neq s\neq\eta}^T \sum_{s\neq\eta}^T \sum_{\eta=1}^T \sum_{i=1}^n v_{is}^2 \tilde{x}_{it} v_{i\eta} = O_p(\frac{1}{nT^2})$;

(13) $s = \eta \neq t \neq \tau : \frac{4}{nT^4}(\tilde{\beta} - \beta)' \sum_{t\neq s\neq\tau}^T \sum_{s\neq\tau}^T \sum_{\tau=1}^T \sum_{i=1}^n v_{is}^2 v_{i\tau} \tilde{x}_{it} = O_p(\frac{1}{nT^2})$;

(14) $\tau = \eta \neq t \neq s : \frac{4}{nT^4}(\tilde{\beta} - \beta)' \sum_{t\neq s\neq\tau}^T \sum_{s\neq\tau}^T \sum_{\tau=1}^T \sum_{i=1}^n v_{is} v_{i\tau}^2 \tilde{x}_{it} = O_p(\frac{1}{nT^2})$;

(15) $t \neq s \neq \tau \neq \eta : \frac{4}{nT^4}(\tilde{\beta} - \beta)' \sum_{t\neq s\neq\tau\neq\eta}^T \sum_{s\neq\tau\neq\eta}^T \sum_{\tau\neq\eta}^T \sum_{\eta=1}^T \sum_{i=1}^n v_{is} v_{i\tau} \tilde{x}_{it} v_{i\eta} = O_p(\frac{1}{nT^2})$.

Similarly, the second term also can be expanded as the following 15 cases:

(1) $t = s = \tau = \eta : \frac{4}{nT^4}(\tilde{\beta} - \beta)' \sum_{t=1}^T \sum_{i\neq j}^n \sum_{j=1}^n v_{it}^2 \tilde{x}_{jt} v_{jt} = O_p(\frac{1}{T^4})$;

(2) $(t = s) \neq (\tau = \eta) : \frac{4}{nT^4}(\tilde{\beta} - \beta)' \sum_{t\neq\tau}^T \sum_{\tau=1}^T \sum_{i\neq j}^n \sum_{j=1}^n v_{it} v_{i\tau} \tilde{x}_{jt} v_{j\tau} = O_p(\frac{1}{T^3\sqrt{nT}})$;

(3) $(t = \tau) \neq (s = \eta) : \frac{4}{nT^4}(\tilde{\beta} - \beta)' \sum_{t\neq s}^T \sum_{s=1}^T \sum_{i\neq j}^n \sum_{j=1}^n v_{is} v_{it} \tilde{x}_{jt} v_{js} = O_p(\frac{1}{T^3\sqrt{nT}})$;

(4) $(t = \eta) \neq (s = \tau) : \frac{4}{nT^4}(\tilde{\beta} - \beta)' \sum_{t\neq s}^T \sum_{s=1}^T \sum_{i\neq j}^n \sum_{j=1}^n v_{is}^2 \tilde{x}_{it} v_{jt} = O_p(\frac{1}{T^3})$;

(5) $(t = s = \tau) \neq \eta : \frac{4}{nT^4}(\tilde{\beta} - \beta)' \sum_{t\neq\eta}^T \sum_{\eta=1}^T \sum_{i\neq j}^n \sum_{j=1}^n v_{it}^2 \tilde{x}_{jt} v_{j\eta} = O_p(\frac{1}{T^3})$;

(6) $(t = s = \eta) \neq \tau : \frac{4}{nT^4}(\tilde{\beta} - \beta)' \sum_{t\neq\tau}^T \sum_{\tau=1}^T \sum_{i\neq j}^n \sum_{j=1}^n v_{it} v_{i\tau} \tilde{x}_{jt} v_{jt} = O_p(\frac{1}{T^3\sqrt{nT}})$;

(7) $(t = \tau = \eta) \neq s : \frac{4}{nT^4}(\tilde{\beta} - \beta)' \sum_{t\neq s}^T \sum_{s=1}^T \sum_{i\neq j} \sum_{j=1}^n v_{is} v_{it} \tilde{x}_{jt} v_{jt} = O_p(\frac{1}{T^3\sqrt{nT}})$;

(8) $t \neq (s = \tau = \eta) : \frac{4}{nT^4}(\tilde{\beta} - \beta)' \sum_{t\neq s}^T \sum_{s=1}^T \sum_{i\neq j}^n \sum_{j=1}^n v_{is}^2 \tilde{x}_{jt} v_{js} = O_p(\frac{1}{T^3})$;

(9) $t = s \neq \tau \neq \eta : \frac{4}{nT^4}(\tilde{\beta} - \beta)' \sum_{t\neq\tau\neq\eta}^T \sum_{\tau\neq\eta}^T \sum_{\tau\neq\eta}^T \sum_{i\neq j}^n \sum_{j=1}^n v_{it} v_{i\tau} \tilde{x}_{jt} v_{j\eta} = O_p(\frac{1}{T^3\sqrt{n}})$;

(10) $t = \tau \neq s \neq \eta : \frac{4}{nT^4}(\tilde{\beta} - \beta)' \sum_{t\neq s\neq\eta}^T \sum_{s\neq\eta}^T \sum_{\eta=1}^T \sum_{i\neq j}^n \sum_{j=1}^n v_{is} v_{it} \tilde{x}_{jt} v_{j\eta} = O_p(\frac{1}{T^3\sqrt{n}})$;

(11) $t = \eta \neq s \neq \tau : \frac{4}{nT^4}(\tilde{\beta} - \beta)' \sum_{t\neq s\neq\tau}^T \sum_{s\neq\tau}^T \sum_{\tau=1}^T \sum_{i\neq j}^n \sum_{j=1}^n v_{is} v_{i\tau} \tilde{x}_{jt} v_{jt} = O_p(\frac{1}{T^3\sqrt{n}})$;

(12) $s = \tau \neq t \neq \eta : \frac{4}{nT^4}(\tilde{\beta} - \beta)' \sum_{t\neq s\neq\eta}^T \sum_{s\neq\eta}^T \sum_{\eta=1}^T \sum_{i\neq j}^n \sum_{j=1}^n v_{is}^2 \tilde{x}_{jt} v_{j\eta} = O_p(\frac{1}{T^2})$;

(13) $s = \eta \neq t \neq \tau : \frac{4}{nT^4}(\tilde{\beta} - \beta)' \sum_{t\neq s\neq\tau}^T \sum_{s\neq\tau}^T \sum_{\tau=1}^T \sum_{i\neq j}^n \sum_{j=1}^n v_{is} v_{i\tau} \tilde{x}_{jt} v_{js} = O_p(\frac{1}{T^2\sqrt{nT}})$;

(14) $\tau = \eta \neq t \neq s : \frac{4}{nT^4}(\tilde{\beta} - \beta)' \sum_{t\neq s\neq\tau}^T \sum_{s\neq\tau}^T \sum_{\tau=1}^T \sum_{i\neq j}^n \sum_{j=1}^n v_{is} v_{i\tau} \tilde{x}_{jt} v_{j\tau} = O_p(\frac{1}{T^2\sqrt{nT}})$;

(15) $t \neq s \neq \tau \neq \eta : \frac{4}{nT^4}(\tilde{\beta} - \beta)' \sum_{t\neq s\neq\tau\neq\eta}^T \sum_{s\neq\tau\neq\eta}^T \sum_{\tau\neq\eta}^T \sum_{\eta=1}^T \sum_{i\neq j}^n \sum_{j=1}^n v_{is} v_{i\tau} \tilde{x}_{jt} v_{j\eta} = O_p(\frac{1}{T^2\sqrt{n}})$.

We then obtain the order with the results above:

$$C_2^8 = O_p(\frac{1}{T^2}).$$

Then

$$C_2 = O_p(\frac{1}{nT}) + O_p(\frac{1}{T^2}) + O_p(\frac{1}{T\sqrt{nT}}).$$

■

G.3 Proof of part c_3

$$\begin{aligned} C_3 =& \frac{2}{nT(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s}^{T}\sum_{s=1}^{T} \tilde{v}_t'\tilde{v}_s\tilde{x}_t'\tilde{x}_s(\tilde{\beta}-\beta) \\ =& \frac{2}{nT(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s}^{T}\sum_{s=1}^{T} (v_t-\bar{v}.)'(v_s-\bar{v}.)\tilde{x}_t'\tilde{x}_s(\tilde{\beta}-\beta) \\ =& \sum_{k=1}^{3} C_3^k. \end{aligned}$$

Where $C_3^1 = \frac{2}{nT(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s}^{T}\sum_{s=1}^{T} v_t'v_s\tilde{x}_t'\tilde{x}_s(\tilde{\beta}-\beta)$;
$C_3^2 = -\frac{4}{nT(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s}^{T}\sum_{s=1}^{T} v_t'\bar{v}.\tilde{x}_t'\tilde{x}_s(\tilde{\beta}-\beta)$ and $C_3^3 = \frac{2}{nT(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s}^{T}\sum_{s=1}^{T} \bar{v}_.'\bar{v}.\tilde{x}_t'\tilde{x}_s(\tilde{\beta}-\beta)$.

Lemma 9 *Under Assumption 1, 2, 3 and the null,*

(1) $C_3^1 = O_p(\frac{1}{nT^2})$;

(2) $C_3^2 = O_p(\frac{1}{T^2})$;

(3) $C_3^3 = O_p(\frac{1}{T^2})$.

Proof. Proof of part (1)

$$C_3^1 = \frac{2}{nT(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s}^{T}\sum_{s=1}^{T} v_t'v_s\tilde{x}_t'\tilde{x}_s(\tilde{\beta}-\beta).$$

Then we consider the term $\frac{1}{nT(T-1)}\sum_{t\neq s}^{T}\sum_{s=1}^{T}v_t'v_s\tilde{x}_t'\tilde{x}_s$.

$$
\begin{aligned}
&\frac{1}{nT(T-1)}\sum_{t\neq s}^{T}\sum_{s=1}^{T}v_t'v_s\tilde{x}_t'\tilde{x}_s=\frac{1}{nT(T-1)}\sum_{t\neq s}^{T}\sum_{s=1}^{T}[(\sum_{i=1}^{n}v_{it}v_{is})(\sum_{j=1}^{n}\tilde{x}_{jt}\tilde{x}_{js})]\\
=&\frac{1}{nT(T-1)}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n}v_{it}v_{is}\tilde{x}_{it}\tilde{x}_{is}+\frac{1}{nT(T-1)}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}v_{it}v_{is}\tilde{x}_{jt}\tilde{x}_{js}\\
=&O_p(\frac{1}{T\sqrt{n}})+O_p(\frac{1}{T}).
\end{aligned}
$$

Since $(\tilde{\beta}-\beta)=O_p(\frac{1}{\sqrt{nT}})$, then we obtain $C_3^1=O_p(\frac{1}{nT^2})$.

Proof of part (2)

$$
\begin{aligned}
C_3^2=&-\frac{4}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}v_t'\bar{v}.\tilde{x}_t'\tilde{x}_s(\tilde{\beta}-\beta)\\
=&-\frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}v_t'v_\tau\tilde{x}_t'\tilde{x}_s(\tilde{\beta}-\beta)\\
=&-\frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{n}v_{it}v_{i\tau}\tilde{x}_{jt}\tilde{x}_{js}(\tilde{\beta}-\beta)\\
=&-\frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n}v_{it}v_{i\tau}\tilde{x}_{it}\tilde{x}_{is}(\tilde{\beta}-\beta)\\
&-\frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}v_{it}v_{i\tau}\tilde{x}_{jt}\tilde{x}_{js}(\tilde{\beta}-\beta).
\end{aligned}
$$

There are three cases for expansion of each term, the first term can be expanded as the following:

(1) $t=\tau\neq s: -\frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n}v_{it}^2\tilde{x}_{it}\tilde{x}_{is}(\tilde{\beta}-\beta)=O_p(\frac{1}{nT^2})$;

(2) $t\neq s=\tau: -\frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n}v_{it}v_{is}\tilde{x}_{it}\tilde{x}_{is}(\tilde{\beta}-\beta)=O_p(\frac{1}{nT^3\sqrt{n}})$;

(3) $t\neq s\neq\tau: -\frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n}v_{it}v_{i\tau}\tilde{x}_{it}\tilde{x}_{is}(\tilde{\beta}-\beta)=O_p(\frac{1}{nT^2\sqrt{n}})$.

Similarly, we also need to distinguish 3 cases for the second term:

(1) $t=\tau\neq s: -\frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}v_{it}^2\tilde{x}_{jt}\tilde{x}_{js}(\tilde{\beta}-\beta)=O_p(\frac{1}{T^2})$;

(2) $t\neq s=\tau: -\frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}v_{it}v_{is}\tilde{x}_{jt}\tilde{x}_{js}(\tilde{\beta}-\beta)=O_p(\frac{1}{T^3\sqrt{n}})$;

(3) $t \neq s \neq \tau : -\frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s}^T \sum_{s\neq\tau}^T \sum_{\tau=1}^T \sum_{i\neq j}^n \sum_{j=1}^n v_{it}v_{i\tau}\tilde{x}_{jt}\tilde{x}_{js}(\tilde{\beta}-\beta) = O_p(\frac{1}{T^2\sqrt{n}})$.

As a result,

$$C_3^2 = O_p(\frac{1}{T^2}).$$

Proof of part (3)

$$\begin{aligned}
C_3^3 =& \frac{2}{nT(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s}^T \sum_{s=1}^T \bar{v}'_{\cdot}\bar{v}_{\cdot}\tilde{x}'_t\tilde{x}_s(\tilde{\beta}-\beta) \\
=& \frac{2}{nT^3(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s}^T \sum_{s=1}^T \sum_{\tau=1}^T \sum_{\eta=1}^T v'_\tau v_\eta \tilde{x}'_t\tilde{x}_s(\tilde{\beta}-\beta) \\
=& \frac{2}{nT^3(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s}^T \sum_{s=1}^T \sum_{\tau=1}^T \sum_{\eta=1}^T \sum_{i=1}^n \sum_{j=1}^n v_{i\tau}v_{i\eta}\tilde{x}_{jt}\tilde{x}_{js}(\tilde{\beta}-\beta) \\
=& \frac{2}{nT^3(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s}^T \sum_{s=1}^T \sum_{\tau=1}^T \sum_{\eta=1}^T \sum_{i=1}^n v_{i\tau}v_{i\eta}\tilde{x}_{it}\tilde{x}_{is}(\tilde{\beta}-\beta) \\
&+ \frac{2}{nT^3(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s}^T \sum_{s=1}^T \sum_{\tau=1}^T \sum_{\eta=1}^T \sum_{i\neq j}^n \sum_{j=1}^n v_{i\tau}v_{i\eta}\tilde{x}_{jt}\tilde{x}_{js}(\tilde{\beta}-\beta).
\end{aligned}$$

Here we have 10 cases to distinguish:

I) Two "="s: (1) $t = \tau = \eta \neq s$; (2) $t \neq s = \tau = \eta$; (3) $t = \tau \neq s = \eta$; (4) $t = \eta \neq s = \tau$

II) One "=": (5) $t = \tau \neq s \neq \eta$; (6) $t = \eta \neq s \neq \tau$; (7) $t \neq s = \tau \neq \eta$; (8) $t \neq s = \eta \neq \tau$; (9) $t \neq s \neq \tau = \eta$;

III) No "=": (10) $t \neq s \neq \tau \neq \eta$. Then $\frac{2}{nT^3(T-1)} \sum_{t\neq s}^T \sum_{s=1}^T \sum_{\tau=1}^T \sum_{\eta=1}^T \sum_{i=1}^n v_{i\tau}v_{i\eta}\tilde{x}_{it}\tilde{x}_{is}$ in the first term can be expanded as the following cases:

(1) $t = \tau = \eta \neq s : \frac{2}{nT^3(T-1)} \sum_{t\neq s}^T \sum_{s=1}^T \sum_{i=1}^n v_{it}^2\tilde{x}_{it}\tilde{x}_{is} = O_p(\frac{1}{T^2})$;

(2) $t \neq s = \tau = \eta : \frac{2}{nT^3(T-1)} \sum_{t\neq s}^T \sum_{s=1}^T \sum_{i=1}^n v_{is}^2\tilde{x}_{it}\tilde{x}_{is} = O_p(\frac{1}{T^2})$;

(3) $t = \tau \neq s = \eta : \frac{2}{nT^3(T-1)} \sum_{t\neq s}^T \sum_{s=1}^T \sum_{i=1}^n v_{it}v_{is}\tilde{x}_{it}\tilde{x}_{is} = O_p(\frac{1}{T^3\sqrt{n}})$;

(4) $t = \eta \neq s = \tau : \frac{2}{nT^3(T-1)} \sum_{t\neq s}^T \sum_{s=1}^T \sum_{i=1}^n v_{is}v_{it}\tilde{x}_{it}\tilde{x}_{is} = O_p(\frac{1}{T^3\sqrt{n}})$;

(5) $t = \tau \neq s \neq \eta : \frac{2}{nT^3(T-1)} \sum_{t\neq s\neq\eta}^T \sum_{s\neq\eta}^T \sum_{\eta=1}^T \sum_{i=1}^n v_{it}v_{i\eta}\tilde{x}_{it}\tilde{x}_{is} = O_p(\frac{1}{T^2\sqrt{n}})$;

(6) $t = \eta \neq s \neq \tau : \frac{2}{nT^3(T-1)} \sum_{t\neq s\neq\tau}^T \sum_{s\neq\tau}^T \sum_{\tau=1}^T \sum_{i=1}^n v_{i\tau}v_{it}\tilde{x}_{it}\tilde{x}_{is} = O_p(\frac{1}{T^2\sqrt{n}})$;

(7) $t \neq s = \tau \neq \eta : \frac{2}{nT^3(T-1)} \sum_{t\neq s\neq\eta}^T \sum_{s\neq\eta}^T \sum_{\eta=1}^T \sum_{i=1}^n v_{is}v_{i\eta}\tilde{x}_{it}\tilde{x}_{is} = O_p(\frac{1}{T^2\sqrt{n}})$;

(8) $t \neq s = \eta \neq \tau$: $\frac{2}{nT^3(T-1)} \sum_{t\neq s\neq \tau}^{T} \sum_{s\neq \tau}^{T} \sum_{\tau=1}^{T} \sum_{i=1}^{n} v_{i\tau} v_{is} \tilde{x}_{it} \tilde{x}_{is} = O_p(\frac{1}{T^2\sqrt{n}})$;

(9) $t \neq s \neq \tau = \eta$: $\frac{2}{nT^3(T-1)} \sum_{t\neq s\neq \tau}^{T} \sum_{s\neq \tau}^{T} \sum_{\tau=1}^{T} \sum_{i=1}^{n} v_{i\tau}^2 \tilde{x}_{it} \tilde{x}_{is} = O_p(\frac{1}{T})$;

(10) $t \neq s \neq \tau \neq \eta$: $\frac{2}{nT^3(T-1)} \sum_{t\neq s\neq \tau\neq\eta}^{T} \sum_{s\neq \tau\neq\eta}^{T} \sum_{\tau\neq\eta}^{T} \sum_{\eta=1}^{T} \sum_{i=1}^{n} v_{i\tau} v_{i\eta} \tilde{x}_{it} \tilde{x}_{is} = O_p(\frac{1}{T\sqrt{n}})$.

Similarly, $\frac{2}{nT^3(T-1)} \sum_{t\neq s}^{T} \sum_{s=1}^{T} \sum_{\tau=1}^{T} \sum_{\eta=1}^{T} \sum_{i\neq j}^{n} \sum_{j=1}^{n} v_{i\tau} v_{i\eta} \tilde{x}_{jt} \tilde{x}_{js}$ in the second term can be expanded into ten cases:

(1) $t = \tau = \eta \neq s$: $\frac{2}{nT^3(T-1)} \sum_{t\neq s}^{T} \sum_{s=1}^{T} \sum_{i\neq j}^{n} \sum_{j=1}^{n} v_{it}^2 \tilde{x}_{jt} \tilde{x}_{js} = O_p(\frac{n}{T^2})$;

(2) $t \neq s = \tau = \eta$: $\frac{2}{nT^3(T-1)} \sum_{t\neq s}^{T} \sum_{s=1}^{T} \sum_{i\neq j}^{n} \sum_{j=1}^{n} v_{is}^2 \tilde{x}_{jt} \tilde{x}_{js} = O_p(\frac{n}{T^2})$;

(3) $t = \tau \neq s = \eta$: $\frac{2}{nT^3(T-1)} \sum_{t\neq s}^{T} \sum_{s=1}^{T} \sum_{i\neq j}^{n} \sum_{j=1}^{n} v_{it} v_{is} \tilde{x}_{jt} \tilde{x}_{js} = O_p(\frac{\sqrt{n}}{T^3})$;

(4) $t = \eta \neq s = \tau$: $\frac{2}{nT^3(T-1)} \sum_{t\neq s}^{T} \sum_{s=1}^{T} \sum_{i\neq j}^{n} \sum_{j=1}^{n} v_{is} v_{it} \tilde{x}_{jt} \tilde{x}_{js} = O_p(\frac{\sqrt{n}}{T^3})$;

(5) $t = \tau \neq s \neq \eta$: $\frac{2}{nT^3(T-1)} \sum_{t\neq s\neq \eta}^{T} \sum_{s\neq \eta}^{T} \sum_{\eta=1}^{T} \sum_{i\neq j}^{n} \sum_{j=1}^{n} v_{it} v_{i\eta} \tilde{x}_{jt} \tilde{x}_{js} = O_p(\frac{\sqrt{n}}{T^2})$;

(6) $t = \eta \neq s \neq \tau$: $\frac{2}{nT^3(T-1)} \sum_{t\neq s\neq \tau}^{T} \sum_{s\neq \tau}^{T} \sum_{\tau=1}^{T} \sum_{i\neq j}^{n} \sum_{j=1}^{n} v_{i\tau} v_{it} \tilde{x}_{jt} \tilde{x}_{js} = O_p(\frac{\sqrt{n}}{T^2})$;

(7) $t \neq s = \tau \neq \eta$: $\frac{2}{nT^3(T-1)} \sum_{t\neq s\neq \eta}^{T} \sum_{s\neq \eta}^{T} \sum_{\eta=1}^{T} \sum_{i\neq j}^{n} \sum_{j=1}^{n} v_{is} v_{i\eta} \tilde{x}_{jt} \tilde{x}_{js} = O_p(\frac{\sqrt{n}}{T^2})$;

(8) $t \neq s = \eta \neq \tau$: $\frac{2}{nT^3(T-1)} \sum_{t\neq s\neq \tau}^{T} \sum_{s\neq \tau}^{T} \sum_{\tau=1}^{T} \sum_{i\neq j}^{n} \sum_{i=1}^{n} v_{i\tau} v_{is} \tilde{x}_{jt} \tilde{x}_{js} = O_p(\frac{\sqrt{n}}{T^2})$;

(9) $t \neq s \neq \tau = \eta$: $\frac{2}{nT^3(T-1)} \sum_{t\neq s\neq \tau}^{T} \sum_{s\neq \tau}^{T} \sum_{\tau=1}^{T} \sum_{i\neq j}^{n} \sum_{j=1}^{n} v_{i\tau}^2 \tilde{x}_{jt} \tilde{x}_{js} = O_p(\frac{n}{T})$;

(10) $t \neq s \neq \tau \neq \eta$: $\frac{2}{nT^3(T-1)} \sum_{t\neq s\neq \tau\neq\eta}^{T} \sum_{s\neq \tau\neq\eta}^{T} \sum_{\tau\neq\eta}^{T} \sum_{\eta=1}^{T} \sum_{i\neq j}^{n} \sum_{j=1}^{n} v_{i\tau} v_{i\eta} \tilde{x}_{jt} \tilde{x}_{js} = O_p(\frac{\sqrt{n}}{T})$.

Since $\tilde{\beta} - \beta = O_p(\frac{1}{\sqrt{nT}})$, then we have $C_3^3 = O_p(\frac{1}{T^2})$. ■

Then from Lemma 9, we have

$$C_3 = O_p(\frac{1}{T^2}).$$

G.4 Proof of part c_4

$$\begin{aligned} C_4 =& \frac{2}{nT(T-1)} (\tilde{\beta} - \beta)' \sum_{t\neq s}^{T} \sum_{s=1}^{T} (\tilde{x}_t' \tilde{v}_s \tilde{v}_s' \tilde{x}_t)(\tilde{\beta} - \beta) \\ =& \frac{2}{nT(T-1)} (\tilde{\beta} - \beta)' \sum_{t\neq s}^{T} \sum_{s=1}^{T} (\tilde{x}_t' (v_s - \bar{v}_.)(v_s - \bar{v}_.)' \tilde{x}_t)(\tilde{\beta} - \beta) \\ =& \sum_{k=1}^{3} C_4^k. \end{aligned}$$

Where $C_4^1 = \frac{2}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\tilde{x}_t' v_s v_s' \tilde{x}_t(\tilde{\beta}-\beta)$;
$C_4^2 = -\frac{4}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\tilde{x}_t' v_s \bar{v}'_. \tilde{x}_t(\tilde{\beta}-\beta)$ and $C_4^3 = \frac{2}{nT}(\tilde{\beta}-\beta)'\sum_{t=1}^{T}\tilde{x}_t'\bar{v}_. \bar{v}'_. \tilde{x}_t(\tilde{\beta}-\beta)$.

Lemma 10 *Under Assumption 1, 2, 3 and the null,*

(1) $C_4^1 = O_p(\frac{1}{nT})$;

(2) $C_4^2 = O_p(\frac{1}{nT^2})$;

(3) $C_4^3 = O_p(\frac{1}{nT^2})$.

Proof. Proof of part (1)

$$
\begin{aligned}
C_4^1 =& \frac{2}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\tilde{x}_t' v_s v_s' \tilde{x}_t(\tilde{\beta}-\beta)\\
=& \frac{2}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{n}\tilde{x}_{it} v_{is} v_{js} \tilde{x}_{jt}(\tilde{\beta}-\beta)\\
=& \frac{2}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n}\tilde{x}_{it}^2 v_{is}^2(\tilde{\beta}-\beta)\\
&+ \frac{2}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n}\tilde{x}_{it} v_{is} v_{js} \tilde{x}_{jt}(\tilde{\beta}-\beta)\\
=& O_p(\frac{1}{nT}) + O_p(\frac{1}{nT^2}) = O_p(\frac{1}{nT}).
\end{aligned}
$$

Proof of part (2)

$$
\begin{aligned}
C_4^2 =& -\frac{4}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\tilde{x}_t' v_s \bar{v}'_. \tilde{x}_t(\tilde{\beta}-\beta)\\
=& -\frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\tilde{x}_t' v_s v_\tau' \tilde{x}_t(\tilde{\beta}-\beta)\\
=& -\frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{i=1}\tilde{x}_{it}^2 v_{is} v_{i\tau}(\tilde{\beta}-\beta)\\
&- \frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{i=1}\tilde{x}_{it} v_{is} v_{j\tau} \tilde{x}_{jt}(\tilde{\beta}-\beta).
\end{aligned}
$$

Now, we consider the term $\frac{1}{nT^2(T-1)}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{i=1}\tilde{x}_{it}^2 v_{is}v_{i\tau}$, there are 3 cases for expanding: (1) $t=\tau\neq s$; (2) $t\neq s=\tau$; (3) $t\neq s\neq\tau$. Then we can distinguish the term to be the following cases:

(1) $t=\tau\neq s$: $\frac{1}{nT^2(T-1)}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n}\tilde{x}_{it}^2 v_{is}v_{it}=O_p(\frac{1}{T^2\sqrt{n}})$;

(2) $t\neq s=\tau$: $\frac{1}{nT^2(T-1)}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n}\tilde{x}_{it}^2 v_{is}^2=O_p(\frac{1}{T})$;

(3) $t\neq s\neq\tau$: $\frac{1}{nT^2(T-1)}\sum_{t\neq s}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}\tilde{x}_{it}^2 v_{is}v_{i\tau}=O_p(\frac{1}{T\sqrt{n}})$.

Then, similarly, we can also expand the term $\frac{1}{nT^2(T-1)}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{i=1}\tilde{x}_{it}v_{is}v_{j\tau}\tilde{x}_{jt}$ to 3 cases:

(1) $t=\tau\neq s$: $\frac{1}{nT^2(T-1)}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}\tilde{x}_{it}v_{is}v_{jt}\tilde{x}_{jt}=O_p(\frac{1}{T^2})$;

(2) $t\neq s=\tau$: $\frac{1}{nT^2(T-1)}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}\tilde{x}_{it}v_{is}v_{js}\tilde{x}_{jt}=O_p(\frac{1}{T\sqrt{T}})$;

(3) $t\neq s\neq\tau$: $\frac{1}{nT^2(T-1)}\sum_{t\neq s}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}\tilde{x}_{it}v_{is}v_{j\tau}\tilde{x}_{jt}=O_p(\frac{1}{T})$.

Hence, with $\tilde{\beta}-\beta=O_p(\frac{1}{\sqrt{nT}})$, then

$$C_4^2=O_p(\frac{1}{nT^2}).$$

Proof of C_4^3

$$\begin{aligned}
C_4^3=&\frac{2}{nT}(\tilde{\beta}-\beta)'\sum_{t=1}^{T}\tilde{x}_t'\bar{v}.\bar{v}'.\tilde{x}_t(\tilde{\beta}-\beta)\\
=&\frac{2}{nT^3}(\tilde{\beta}-\beta)'\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\tilde{x}_t'v_s v_\tau'\tilde{x}_t(\tilde{\beta}-\beta)\\
=&\frac{2}{nT^3}(\tilde{\beta}-\beta)'\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n}\tilde{x}_{it}^2 v_{is}v_{i\tau}(\tilde{\beta}-\beta)\\
&+\frac{2}{nT^3}(\tilde{\beta}-\beta)'\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}\tilde{x}_{it}v_{is}v_{j\tau}\tilde{x}_{jt}(\tilde{\beta}-\beta).
\end{aligned}$$

We need to distinguish five cases of (t, s, τ):I) Two "="s: (1) $t=s=\tau$; II) One "=": (2) $t=s\neq\tau$; (3) $t=\tau\neq s$; (4) $t\neq s=\tau$; III) No "=": (5) $t\neq s\neq\tau$. Then the term $\frac{1}{nT^3}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n}\tilde{x}_{it}^2 v_{is}v_{i\tau}$ in the first term can be distinguished into five cases:

(1) $t=s=\tau$: $\frac{1}{nT^3}\sum_{t=1}^{T}\sum_{i=1}^{n}\tilde{x}_{it}^2 v_{it}^2=O_p(\frac{1}{T^2})$;

(2) $t = s \neq \tau : \frac{1}{nT^3}\sum_{t\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n}\tilde{x}_{it}^2 v_{it}v_{i\tau} = O_p(\frac{1}{T^2\sqrt{n}})$;

(3) $t = \tau \neq s : \frac{1}{nT^3}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n}\tilde{x}_{it}^2 v_{it}v_{is} = O_p(\frac{1}{T^2\sqrt{n}})$;

(4) $t \neq s = \tau : \frac{1}{nT^3}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n}\tilde{x}_{it}^2 v_{is}^2 = O_p(\frac{1}{T})$;

(5) $t \neq s \neq \tau : \frac{1}{nT^3}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n}\tilde{x}_{it}^2 v_{is}v_{i\tau} = O_p(\frac{1}{T\sqrt{n}})$.

The term $\frac{1}{nT^3}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}\tilde{x}_{it}v_{is}v_{j\tau}\tilde{x}_{jt}$ can be similarly distinguished:

(1) $t = s = \tau : \frac{1}{nT^3}\sum_{t=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}\tilde{x}_{it}v_{it}v_{jt}\tilde{x}_{jt} = O_p(\frac{1}{T^2\sqrt{T}})$;

(2) $t = s \neq \tau : \frac{1}{nT^3}\sum_{t\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}\tilde{x}_{it}v_{it}v_{j\tau}\tilde{x}_{jt} = O_p(\frac{1}{T^2})$;

(3) $t = \tau \neq s : \frac{1}{nT^3}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}\tilde{x}_{it}v_{is}v_{jt}\tilde{x}_{jt} = O_p(\frac{1}{T^2})$;

(4) $t \neq s = \tau : \frac{1}{nT^3}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i\neq j}^{n}\sum_{i=1}^{n}\tilde{x}_{it}v_{is}v_{js}\tilde{x}_{jt}^2 = O_p(\frac{1}{T^2})$;

(5) $t \neq s \neq \tau : \frac{1}{nT^3}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{i=1}^{n}\tilde{x}_{it}v_{is}v_{j\tau}\tilde{x}_{jt} = O_p(\frac{1}{T})$.

With $\tilde{\beta} - \beta = O_p(\frac{1}{\sqrt{nT}})$, we obtain

$$C_4^3 = O_p(\frac{1}{nT^2}).$$

■

From Lemma 10, we then have $C_4 = O_p(\frac{1}{nT})$.

G.5 Proof of part c_5

$$\begin{aligned} C_5 =& \frac{2}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}(\tilde{x}_t'(v_s-\bar{v}_.)(v_t-\bar{v}_.)'\tilde{x}_s)(\tilde{\beta}-\beta) \\ =& \sum_{k=1}^{3} C_5^k. \end{aligned}$$

Where (1) $C_5^1 = \frac{2}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}(\tilde{x}_t'v_s v_t'\tilde{x}_s)(\tilde{\beta}-\beta)$; (2) $C_5^2 = -\frac{4}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}(\tilde{x}_t'v_s\bar{v}_.'\tilde{x}_s)(\tilde{\beta}-\beta)$ and (3) $C_5^3 = \frac{2}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}(\tilde{x}_t'\bar{v}_.\bar{v}_.'\tilde{x}_s)(\tilde{\beta}-\beta)$.

Lemma 11 *Under Assumption 1, 2, 3 and the null,*

(1) $C_5^1 = O_p(\frac{1}{nT^2})$;

(2) $C_5^2 = O_p(\frac{1}{nT^2})$;

(3) $C_5^3 = O_p(\frac{1}{nT^2})$.

Proof. Proof of part (1)

$$\begin{aligned}C_5^1 =&\frac{2}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n}(\tilde{x}_{it}v_{is}v_{it}\tilde{x}_{is})(\tilde{\beta}-\beta)\\&+\frac{2}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}(\tilde{x}_{it}v_{is}v_{jt}\tilde{x}_{js})(\tilde{\beta}-\beta)\\=&O_p(\frac{1}{nT^2\sqrt{n}})+O_p(\frac{1}{nT^2})=O_p(\frac{1}{nT^2}).\end{aligned}$$

Proof of part (2)

$$\begin{aligned}C_5^2 =&-\frac{4}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}(\tilde{x}_t'v_s(\frac{1}{T}\sum_{\tau=1}^{T}v_\tau)'\tilde{x}_s)(\tilde{\beta}-\beta)\\=&-\frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}(\tilde{x}_t'v_sv_\tau'\tilde{x}_s)(\tilde{\beta}-\beta)\\=&-\frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n}(\tilde{x}_{it}v_{is}v_{i\tau}\tilde{x}_{is})(\tilde{\beta}-\beta)\\&-\frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}(\tilde{x}_{it}v_{is}v_{j\tau}\tilde{x}_{js})(\tilde{\beta}-\beta).\end{aligned}$$

The term $-\frac{4}{nT^2(T-1)}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n}\tilde{x}_{it}v_{is}v_{i\tau}\tilde{x}_{is}$ can be expanded as the following 3 expressions:

(1) $t=\tau\neq s: -\frac{4}{nT^2(T-1)}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\tilde{x}_{it}v_{is}v_{it}\tilde{x}_{is}=O_p(\frac{1}{T^2\sqrt{n}})$;

(2) $t\neq(s=\tau): -\frac{4}{nT^2(T-1)}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\tilde{x}_{it}v_{is}^2\tilde{x}_{is}=O_p(\frac{1}{T})$;

(3) $t\neq(s=\tau): -\frac{4}{nT^2(T-1)}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\tilde{x}_{it}v_{is}v_{i\tau}\tilde{x}_{is}=O_p(\frac{1}{T\sqrt{n}})$.

Similarly, $-\frac{2}{nT^2(T-1)}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}\tilde{x}_{it}v_{is}v_{j\tau}\tilde{x}_{js}$ can be expanded as the following terms:

(1) $t=\tau\neq s: -\frac{4}{nT^2(T-1)}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}\tilde{x}_{it}v_{is}v_{jt}\tilde{x}_{js}=O_p(\frac{1}{T^2})$;

(2) $t \neq (s = \tau) : -\frac{4}{nT^2(T-1)} \sum_{t\neq s}^T \sum_{s=1}^T \sum_{i\neq j}^n \sum_{j=1}^n \tilde{x}_{it} v_{is} v_{js} \tilde{x}_{js} = O_p(\frac{1}{T\sqrt{T}})$;

(3) $t \neq s \neq \tau : -\frac{4}{nT^2(T-1)} \sum_{t\neq s}^T \sum_{s=1}^T \sum_{\tau=1}^T \sum_{i\neq j}^n \sum_{j=1}^n \tilde{x}_{it} v_{is} v_{j\tau} \tilde{x}_{js} = O_p(\frac{1}{T})$.

Since $\tilde{\beta} - \beta = O_p(\frac{1}{\sqrt{nT}})$ and with results above, we can get

$$C_5^2 = O_p(\frac{1}{nT^2}).$$

Proof of (3)

$$\begin{aligned} C_5^3 =& \frac{2}{nT(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s}^T \sum_{s=1}^T (\tilde{x}_t'(\frac{1}{T}\sum_{\tau=1}^{\tau} v_\tau)(\frac{1}{T}\sum_{\eta=1}^T v_\eta)' \tilde{x}_s)(\tilde{\beta}-\beta) \\ =& \frac{2}{nT^3(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s}^T \sum_{s=1}^T \sum_{\tau=1}^T \sum_{\eta=1}^T \tilde{x}_t' v_\tau v_\eta' \tilde{x}_s (\tilde{\beta}-\beta) \\ =& \frac{2}{nT^3(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s}^T \sum_{s=1}^T \sum_{\tau=1}^T \sum_{\eta=1}^T \sum_{i=1}^n \tilde{x}_{it} v_{i\tau} v_{i\eta} \tilde{x}_{is} (\tilde{\beta}-\beta) \\ &+ \frac{2}{nT^3(T-1)}(\tilde{\beta}-\beta)' \sum_{t\neq s}^T \sum_{s=1}^T \sum_{\tau=1}^T \sum_{\eta=1}^T \sum_{i=1}^n \tilde{x}_{it} v_{i\tau} v_{j\eta} \tilde{x}_{js} (\tilde{\beta}-\beta). \end{aligned}$$

Then we need to distinguish the cases of $(t \neq s, \tau, \eta)$:

I) Two "="s: (1) $(t = \tau = \eta) \neq s$; (2) $t \neq s = \tau = \eta$; (3) $t = \tau \neq s = \eta$; (4) $t = \eta \neq s = \tau$;

2) One "=": (5) $t = \tau \neq s \neq \eta$; (6) $t = \eta \neq s \neq \tau$; (7) $t \neq (s = \tau) \neq \eta$; (8) $t \neq (s = \eta) \neq \tau$;(9)$t \neq s \neq (\tau \neq \eta)$;

3) No "=": (10) $t \neq s \neq \tau \neq \eta$.

We need to calculate $\frac{2}{nT^3(T-1)} \sum_{t\neq s}^T \sum_{s=1}^T \sum_{\tau=1}^T \sum_{\eta=1}^T \sum_{i=1}^n \tilde{x}_{it} v_{i\tau} v_{i\eta} \tilde{x}_{is}$ to get the order of C_5^3,

, it can be expanded as the following ten terms:

(1) $(t = \tau = \eta) \neq s : \frac{2}{nT^3(T-1)} \sum_{t\neq s}^T \sum_{s=1}^T \sum_{i=1}^n \tilde{x}_{it} v_{it}^2 \tilde{x}_{is} = O_p(\frac{1}{T^2})$;

(2) $t \neq s = \tau = \eta : \frac{2}{nT^3(T-1)} \sum_{t\neq s}^T \sum_{s=1}^T \sum_{i=1}^n \tilde{x}_{it} v_{is}^2 \tilde{x}_{is} = O_p(\frac{1}{T^2})$;

(3) $t = \tau \neq s = \eta : \frac{2}{nT^3(T-1)} \sum_{t\neq s}^T \sum_{s=1}^T \sum_{i=1}^n \tilde{x}_{it} v_{it} v_{is} \tilde{x}_{is} = O_p(\frac{1}{T^3\sqrt{n}})$;

(4) $t = \eta \neq s = \tau : \frac{2}{nT^3(T-1)} \sum_{t\neq s}^T \sum_{s=1}^T \sum_{i=1}^n \tilde{x}_{it} v_{is} v_{it} \tilde{x}_{is} = O_p(\frac{1}{T^3\sqrt{n}})$;

(5) $t = \tau \neq s \neq \eta : \frac{2}{nT^3(T-1)} \sum_{t\neq s\neq \eta}^T \sum_{s\neq \eta}^T \sum_{\eta=1}^T \sum_{i=1}^n \tilde{x}_{it} v_{it} v_{i\eta} \tilde{x}_{is} = O_p(\frac{1}{T^2\sqrt{n}})$;

(6) $t=\eta\neq s\neq\tau$: $\frac{2}{nT^3(T-1)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n}\tilde{x}_{it}v_{i\tau}v_{it}\tilde{x}_{is}=O_p(\frac{1}{T^2\sqrt{n}})$;

(7) $t\neq(s=\tau)\neq\eta$: $\frac{2}{nT^3(T-1)}\sum_{t\neq s\neq\eta}^{T}\sum_{s\neq\eta}^{T}\sum_{\eta=1}^{T}\sum_{i=1}^{n}\tilde{x}_{it}v_{is}v_{i\eta}\tilde{x}_{is}=O_p(\frac{1}{T^2\sqrt{n}})$;

(8) $t\neq(s=\eta)\neq\tau$: $\frac{2}{nT^3(T-1)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n}\tilde{x}_{it}v_{i\tau}v_{is}\tilde{x}_{is}=O_p(\frac{1}{T^2\sqrt{n}})$;

(9) $t\neq s\neq(\tau=\eta)$: $\frac{2}{nT^3(T-1)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n}\tilde{x}_{it}v_{i\tau}^2\tilde{x}_{is}=O_p(\frac{1}{T})$;

(10) $t\neq s\neq\tau\neq\eta$: $\frac{2}{nT^3(T-1)}\sum_{t\neq s\neq\tau\neq\eta}^{T}\sum_{s\neq\tau\neq\eta}^{T}\sum_{\tau\neq\eta}^{T}\sum_{\eta=1}^{T}\sum_{i=1}^{n}\tilde{x}_{it}v_{i\tau}v_{i\eta}\tilde{x}_{is}=O_p(\frac{1}{T\sqrt{n}})$.

Moreover, $\frac{2}{nT^3(T-1)}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{\eta=1}^{T}\sum_{i=1}^{n}\tilde{x}_{it}v_{i\tau}v_{j\eta}\tilde{x}_{js}$ can be expanded as the following expressions:

(1) $(t=\tau=\eta)\neq s$: $\frac{2}{nT^3(T-1)}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}\tilde{x}_{it}v_{it}v_{jt}\tilde{x}_{js}=O_p(\frac{1}{T^2\sqrt{T}})$;

(2) $t\neq s=\tau=\eta$: $\frac{2}{nT^3(T-1)}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}\tilde{x}_{it}v_{is}v_{js}\tilde{x}_{js}=O_p(\frac{1}{T^2\sqrt{T}})$;

(3) $t=\tau\neq s=\eta$: $\frac{2}{nT^3(T-1)}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}\tilde{x}_{it}v_{it}v_{js}\tilde{x}_{js}=O_p(\frac{1}{T^3})$;

(4) $t=\eta\neq s=\tau$: $\frac{2}{nT^3(T-1)}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}\tilde{x}_{it}v_{is}v_{jt}\tilde{x}_{js}=O_p(\frac{1}{T^3})$;

(5) $t=\tau\neq s\neq\eta$: $\frac{2}{nT^3(T-1)}\sum_{t\neq s\neq\eta}^{T}\sum_{s\neq\eta}^{T}\sum_{\eta=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}\tilde{x}_{it}v_{it}v_{j\eta}\tilde{x}_{js}=O_p(\frac{1}{T^2})$;

(6) $t=\eta\neq s\neq\tau$: $\frac{2}{nT^3(T-1)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}\tilde{x}_{it}v_{i\tau}v_{jt}\tilde{x}_{js}=O_p(\frac{1}{T^2})$;

(7) $t\neq(s=\tau)\neq\eta$: $\frac{2}{nT^3(T-1)}\sum_{t\neq s\neq\eta}^{T}\sum_{s\neq\eta}^{T}\sum_{\eta=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}\tilde{x}_{it}v_{is}v_{j\eta}\tilde{x}_{js}=O_p(\frac{1}{T^2})$;

(8) $t\neq(s=\eta)\neq\tau$: $\frac{2}{nT^3(T-1)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}\tilde{x}_{it}v_{i\tau}v_{js}\tilde{x}_{js}=O_p(\frac{1}{T^2})$;

(9) $t\neq s\neq(\tau=\eta)$: $\frac{2}{nT^3(T-1)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}\tilde{x}_{it}v_{i\tau}v_{j\tau}\tilde{x}_{js}=O_p(\frac{1}{T\sqrt{T}})$;

(10) $t\neq s\neq\tau\neq\eta$: $\frac{2}{nT^3(T-1)}\sum_{t\neq s\neq\tau\neq\eta}^{T}\sum_{s\neq\tau\neq\eta}^{T}\sum_{\tau\neq\eta}^{T}\sum_{\eta=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}\tilde{x}_{it}v_{i\tau}v_{j\eta}\tilde{x}_{js}=O_p(\frac{1}{T})$.

Since $\tilde{\beta}-\beta=O_p(\frac{1}{\sqrt{nT}})$, then we have

$$C_5^3=O_p(\frac{1}{nT^2}).$$

■

Then Lemma 11 is proved, then we have

$$C_5=O_p(\frac{1}{nT^2}).$$

G.6 Proof of part c_6

$$
\begin{aligned}
C_6 =& -\frac{4}{nT(T-1)}(\tilde{\beta}-\beta)'(\sum_{t\neq s}^{T}\sum_{s=1}^{T}\tilde{x}_t'\tilde{v}_s(\tilde{\beta}-\beta)\tilde{x}_t'\tilde{x}_s)(\tilde{\beta}-\beta) \\
=& -\frac{4}{nT(T-1)}(\tilde{\beta}-\beta)'(\sum_{t\neq s}^{T}\sum_{s=1}^{T}\tilde{x}_t'(v_s-\bar{v}.)(\tilde{\beta}-\beta)\tilde{x}_t'\tilde{x}_s)(\tilde{\beta}-\beta) \\
=& -\frac{4}{nT(T-1)}(\tilde{\beta}-\beta)'(\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{n}\tilde{x}_{it}v_{is}(\tilde{\beta}-\beta)\tilde{x}_{jt}\tilde{x}_{js})(\tilde{\beta}-\beta) \\
& -\frac{4}{nT^2(T-1)}(\tilde{\beta}-\beta)'(\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{n}\tilde{x}_{it}v_{i\tau}(\tilde{\beta}-\beta)\tilde{x}_{jt}\tilde{x}_{js})(\tilde{\beta}-\beta).
\end{aligned}
$$

Since

$$
\begin{aligned}
& -\frac{4}{nT(T-1)}(\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{n}\tilde{x}_{it}v_{is}\tilde{x}_{jt}\tilde{x}_{js}) \\
=& -\frac{4}{nT(T-1)}(\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n}\tilde{x}_{it}^2v_{is}\tilde{x}_{is}) - \frac{4}{nT(T-1)}(\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}\tilde{x}_{it}v_{is}\tilde{x}_{jt}\tilde{x}_{js}) \\
=& O_p(\frac{1}{\sqrt{nT}}) + O_p(\frac{\sqrt{n}}{\sqrt{T}}),
\end{aligned}
$$

and

$$
\begin{aligned}
& -\frac{4}{nT^2(T-1)}(\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{n}\tilde{x}_{it}v_{i\tau}\tilde{x}_{jt}\tilde{x}_{js}) \\
=& -\frac{4}{nT^2(T-1)}(\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{n}\tilde{x}_{it}v_{it}\tilde{x}_{jt}\tilde{x}_{js}) - \frac{4}{nT^2(T-1)}(\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{n}\tilde{x}_{it}v_{is}\tilde{x}_{jt}\tilde{x}_{js}) \\
& -\frac{4}{nT^2(T-1)}(\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{n}\tilde{x}_{it}v_{i\tau}\tilde{x}_{jt}\tilde{x}_{js}) \\
=& O_p(\frac{\sqrt{n}}{T\sqrt{T}}) + O_p(\frac{\sqrt{n}}{T\sqrt{T}}) + O_p(\frac{\sqrt{n}}{\sqrt{T}}).
\end{aligned}
$$

With $\tilde{\beta}-\beta=O_p(\frac{1}{\sqrt{nT}})$, then we have

$$C_6=O_p(\frac{1}{nT^2}).$$

G.7 Proof of part c_7

$$\begin{aligned}
C_7=&\frac{1}{nT(T-1)}(\tilde{\beta}-\beta)'(\sum_{t\neq s}^{T}\sum_{s=1}^{T}\tilde{x}_t'\tilde{x}_s(\tilde{\beta}-\beta)(\tilde{\beta}-\beta)'\tilde{x}_t'\tilde{x}_s)(\tilde{\beta}-\beta)\\
=&\frac{1}{nT(T-1)}(\tilde{\beta}-\beta)'(\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{n}\tilde{x}_{it}'\tilde{x}_{is}(\tilde{\beta}-\beta)(\tilde{\beta}-\beta)'\tilde{x}_{jt}'\tilde{x}_{js})(\tilde{\beta}-\beta)\\
=&\frac{1}{nT(T-1)}(\tilde{\beta}-\beta)'(\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n}\tilde{x}_{it}'\tilde{x}_{is}(\tilde{\beta}-\beta)(\tilde{\beta}-\beta)'\tilde{x}_{it}'\tilde{x}_{is})(\tilde{\beta}-\beta)\\
&+\frac{1}{nT(T-1)}(\tilde{\beta}-\beta)'(\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}\tilde{x}_{it}'\tilde{x}_{is}(\tilde{\beta}-\beta)(\tilde{\beta}-\beta)'\tilde{x}_{jt}'\tilde{x}_{js})(\tilde{\beta}-\beta)\\
=&O_p(\frac{1}{n^2T^2})+O_p(\frac{1}{nT^2})=O_p(\frac{1}{nT^2}).
\end{aligned}$$

H Proof of part (2) of Proposition 2

$$\begin{aligned}
&\frac{1}{n}\hat{M}_{4,T}=\frac{1}{nT(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\hat{v}_t'\hat{v}_s\hat{v}_s'\hat{v}_\tau\\
=&\frac{1}{nT(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}(\tilde{v}_t'\tilde{v}_s-\tilde{v}_t'\tilde{x}_s(\tilde{\beta}-\beta)-(\tilde{\beta}-\beta)'\tilde{x}_t'\tilde{v}_s+(\tilde{\beta}-\beta)'\tilde{x}_t'\tilde{x}_s(\tilde{\beta}-\beta))\\
&(\tilde{v}_s'\tilde{v}_\tau-\tilde{v}_s'\tilde{x}_\tau(\tilde{\beta}-\beta)-(\tilde{\beta}-\beta)'\tilde{x}_s'\tilde{v}_\tau+(\tilde{\beta}-\beta)'\tilde{x}_s\tilde{x}_\tau(\tilde{\beta}-\beta))=\sum_{k=1}^{10}D_k.
\end{aligned}$$

Where

$D_1=\frac{1}{nT(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\tilde{v}_t'\tilde{v}_s\tilde{v}_s'\tilde{v}_\tau;$

$D_2=-\frac{2}{nT(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\tilde{v}_t'\tilde{v}_s\tilde{v}_s'\tilde{x}_\tau(\tilde{\beta}-\beta);$

$D_3=-\frac{2}{nT(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\tilde{v}_t'\tilde{v}_s(\tilde{\beta}-\beta)'\tilde{x}_s'\tilde{v}_\tau;$

$D_4 = \frac{2}{nT(T-1)(T-2)} \sum_{t\neq s\neq \tau}^{T} \sum_{s\neq\tau}^{T} \sum_{\tau=1}^{T} \tilde{v}_t'\tilde{v}_s(\tilde{\beta}-\beta)'\tilde{x}_s'\tilde{x}_\tau(\tilde{\beta}-\beta);$

$D_5 = \frac{2}{nT(T-1)(T-2)} \sum_{t\neq s\neq \tau}^{T} \sum_{s\neq\tau}^{T} \sum_{\tau=1}^{T} \tilde{v}_t'\tilde{x}_s(\tilde{\beta}-\beta)\tilde{v}_s'\tilde{x}_\tau(\tilde{\beta}-\beta);$

$D_6 = \frac{1}{nT(T-1)(T-2)} \sum_{t\neq s\neq \tau}^{T} \sum_{s\neq\tau}^{T} \sum_{\tau=1}^{T} \tilde{v}_t'\tilde{x}_s(\tilde{\beta}-\beta)(\tilde{\beta}-\beta)'\tilde{x}_s'\tilde{v}_\tau;$

$D_7 = -\frac{2}{nT(T-1)(T-2)} \sum_{t\neq s\neq \tau}^{T} \sum_{s\neq\tau}^{T} \sum_{\tau=1}^{T} \tilde{v}_t'\tilde{x}_s(\tilde{\beta}-\beta)(\tilde{\beta}-\beta)'\tilde{x}_s'\tilde{x}_\tau(\tilde{\beta}-\beta);$

$D_8 = \frac{1}{nT(T-1)(T-2)} \sum_{t\neq s\neq \tau}^{T} \sum_{s\neq\tau}^{T} \sum_{\tau=1}^{T} (\tilde{\beta}-\beta)'\tilde{x}_t'\tilde{v}_s\tilde{v}_s'\tilde{x}_\tau(\tilde{\beta}-\beta);$

$D_9 = -\frac{2}{nT(T-1)(T-2)} \sum_{t\neq s\neq \tau}^{T} \sum_{s\neq\tau}^{T} \sum_{\tau=1}^{T} (\tilde{\beta}-\beta)'\tilde{x}_t'\tilde{v}_s(\tilde{\beta}-\beta)'\tilde{x}_s'\tilde{x}_\tau(\tilde{\beta}-\beta);$

$D_{10} = \frac{1}{nT(T-1)(T-2)} \sum_{t\neq s\neq \tau}^{T} \sum_{s\neq\tau}^{T} \sum_{\tau=1}^{T} (\tilde{\beta}-\beta)'\tilde{x}_t\tilde{x}_s(\tilde{\beta}-\beta)(\tilde{\beta}-\beta)'\tilde{x}_s'\tilde{x}_\tau(\tilde{\beta}-\beta).$

Lemma 12 *Under Assumption 1, 2, 3 and the null,*

(d_1) $D_1 = \frac{1}{n}M_{4,T} - \frac{2\sigma_v^4}{T} + \frac{T-1}{T^2}\sigma_v^4 + 4\frac{n-1}{T^2}\sigma_v^4 - \frac{3n}{T^2}\sigma_v^4 + O_p(\frac{1}{T\sqrt{n}}) + O_p(\frac{\sqrt{n}}{T^2});$

(d_2) $D_2 = O_p(\frac{1}{nT}) + O_p(\frac{1}{T^2}) + O_p(\frac{1}{T\sqrt{nT}});$

(d_3) $D_3 = O_p(\frac{1}{T^2});$

(d_4) $D_4 = O_p(\frac{1}{T^2});$

(d_5) $D_5 = O_p(\frac{1}{nT^2});$

(d_6) $D_6 = O_p(\frac{1}{nT^2});$

(d_7) $D_7 = O_p(\frac{1}{nT^2});$

(d_8) $D_8 = O_p(\frac{1}{nT});$

(d_9) $D_9 = O_p(\frac{1}{nT^2});$

(d_{10}) $D_{10} = O_p(\frac{1}{nT^2}).$

H.1 Proof of part d_1

$$
\begin{aligned}
D_1 =& \frac{1}{nT(T-1)(T-2)} \sum_{t\neq s\neq\tau}^{T} \sum_{s\neq\tau}^{T} \sum_{\tau=1}^{T} \tilde{v}_t'\tilde{v}_s\tilde{v}_s'\tilde{v}_\tau \\
=& \frac{1}{nT(T-1)(T-2)} \sum_{t\neq s\neq\tau}^{T} \sum_{s\neq\tau}^{T} \sum_{\tau=1}^{T} (v_t-\bar{v}_.)'(v_s-\bar{v}_.)(v_s-\bar{v}_.)'(v_\tau-\bar{v}_.) \\
=& \frac{1}{n}M_{4,T} + \sum_{k=1}^{7} D_1^k.
\end{aligned}
$$

Where $D_1^1 = -\frac{2}{nT(T-1)}\sum_{t\neq s}^T\sum_{s=1}^T v_t'v_sv_s'\bar{v}_.$; $D_1^2 = -\frac{2}{nT(T-1)(T-2)}\sum_{t\neq s\neq\tau}^T\sum_{s\neq\tau}^T\sum_{\tau=1}^T v_t'v_s\bar{v}_.'v_\tau$; $D_1^3 = \frac{2}{nT(T-1)}\sum_{t\neq s}^T\sum_{s=1}^T v_t'v_s\bar{v}_.'\bar{v}_.$; $D_1^4 = \frac{3}{nT(T-1)}\sum_{t\neq s}^T\sum_{s=1}^T v_t'\bar{v}_.v_s'\bar{v}_.$; $D_1^5 = \frac{1}{nT}\sum_{s=1}^T \bar{v}_.'v_sv_s'\bar{v}_.$; $D_1^6 = -\frac{3}{n}\bar{v}_.'\bar{v}_.\bar{v}_.'\bar{v}_.$.

Lemma 13 *Under Assumption 1, 2, 3 and the null,*

(1) $D_1^1 = -\frac{2\sigma_v^4}{T} + O_p(\frac{1}{T\sqrt{n}}) + O_p(\frac{\sqrt{n}}{T^2})$;

(2) $D_1^2 = O_p(\frac{\sqrt{n}}{T^2})$;

(3) $D_1^3 = O_p(\frac{\sqrt{n}}{T^2})$;

(4) $D_1^4 = 3\frac{n-1}{T^2}\sigma_v^4 + O_p(\frac{\sqrt{n}}{T^2})$;

(5) $D_1^5 = \frac{T-1}{T^2}\sigma_v^4 + \frac{n-1}{T^2}\sigma_v^4 + O_p(\frac{1}{T\sqrt{n}}) + O_p(\frac{\sqrt{n}}{T^2})$;

(6) $D_1^6 = -\frac{3n}{T^2}\sigma_v^4 + O_p(\frac{\sqrt{n}}{T^2})$.

Proof. It is easy to see that $D_1^1 = \frac{1}{2}C_1^1 = -\frac{2\sigma_v^4}{T} + O_p(\frac{1}{T\sqrt{n}}) + O_p(\sqrt{n}T^2)$;$D_1^3 = C_1^2 = O_p(\frac{\sqrt{n}}{T^2})$;$D_1^4 = \frac{3}{2}C_1^4 = \frac{3(n-1)}{T^2}\sigma_v^4 + O_p(\frac{\sqrt{n}}{T^2})$ and $D_1^6 = C_1^5 = -\frac{3n}{T^2}\sigma_v^4 + O_p(\frac{\sqrt{n}}{T^2})$. We have two part left to prove.Let's first consider part (2),

$$
\begin{aligned}
D_1^2 = & -\frac{2}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{\eta=1}^{T}v_t'v_sv_\eta'v_\tau \\
= & -\frac{2}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{\eta=1}^{T}\sum_{i=1}^{n}v_{it}v_{is}v_{i\tau}v_{i\eta} \\
& -\frac{2}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{\eta=1}^{T}\sum_{i=1}^{n}v_{it}v_{is}v_{j\tau}v_{j\eta}.
\end{aligned}
$$

There are four cases for each term, four terms for the first one are:

(1) $t=\eta\neq s\neq\tau: -\frac{2}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n}v_{it}^2v_{is}v_{i\tau}=O_p(\frac{1}{T^2\sqrt{n}})$;

(2) $t\neq s=\eta\neq\tau: -\frac{2}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n}v_{it}v_{is}^2v_{i\tau}=O_p(\frac{1}{T^2\sqrt{n}})$;

(3) $t\neq s\neq\tau=\eta: -\frac{2}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n}v_{it}v_{is}v_{i\tau}^2=O_p(\frac{1}{T^2\sqrt{n}})$;

(4) $t\neq s\neq\tau\neq\eta: -\frac{2}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau\neq\eta}^{T}\sum_{s\neq\tau\neq\eta}^{T}\sum_{\tau\neq\eta}^{T}\sum_{\eta=1}^{T}\sum_{i=1}^{n}v_{it}v_{is}v_{i\tau}v_{i\eta}=O_p(\frac{1}{T^2\sqrt{n}})$.

Similarly, we also have

(1) $t=\eta\neq s\neq\tau: -\frac{2}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}v_{it}v_{is}v_{jt}v_{j\tau}=O_p(\frac{1}{T^2\sqrt{T}})$;

(2) $t\neq s=\eta\neq\tau: -\frac{2}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}v_{it}v_{is}v_{js}v_{j\tau}=O_p(\frac{1}{T^2\sqrt{T}})$;

(3) $t\neq s\neq\tau=\eta: -\frac{2}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}v_{it}v_{is}v_{j\tau}^2=O_p(\frac{\sqrt{n}}{T^2})$;

(4) $t\neq s\neq\tau\neq\eta: -\frac{2}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau\neq\eta}^{T}\sum_{s\neq\tau\neq\eta}^{T}\sum_{\tau\neq\eta}^{T}\sum_{\eta=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}v_{it}v_{is}v_{j\tau}v_{j\eta}=O_p(\frac{1}{T^2})$.

The results above lead to $D_1^2 = O_p(\frac{\sqrt{n}}{T^2})$. Next we consider part (5),

$$\begin{aligned}
D_1^5 =& \frac{1}{nT}\sum_{s=1}^{T}(\frac{1}{T}\sum_{t=1}^{T} v_t)' v_s v_s' (\frac{1}{T}\sum_{\tau=1}^{T} v_\tau) = \frac{1}{nT^3}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T} v_t' v_s v_s' v_\tau \\
=& \frac{1}{nT^3}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T} v_t' v_s v_s' v_\tau + \frac{1}{nT^3}\sum_{s=1}^{T}\sum_{\tau=1}^{T} v_s' v_s v_s' v_\tau \\
=& -\frac{T-1}{2T} D_1^1 + \frac{1}{nT^3}\sum_{s=1}^{T}\sum_{i=1}^{n} v_{is}^4 + \frac{1}{nT^3}\sum_{s=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{is}^2 v_{js}^2 \\
&+ \frac{1}{nT^3}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n} v_{is}^3 v_{i\tau} + \frac{1}{nT^3}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{is}^2 v_{js} v_{j\tau} \\
=& \frac{T-1}{T^2}\sigma_v^4 + O_p(\frac{1}{T\sqrt{n}}) + O_p(\frac{\sqrt{n}}{T^2}) + O_p(\frac{1}{T^2}) + \frac{n-1}{T^2}\sigma_v^4 \\
&+ O_p(\frac{1}{T^2\sqrt{T}}) + O_p(\frac{1}{T^2\sqrt{n}}) + O_p(\frac{\sqrt{n}}{T^2}) \\
=& \frac{T-1}{T^2}\sigma_v^4 + \frac{n-1}{T^2}\sigma_v^4 + O_p(\frac{1}{T\sqrt{n}}) + O_p(\frac{\sqrt{n}}{T^2}).
\end{aligned}$$

■

With Lemma 13, we have

$$D_1 = \frac{1}{n}M_{4,T} + \frac{T-1}{T^2}\sigma_v^4 + 4\frac{n-1}{T^2}\sigma_v^4 - \frac{3n}{T^2}\sigma_v^4 + O_p(\frac{1}{T\sqrt{n}}) + O_p(\frac{\sqrt{n}}{T^2}).$$

H.2 Proof of part d_2

$$\begin{aligned}
D_2 =& -\frac{2}{nT(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}(v_t-\bar{v}.)'(v_s-\bar{v}.)(v_s-\bar{v}.)'\tilde{x}_\tau(\tilde{\beta}-\beta) \\
=& 2\sum_{k=1}^{7} D_2^k.
\end{aligned}$$

Where

$D_2^1 = -\frac{1}{nT(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T} v_t' v_s v_s' \tilde{x}_\tau(\tilde{\beta}-\beta)$;

$D_2^2 = \frac{1}{nT(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T} v_t'\bar{v}.v_s'\tilde{x}_\tau(\tilde{\beta}-\beta)$;

$D_2^3 = \frac{1}{nT(T-1)} \sum_{s\neq\tau}^T \sum_{\tau=1}^T \bar{v}_\cdot' v_s v_s' \tilde{x}_\tau(\tilde{\beta} - \beta)$;
$D_2^4 = -\frac{1}{nT(T-1)} \sum_{s\neq\tau}^T \sum_{\tau=1}^T \bar{v}_\cdot' \bar{v}_\cdot v_s' \tilde{x}_\tau(\tilde{\beta} - \beta)$;
$D_2^5 = \frac{1}{nT(T-1)(T-2)} \sum_{t\neq s\neq\tau}^T \sum_{s\neq\tau}^T \sum_{\tau=1}^T v_t' v_s \bar{v}_\cdot' \tilde{x}_\tau(\tilde{\beta} - \beta)$;
$D_2^6 = -\frac{2}{nT(T-1)} \sum_{t\neq\tau}^T \sum_{\tau=1}^T v_t' \bar{v}_\cdot \bar{v}_\cdot' \tilde{x}_\tau(\tilde{\beta} - \beta)$;
$D_2^7 = \frac{1}{nT} \sum_{\tau=1}^T \bar{v}_\cdot' \bar{v}_\cdot \bar{v}_\cdot' \tilde{x}_\tau(\tilde{\beta} - \beta)$.

Lemma 14 *Under Assumption 1, 2, 3 and the null,*

(1) $D_2^1 = O_p(\frac{1}{nT}) + O_p(\frac{1}{T\sqrt{nT}})$;

(2) $D_2^2 = O_p(\frac{1}{T^2})$;

(3) $D_2^3 = O_p(\frac{1}{nT}) + O_p(\frac{1}{T^2}) + O_p(\frac{1}{T\sqrt{nT}})$;

(4) $D_2^4 = O_p(\frac{1}{T^2})$;

(5) $D_2^5 = O_p(\frac{1}{T^2\sqrt{n}})$;

(6) $D_2^6 = O_p(\frac{1}{T^2})$;

(7) $D_2^7 = O_p(\frac{1}{T^2})$.

Proof. Proof of part (1)

$$\begin{aligned}
D_2^1 = &- \frac{1}{nT(T-1)(T-2)} \sum_{t\neq s\neq\tau}^T \sum_{s\neq\tau}^T \sum_{\tau=1}^T \sum_{i=1}^n v_{it} v_{is}^2 \tilde{x}_{i\tau}(\tilde{\beta} - \beta) \\
&- \frac{1}{nT(T-1)(T-2)} \sum_{t\neq s\neq\tau}^T \sum_{s\neq\tau}^T \sum_{\tau=1}^T \sum_{i\neq j}^n \sum_{j=1}^n v_{it} v_{is} v_{js} \tilde{x}_{j\tau}(\tilde{\beta} - \beta) \\
=&O_p(\frac{1}{nT}) + O_p(\frac{1}{T\sqrt{nT}}).
\end{aligned}$$

Proof of part (2)

$$
\begin{aligned}
D_2^2 =& \frac{1}{nT(T-1)(T-2)} \sum_{t\neq s\neq\tau}^{T} \sum_{s\neq\tau}^{T} \sum_{\tau=1}^{T} v_t'(\frac{1}{T}\sum_{\eta=1}^{T} v_\eta) v_s' \tilde{x}_\tau (\tilde{\beta}-\beta) \\
=& \frac{1}{nT^2(T-1)(T-2)} \sum_{t\neq s\neq\tau}^{T} \sum_{s\neq\tau}^{T} \sum_{\tau=1}^{T} \sum_{\eta=1}^{T} v_t' v_\eta v_s' \tilde{x}_\tau (\tilde{\beta}-\beta) \\
=& \frac{1}{nT^2(T-1)(T-2)} \sum_{t\neq s\neq\tau}^{T} \sum_{s\neq\tau}^{T} \sum_{\tau=1}^{T} \sum_{\eta=1}^{T} \sum_{i=1}^{n} v_{it} v_{i\eta} v_{is} \tilde{x}_{i\tau} (\tilde{\beta}-\beta) \\
&+ \frac{1}{nT^2(T-1)(T-2)} \sum_{t\neq s\neq\tau}^{T} \sum_{s\neq\tau}^{T} \sum_{\tau=1}^{T} \sum_{\eta=1}^{T} \sum_{i\neq j}^{n} \sum_{j=1}^{n} v_{jt} v_{j\eta} v_{js} \tilde{x}_{j\tau} (\tilde{\beta}-\beta).
\end{aligned}
$$

We distinct four cases of the first term: (1) $(t=\eta)\neq s\neq\tau$; (2) $t\neq(s=\eta)\neq\tau$; (3) $t\neq s\neq(\tau=\eta)$; (4) $t\neq s\neq\tau\neq\eta$. Then the first term can be calculated respectively as the following:

(1) $(t=\eta)\neq s\neq\tau$: $\frac{1}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{\eta=1}^{T}\sum_{i=1}^{n} v_{it}^2 v_{is}\tilde{x}_{i\tau}(\tilde{\beta}-\beta) = O_p(\frac{1}{nT^2})$;

(2) $t\neq(s=\eta)\neq\tau$: $\frac{1}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{\eta=1}^{T}\sum_{i=1}^{n} v_{it} v_{is}^2\tilde{x}_{i\tau}(\tilde{\beta}-\beta) = O_p(\frac{1}{nT^2})$;

(3) $t\neq s\neq(\tau=\eta)$: $\frac{1}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{\eta=1}^{T}\sum_{i=1}^{n} v_{it} v_{i\tau} v_{is}\tilde{x}_{i\tau}(\tilde{\beta}-\beta) = O_p(\frac{1}{nT^3})$;

(4) $t\neq s\neq\tau\neq\eta$: $\frac{1}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{\eta=1}^{T}\sum_{i=1}^{n} v_{it} v_{i\eta} v_{is}\tilde{x}_{i\tau}(\tilde{\beta}-\beta) = O_p(\frac{1}{nT^2})$.

We can get four cases of the second term as well:

(1) $(t=\eta)\neq s\neq\tau$: $\frac{1}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{\eta=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it}^2 v_{js}\tilde{x}_{j\tau}(\tilde{\beta}-\beta) = O_p(\frac{1}{T^2})$;

(2) $t\neq(s=\eta)\neq\tau$: $\frac{1}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{\eta=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it} v_{is} v_{js}\tilde{x}_{j\tau}(\tilde{\beta}-\beta) = O_p(\frac{1}{T^2\sqrt{nT}})$;

(3) $t\neq s\neq(\tau=\eta)$: $\frac{1}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{\eta=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it} v_{i\tau} v_{js}\tilde{x}_{j\tau}(\tilde{\beta}-\beta) = O_p(\frac{1}{T^3\sqrt{n}})$;

(4) $t\neq s\neq\tau\neq\eta$: $\frac{1}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{\eta=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it} v_{i\eta} v_{js}\tilde{x}_{j\tau}(\tilde{\beta}-\beta) = O_p(\frac{1}{T^2\sqrt{n}})$.

Then we have

$$D_2^2 = O_p(\frac{1}{T^2}).$$

Proof of part (3)

$$\begin{aligned}
D_2^3 =& \frac{1}{nT(T-1)} \sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}(\frac{1}{T}\sum_{t=1}^{T} v_t)' v_s v_s' \tilde{x}_\tau(\tilde{\beta}-\beta) \\
=& \frac{1}{nT^2(T-1)} \sum_{t=1}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T} v_t' v_s v_s' \tilde{x}_\tau(\tilde{\beta}-\beta) \\
=& \frac{1}{nT^2(T-1)} \sum_{t=1}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n} v_{it} v_{is}^2 \tilde{x}_{i\tau}(\tilde{\beta}-\beta) \\
&+ \frac{1}{nT^2(T-1)} \sum_{t=1}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it} v_{is} v_{js} \tilde{x}_{j\tau}(\tilde{\beta}-\beta).
\end{aligned}$$

There are three cases needed to be distinguished: (1) $t = s \neq \tau$; (2) $t \neq s = \tau$; (3) $t \neq s \neq \tau$.

Then the first term can be expanded as the following:

(1) $t = s \neq \tau : \frac{1}{nT^2(T-1)} \sum_{t=1}^{T}\sum_{t\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n} v_{it}^3 \tilde{x}_{i\tau}(\tilde{\beta}-\beta) = O_p(\frac{1}{nT^2})$;

(2) $s \neq t = \tau : \frac{1}{nT^2(T-1)} \sum_{t=1}^{T}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n} v_{it} v_{is}^2 \tilde{x}_{it}(\tilde{\beta}-\beta) = O_p(\frac{1}{nT^2})$;

(3) $t \neq s \neq \tau : \frac{1}{nT^2(T-1)} \sum_{t=1}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n} v_{it} v_{is}^2 \tilde{x}_{i\tau}(\tilde{\beta}-\beta) = O_p(\frac{1}{nT})$.

Similarly, the second term can be also expanded as the following:

(1) $t = s \neq \tau : \frac{1}{nT^2(T-1)} \sum_{t=1}^{T}\sum_{t\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it}^2 v_{jt} \tilde{x}_{j\tau}(\tilde{\beta}-\beta) = O_p(\frac{1}{T^2})$;

(2) $s \neq t = \tau : \frac{1}{nT^2(T-1)} \sum_{t=1}^{T}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it} v_{is} v_{js} \tilde{x}_{jt}(\tilde{\beta}-\beta) = O_p(\frac{1}{T^2\sqrt{nT}})$;

(3) $t \neq s \neq \tau : \frac{1}{nT^2(T-1)} \sum_{t=1}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it} v_{is} v_{js} \tilde{x}_{j\tau}(\tilde{\beta}-\beta) = O_p(\frac{1}{T\sqrt{nT}})$.

Hence, we have

$$D_2^3 = O_p(\frac{1}{nT}) + O_p(\frac{1}{T^2}) + O_p(\frac{1}{T\sqrt{nT}}).$$

Proof of part (5)

$$\begin{aligned}
D_2^5 =& \frac{1}{nT(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T} v_t' v_s(\frac{1}{T}\sum_{\eta=1}^{T} v_\eta)'\tilde{x}_\tau(\tilde{\beta}-\beta)\\
=& \frac{1}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{\eta=1}^{T} v_t' v_s v_\eta'\tilde{x}_\tau(\tilde{\beta}-\beta)\\
=& \frac{1}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{\eta=1}^{T}\sum_{i=1}^{n} v_{it}v_{is}v_{i\eta}\tilde{x}_{i\tau}(\tilde{\beta}-\beta)\\
&+\frac{1}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{\eta=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it}v_{is}v_{j\eta}\tilde{x}_{j\tau}(\tilde{\beta}-\beta).
\end{aligned}$$

Here, we need to distinguish 4 cases for expanding terms: (1) $t=\eta\neq s\neq\tau$; (2) $t\neq(s=\eta)\neq\tau$; (3) $t\neq s\neq(\tau=\eta)$; (4) $t\neq s\neq\tau\neq\eta$. So we have four terms for the first term of B_2^5:

(1) $t=\eta\neq s\neq\tau:\frac{1}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n} v_{it}^2 v_{is}\tilde{x}_{i\tau}(\tilde{\beta}-\beta)=O_p(\frac{1}{nT^2})$;

(2) $t\neq(s=\eta)\neq\tau:\frac{1}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n} v_{it}v_{is}^2\tilde{x}_{i\tau}(\tilde{\beta}-\beta)=O_p(\frac{1}{nT^2})$;

(3) $t\neq s\neq(\tau=\eta):\frac{1}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n} v_{it}v_{is}v_{i\tau}\tilde{x}_{i\tau}(\tilde{\beta}-\beta)=O_p(\frac{1}{nT^3})$;

(4) $t\neq s\neq\tau\neq\eta:\frac{1}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n} v_{it}v_{is}v_{i\eta}\tilde{x}_{i\tau}(\tilde{\beta}-\beta)=O_p(\frac{1}{nT^2})$.

And we also have 4 cases for the second term:

(1) $t=\eta\neq s\neq\tau:\frac{1}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it}v_{is}v_{jt}\tilde{x}_{j\tau}(\tilde{\beta}-\beta)=O_p(\frac{1}{T^2\sqrt{nT}})$;

(2) $t\neq(s=\eta)\neq\tau:\frac{1}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it}v_{is}v_{js}\tilde{x}_{j\tau}(\tilde{\beta}-\beta)=O_p(\frac{1}{T^2\sqrt{nT}})$;

(3) $t\neq s\neq(\tau=\eta):\frac{1}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it}v_{is}v_{j\tau}\tilde{x}_{j\tau}(\tilde{\beta}-\beta)=O_p(\frac{1}{T^3\sqrt{n}})$;

(4) $t\neq s\neq\tau\neq\eta:\frac{1}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it}v_{is}v_{j\eta}\tilde{x}_{j\tau}(\tilde{\beta}-\beta)=O_p(\frac{1}{T^2\sqrt{n}})$.

Hence

$$D_2^5=O_p(\frac{1}{T^2\sqrt{n}}).$$

For the left parts, we could easily to show $D_2^4 = \frac{1}{4}C_2^7 = O_p(\frac{1}{T})$; $D_2^6 = \frac{1}{4}C_2^6 = O_p(\frac{1}{T^2})$ and $D_2^7 = \frac{1}{4}C_2^8 = O_p(\frac{1}{T^2})$; then Lemma 14 is proved. ■

With Lemma 14, we have

$$D_2 = O_p(\frac{1}{nT}) + O_p(\frac{1}{T\sqrt{nT}}) + O_p(\frac{1}{T^2}).$$

H.3 Proof of part d_3

$$\begin{aligned}D_3 =& -\frac{2}{nT(T-1)(T-2)}(\tilde{\beta}-\beta)' \sum_{t\neq s\neq \tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}(v_t-\bar{v}.)'(v_s-\bar{v}.)\tilde{x}_s'(v_\tau-\bar{v}.)\\ =&2\sum_{k=1}^{8}D_3^k.\end{aligned}$$

Where

$D_3^1 = -\frac{1}{nT(T-1)(T-2)}(\tilde{\beta}-\beta)'\sum_{t\neq s\neq\tau}^T\sum_{s\neq\tau}^T\sum_{\tau=1}^T v_t'v_s\tilde{x}_s'v_\tau$;

$D_3^2 = \frac{1}{nT(T-1)(T-2)}(\tilde{\beta}-\beta)'\sum_{t\neq s\neq\tau}^T\sum_{s\neq\tau}^T\sum_{\tau=1}^T v_t'\bar{v}.\tilde{x}_s'v_\tau$;

$D_3^3 = \frac{1}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{s\neq\tau}^T\sum_{\tau=1}^T \bar{v}_.'v_s\tilde{x}_s'v_\tau$;

$D_3^4 = -\frac{1}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{s\neq\tau}^T\sum_{\tau=1}^T \bar{v}_.'\bar{v}.\tilde{x}_s'v_\tau$;

$D_3^5 = \frac{1}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^T\sum_{s=1}^T v_t'v_s\tilde{x}_s'\bar{v}.$;

$D_3^6 = -\frac{1}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq s}^T\sum_{s=1}^T v_t'\bar{v}.\tilde{x}_s'\bar{v}.$;

$D_3^7 = -\frac{1}{nT}(\tilde{\beta}-\beta)'\sum_{s=1}^T \bar{v}_.'v_s\tilde{x}_s'\bar{v}.$;

$D_3^8 = \frac{1}{nT}(\tilde{\beta}-\beta)'\sum_{s=1}^T \bar{v}_.'\bar{v}.\tilde{x}_s'\bar{v}..$

Lemma 15 *Under Assumption 1, 2, 3 and the null,*

(1) $D_3^1 = O_p(\frac{1}{T^2\sqrt{n}})$;

(2) $D_3^2 = O_p(\frac{1}{T^2})$;

(3) $D_3^3 = O_p(\frac{1}{T^2})$;

(4) $D_3^4 = O_p(\frac{1}{T^2})$;

(5) $D_3^5 = O_p(\frac{1}{T^2})$;

(6) $D_3^6 = O_p(\frac{1}{T^2})$;

(7) $D_3^7 = O_p(\frac{1}{T^2})$;

(8) $D_3^8 = O_p(\frac{1}{T^2})$.

Proof. First consider part (1),

$$\begin{aligned}
D_3^1 =& -\frac{1}{nT(T-1)(T-2)}(\tilde{\beta}-\beta)'\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T} v_t' v_s \tilde{x}_s' v_\tau \\
=& -\frac{1}{nT(T-1)(T-2)}(\tilde{\beta}-\beta)'\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n} v_{it} v_{is} \tilde{x}_{is} v_{i\tau} \\
& -\frac{1}{nT(T-1)(T-2)}(\tilde{\beta}-\beta)'\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it} v_{is} \tilde{x}_{js} v_{j\tau} \\
=& O_p(\frac{1}{nT^2}) + O_p(\frac{1}{T^2\sqrt{n}}) = O_p(\frac{1}{T^2\sqrt{n}}).
\end{aligned}$$

Then consider part (2),

$$\begin{aligned}
D_3^2 =& \frac{1}{nT(T-1)(T-2)}(\tilde{\beta}-\beta)'\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T} v_t'(\frac{1}{T}\sum_{\eta=1}^{T} v_\eta)\tilde{x}_s' v_\tau \\
=& \frac{1}{nT^2(T-1)(T-2)}(\tilde{\beta}-\beta)'\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{\eta=1}^{T} v_t' v_\eta \tilde{x}_s' v_\tau \\
=& \frac{1}{nT^2(T-1)(T-2)}(\tilde{\beta}-\beta)'\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{\eta=1}^{T}\sum_{i=1}^{n} v_{it} v_{i\eta} \tilde{x}_{is} v_{i\tau} \\
& + \frac{1}{nT^2(T-1)(T-2)}(\tilde{\beta}-\beta)'\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{\eta=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it} v_{i\eta} \tilde{x}_{js} v_{j\tau}.
\end{aligned}$$

We need to distinguish four cases for each term, the first term can be expanded as the following:

(1) $(t=\eta)\neq s\neq\tau$: $\frac{1}{nT^2(T-1)(T-2)}(\tilde{\beta}-\beta)'\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n}v_{it}^2\tilde{x}_{is}v_{i\tau}=O_p(\frac{1}{nT^2})$;

(2) $t\neq(s=\eta)\neq\tau$: $\frac{1}{nT^2(T-1)(T-2)}(\tilde{\beta}-\beta)'\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n}v_{it}v_{is}\tilde{x}_{is}v_{i\tau}=O_p(\frac{1}{nT^3})$;

(3) $t\neq s\neq(\tau=\eta)$: $\frac{1}{nT^2(T-1)(T-2)}(\tilde{\beta}-\beta)'\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n}v_{it}\tilde{x}_{is}v_{i\tau}^2=O_p(\frac{1}{nT^2})$;

(4) $t\neq s\neq\tau\neq\eta$: $\frac{1}{nT^2(T-1)(T-2)}(\tilde{\beta}-\beta)'\sum_{t\neq s\neq\tau\neq\eta}^{T}\sum_{s\neq\tau\neq\eta}^{T}\sum_{\tau\neq\eta}^{T}\sum_{\eta=1}^{T}\sum_{i=1}^{n}v_{it}v_{i\eta}\tilde{x}_{is}v_{i\tau}=O_p(\frac{1}{nT^2})$.

Similarly, the second term can be also expand as the following four terms:

(1) $(t=\eta)\neq s\neq\tau$: $\frac{1}{nT^2(T-1)(T-2)}(\tilde{\beta}-\beta)'\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{\eta=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}v_{it}^2\tilde{x}_{js}v_{j\tau}=O_p(\frac{1}{T^2})$;

(2) $t\neq(s=\eta)\neq\tau$: $\frac{1}{nT^2(T-1)(T-2)}(\tilde{\beta}-\beta)'\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{\eta=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}v_{it}v_{is}\tilde{x}_{js}v_{j\tau}=O_p(\frac{1}{T^3\sqrt{n}})$;

(3) $t\neq s\neq(\tau=\eta)$: $\frac{1}{nT^2(T-1)(T-2)}(\tilde{\beta}-\beta)'\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{\eta=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}v_{it}v_{i\tau}\tilde{x}_{js}v_{j\tau}=O_p(\frac{1}{T^2\sqrt{nT}})$;

(4) $t\neq s\neq\tau\neq\eta$: $\frac{1}{nT^2(T-1)(T-2)}(\tilde{\beta}-\beta)'\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{\eta=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}v_{it}v_{i\eta}\tilde{x}_{js}v_{j\tau}=O_p(\frac{1}{T^2\sqrt{n}})$.

With the results above we have:

$$D_3^2=O_p(\frac{1}{T^2}).$$

For the rest parts, we could show that $D_3^3=\frac{1}{4}C_2^3=O_p(\frac{1}{T^2})$; $D_3^4=\frac{1}{4}C_2^7=O_p(\frac{1}{T^2})$; $D_3^5=\frac{1}{4}C_2^2=O_p(\frac{1}{T^2})$; $D_3^6=\frac{1}{4}C_2^6=O_p(\frac{1}{T^2})$;$D_3^7=\frac{1}{4}C_2^4=O_p(\frac{1}{T^2})$ and $D_3^8=\frac{1}{4}C_2^8=O_p(\frac{1}{T^2})$. Lemma 15 then is proved. ■

From the results above, then we have

$$D_3=O_p(\frac{1}{T^2}).$$

H.4 Proof of part d_4

$$
\begin{aligned}
D_4 =& \frac{2}{nT(T-1)(T-2)} \sum_{t\neq s\neq \tau}^{T} \sum_{s\neq\tau}^{T} \sum_{\tau=1}^{T} (v_t' - \bar{v}.)'(v_s - \bar{v}.)(\tilde{\beta}-\beta)'\tilde{x}_s'\tilde{x}_\tau(\tilde{\beta}-\beta) \\
=& \frac{2}{nT(T-1)(T-2)} (\tilde{\beta}-\beta)' \sum_{t\neq s\neq \tau}^{T} \sum_{s\neq\tau}^{T} \sum_{\tau=1}^{T} (v_t' - \bar{v}.)'(v_s - \bar{v}.)\tilde{x}_s'\tilde{x}_\tau(\tilde{\beta}-\beta) \\
=& 2\sum_{k=1}^{4} D_4^k.
\end{aligned}
$$

Where

$D_4^1 = \frac{1}{nT(T-1)(T-2)}(\tilde{\beta}-\beta)' \sum_{t\neq s\neq\tau}^T \sum_{s\neq\tau}^T \sum_{\tau=1}^T v_t' v_s \tilde{x}_s' \tilde{x}_\tau (\tilde{\beta}-\beta)$;

$D_4^2 = -\frac{1}{nT(T-1)(T-2)}(\tilde{\beta}-\beta)' \sum_{t\neq s\neq\tau}^T \sum_{s\neq\tau}^T \sum_{\tau=1}^T v_t' \bar{v}. \tilde{x}_s' \tilde{x}_\tau (\tilde{\beta}-\beta)$;

$D_4^3 = -\frac{1}{nT(T-1)}(\tilde{\beta}-\beta)' \sum_{s\neq\tau}^T \sum_{\tau=1}^T \bar{v}_.' v_s \tilde{x}_s' \tilde{x}_\tau (\tilde{\beta}-\beta)$;

$D_4^4 = \frac{1}{nT(T-1)}(\tilde{\beta}-\beta)' \sum_{s\neq\tau}^T \sum_{\tau=1}^T \bar{v}_.' \bar{v}. \tilde{x}_s' \tilde{x}_\tau (\tilde{\beta}-\beta)$.

Lemma 16 *Under Assumption 1, 2, 3 and the null,*

(1) $D_4^1 = O_p(\frac{1}{T^2\sqrt{n}})$;

(2) $D_4^2 = O_p(\frac{1}{T^2})$;

(3) $D_4^3 = O_p(\frac{1}{T^2})$;

(4) $D_4^4 = O_p(\frac{1}{T^2})$.

Proof. Proof of part (1)

$$
\begin{aligned}
D_4^1 =& \frac{1}{nT(T-1)(T-2)}(\tilde{\beta}-\beta)' \sum_{t\neq s\neq \tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T} v_t' v_s \tilde{x}_s' \tilde{x}_\tau(\tilde{\beta}-\beta) \\
=& \frac{1}{nT(T-1)(T-2)}(\tilde{\beta}-\beta)' \sum_{t\neq s\neq \tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n} v_{it} v_{is} \tilde{x}_{is} \tilde{x}_{i\tau}(\tilde{\beta}-\beta) \\
&+ \frac{1}{nT(T-1)(T-2)}(\tilde{\beta}-\beta)' \sum_{t\neq s\neq \tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it} v_{is} \tilde{x}_{js} \tilde{x}_{j\tau}(\tilde{\beta}-\beta) \\
=& O_p(\frac{1}{nT^2\sqrt{n}}) + O_p(\frac{1}{T^2\sqrt{n}}) = O_p(\frac{1}{T^2\sqrt{n}}).
\end{aligned}
$$

Proof of part (2)

$$
\begin{aligned}
D_4^2 =& -\frac{1}{nT(T-1)(T-2)}(\tilde{\beta}-\beta)' \sum_{t\neq s\neq \tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T} v_t' (\frac{1}{T}\sum_{\eta=1}^{T} v_\eta) \tilde{x}_s' \tilde{x}_\tau(\tilde{\beta}-\beta) \\
=& -\frac{1}{nT^2(T-1)(T-2)}(\tilde{\beta}-\beta)' \sum_{t\neq s\neq \tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{\eta=1}^{T} v_t' v_\eta \tilde{x}_s' \tilde{x}_\tau(\tilde{\beta}-\beta) \\
=& -\frac{1}{nT^2(T-1)(T-2)}(\tilde{\beta}-\beta)' \sum_{t\neq s\neq \tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{\eta=1}^{T}\sum_{i=1}^{n} v_{it} v_{i\eta} \tilde{x}_{is} \tilde{x}_{i\tau}(\tilde{\beta}-\beta) \\
&- \frac{1}{nT^2(T-1)(T-2)}(\tilde{\beta}-\beta)' \sum_{t\neq s\neq \tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{\eta=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it} v_{i\eta} \tilde{x}_{js} \tilde{x}_{j\tau}(\tilde{\beta}-\beta).
\end{aligned}
$$

The term $\frac{1}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{\eta=1}^{T}\sum_{i=1}^{n} v_{it}v_{i\eta}\tilde{x}_{is}\tilde{x}_{i\tau}$ can be expanded as the following:

(1) $t=\eta\neq s\neq\tau : -\frac{1}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n} v_{it}^2\tilde{x}_{is}\tilde{x}_{i\tau} = O_p(\frac{1}{T})$;

(2) $t\neq(s=\eta)\neq\tau : -\frac{1}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n} v_{it}v_{is}\tilde{x}_{is}\tilde{x}_{i\tau} = O_p(\frac{1}{T^2\sqrt{n}})$;

(3) $t\neq s\neq(\eta=\tau) = -\frac{1}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n} v_{it}v_{i\tau}\tilde{x}_{is}\tilde{x}_{i\tau} = O_p(\frac{1}{T^2\sqrt{n}})$;

(4) $t\neq s\neq\tau\neq\eta : -\frac{1}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau\neq\eta}^{T}\sum_{s\neq\tau\neq\eta}^{T}\sum_{\tau\neq\eta}^{T}\sum_{\eta=1}\sum_{i=1}^{n} v_{it}v_{i\eta}\tilde{x}_{is}\tilde{x}_{i\tau} = O_p(\frac{1}{T\sqrt{n}})$.

And the term $\frac{1}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{\eta=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it}v_{i\eta}\tilde{x}_{js}\tilde{x}_{j\tau}$ can be also expanded as the following four terms:

(1) $t=\eta\neq s\neq\tau : -\frac{1}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it}^2\tilde{x}_{js}\tilde{x}_{j\tau} = O_p(\frac{n}{T})$;

(2) $t \neq (s=\eta) \neq \tau: -\frac{1}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq \tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it}v_{is}\tilde{x}_{js}\tilde{x}_{j\tau} = O_p(\frac{\sqrt{n}}{T^2})$;

(3) $t \neq s \neq (\eta=\tau) = -\frac{1}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq \tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it}v_{i\tau}\tilde{x}_{is}\tilde{x}_{i\tau} = O_p(\frac{\sqrt{n}}{T^2})$;

(4) $t \neq s \neq \tau \neq \eta: -\frac{1}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq \tau\neq\eta}^{T}\sum_{s\neq\tau\neq\eta}^{T}\sum_{\tau\neq\eta}^{T}\sum_{\eta=1}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it}v_{i\eta}\tilde{x}_{is}\tilde{x}_{i\tau} = O_p(\frac{\sqrt{n}}{T})$.

Since $\tilde{\beta}-\beta = O_p(\frac{1}{\sqrt{nT}})$, and with the results above, we have

$$D_4^2 = O_p(\frac{1}{T^2}).$$

Proof of part(3)

$$\begin{aligned} D_4^3 =& -\frac{1}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t=1}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}(\frac{1}{T}\sum_{t=1}^{T}v_t)'v_s\tilde{x}_s'\tilde{x}_\tau(\tilde{\beta}-\beta) \\ =& -\frac{1}{nT^2(T-1)}(\tilde{\beta}-\beta)'\sum_{t=1}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}v_t'v_s\tilde{x}_s'\tilde{x}_\tau(\tilde{\beta}-\beta) \\ =& -\frac{1}{nT^2(T-1)}(\tilde{\beta}-\beta)'\sum_{t=1}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n}v_{it}v_{is}\tilde{x}_{is}\tilde{x}_{i\tau}(\tilde{\beta}-\beta) \\ & -\frac{1}{nT^2(T-1)}(\tilde{\beta}-\beta)'\sum_{t=1}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}v_{it}v_{is}\tilde{x}_{is}\tilde{x}_{i\tau}(\tilde{\beta}-\beta). \end{aligned}$$

The term $-\frac{1}{nT^2(T-1)}\sum_{t=1}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n} v_{it}v_{is}\tilde{x}_{is}\tilde{x}_{i\tau}$ can be calculated by the following expressions:

(1) $t = s \neq \tau: -\frac{1}{nT^2(T-1)}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n} v_{is}^2\tilde{x}_{is}\tilde{x}_{i\tau} = O_p(\frac{1}{T})$;

(2) $t \neq (s=\tau): -\frac{1}{nT^2(T-1)}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n} v_{i\tau}v_{is}\tilde{x}_{is}\tilde{x}_{i\tau} = O_p(\frac{1}{T^2\sqrt{n}})$;

(3) $t \neq s \neq \tau: -\frac{1}{nT^2(T-1)}\sum_{t\neq s\neq\tau}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n} v_{it}v_{is}\tilde{x}_{is}\tilde{x}_{i\tau} = O_p(\frac{1}{T\sqrt{n}})$.

Similarly, the term $-\frac{1}{nT^2(T-1)}\sum_{t=1}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it}v_{is}\tilde{x}_{js}\tilde{x}_{j\tau}$ can be also expanded as the following expressions:

(1) $t = s \neq \tau: -\frac{1}{nT^2(T-1)}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{is}^2\tilde{x}_{js}\tilde{x}_{j\tau} = O_p(\frac{n}{T})$;

(2) $t \neq (s=\tau): -\frac{1}{nT^2(T-1)}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{i\tau}v_{is}\tilde{x}_{js}\tilde{x}_{j\tau} = O_p(\frac{\sqrt{n}}{T^2})$;

(3) $t \neq s \neq \tau: -\frac{1}{nT^2(T-1)}\sum_{t\neq s\neq\tau}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n} v_{it}v_{is}\tilde{x}_{js}\tilde{x}_{j\tau} = O_p(\frac{\sqrt{n}}{T})$.

Since $\tilde{\beta} - \beta = O_p(\frac{1}{\sqrt{nT}})$, and with the results above, we have

$$D_4^3 = O_p(\frac{1}{T^2}).$$

Proof of part (4)

$$\begin{aligned}
D_4^4 =& \frac{1}{nT(T-1)}(\tilde{\beta} - \beta)' \sum_{s \neq \tau}^{T} \sum_{\tau=1}^{T} (\frac{1}{T} \sum_{t=1}^{T} v_t)' (\frac{1}{T} \sum_{\eta=1}^{T} v_\eta) \tilde{x}_s' \tilde{x}_\tau (\tilde{\beta} - \beta) \\
=& \frac{1}{nT^3(T-1)}(\tilde{\beta} - \beta)' \sum_{t=1}^{T} \sum_{s \neq \tau}^{T} \sum_{\tau=1}^{T} \sum_{\eta=1}^{T} v_t' v_\eta \tilde{x}_s' \tilde{x}_\tau (\tilde{\beta} - \beta) \\
=& \frac{1}{2} C_3^3 = O_p(\frac{1}{T^2}).
\end{aligned}$$

■

As a result, we have

$$D_4 = O_p(\frac{1}{T^2}).$$

H.5 Proof of part d_5

$$\begin{aligned}
D_5 =& \frac{2}{nT(T-1)(T-2)} \sum_{t \neq s \neq \tau}^{T} \sum_{s \neq \tau}^{T} \sum_{\tau=1}^{T} (v_t - \bar{v}.)' \tilde{x}_s (\tilde{\beta} - \beta)(v_s - \bar{v}.)' \tilde{x}_\tau (\tilde{\beta} - \beta) \\
=& \frac{2}{nT(T-1)(T-2)}(\tilde{\beta} - \beta)' \sum_{t \neq s \neq \tau}^{T} \sum_{s \neq \tau}^{T} \sum_{\tau=1}^{T} \tilde{x}_s' (v_t - \bar{v}.)(v_s - \bar{v}.)' \tilde{x}_\tau (\tilde{\beta} - \beta) \\
=& 2 \sum_{k=1}^{4} D_5^k.
\end{aligned}$$

Where

$D_5^1 = \frac{1}{nT(T-1)(T-2)}(\tilde{\beta} - \beta)' \sum_{t \neq s \neq \tau}^{T} \sum_{s \neq \tau}^{T} \sum_{\tau=1}^{T} \tilde{x}_s' v_t v_s' \tilde{x}_\tau (\tilde{\beta} - \beta)$;

$D_5^2 = -\frac{1}{nT(T-1)(T-2)}(\tilde{\beta} - \beta)' \sum_{t \neq s \neq \tau}^{T} \sum_{s \neq \tau}^{T} \sum_{\tau=1}^{T} \tilde{x}_s' v_t \bar{v}_.' \tilde{x}_\tau (\tilde{\beta} - \beta)$;

$D_5^3 = -\frac{1}{nT(T-1)}(\tilde{\beta} - \beta)' \sum_{s \neq \tau}^{T} \sum_{\tau=1}^{T} \tilde{x}_s' \bar{v}. v_s' \tilde{x}_\tau (\tilde{\beta} - \beta)$;

$D_5^4 = \frac{1}{nT(T-1)}(\tilde{\beta} - \beta)' \sum_{s \neq \tau}^{T} \sum_{\tau=1}^{T} \tilde{x}_s' \bar{v}. \bar{v}_.' \tilde{x}_\tau (\tilde{\beta} - \beta)$.

Lemma 17 *Under Assumption 1, 2, 3 and the null,*

(1) $D_5^1 = O_p(\frac{1}{nT^2})$;

(2) $D_5^2 = O_p(\frac{1}{nT^2})$;

(3) $D_5^3 = O_p(\frac{1}{nT^2})$;

(4) $D_5^4 = O_p(\frac{1}{nT^2})$.

Proof. Proof of part (1)

$$
\begin{aligned}
D_5^1 =& \frac{1}{nT(T-1)(T-2)}(\tilde{\beta}-\beta)' \sum_{t\neq s\neq \tau}^{T} \sum_{s\neq\tau}^{T} \sum_{\tau=1}^{T} \sum_{i=1}^{n} \tilde{x}_{is} v_{it} v_{is} \tilde{x}_{i\tau} (\tilde{\beta}-\beta) \\
&+ \frac{1}{nT(T-1)(T-2)}(\tilde{\beta}-\beta)' \sum_{t\neq s\neq \tau}^{T} \sum_{s\neq\tau}^{T} \sum_{\tau=1}^{T} \sum_{i\neq j}^{n} \sum_{j=1} \tilde{x}_{is} v_{it} v_{js} \tilde{x}_{j\tau} (\tilde{\beta}-\beta) \\
=& O_p(\frac{1}{nT^2\sqrt{n}}) + O_p(\frac{1}{nT^2}) = O_p(\frac{1}{nT^2}).
\end{aligned}
$$

Proof of part (2)

$$
\begin{aligned}
D_5^2 =& -\frac{1}{nT(T-1)(T-2)}(\tilde{\beta}-\beta)' \sum_{t\neq s\neq \tau}^{T} \sum_{s\neq\tau}^{T} \sum_{\tau=1}^{T} \tilde{x}_s' v_t (\frac{1}{T}\sum_{\eta=1}^{T} v_\eta)' \tilde{x}_\tau (\tilde{\beta}-\beta) \\
=& -\frac{1}{nT^2(T-1)(T-2)}(\tilde{\beta}-\beta)' \sum_{t\neq s\neq \tau}^{T} \sum_{s\neq\tau}^{T} \sum_{\tau=1}^{T} \sum_{\eta=1}^{T} \tilde{x}_s' v_t v_\eta' \tilde{x}_\tau (\tilde{\beta}-\beta) \\
=& -\frac{1}{nT^2(T-1)(T-2)}(\tilde{\beta}-\beta)' \sum_{t\neq s\neq \tau}^{T} \sum_{s\neq\tau}^{T} \sum_{\tau=1}^{T} \sum_{\eta=1}^{T} \sum_{i=1}^{n} \tilde{x}_{is} v_{it} v_{i\eta} \tilde{x}_{i\tau} (\tilde{\beta}-\beta) \\
&- \frac{1}{nT^2(T-1)(T-2)}(\tilde{\beta}-\beta)' \sum_{t\neq s\neq \tau}^{T} \sum_{s\neq\tau}^{T} \sum_{\tau=1}^{T} \sum_{\eta=1}^{T} \sum_{i\neq j}^{n} \sum_{j=1}^{n} \tilde{x}_{is} v_{it} v_{j\eta} \tilde{x}_{j\tau} (\tilde{\beta}-\beta).
\end{aligned}
$$

$\frac{1}{nT^2(T-1)(T-2)} \sum_{t\neq s\neq\tau}^{T} \sum_{s\neq\tau}^{T} \sum_{\tau=1}^{T} \sum_{\eta=1}^{T} \sum_{i=1}^{n} \tilde{x}_{is} v_{it} v_{i\eta} \tilde{x}_{i\tau}$ can be expanded as the following terms:

(1) $t=\eta\neq s\neq\tau$: $\frac{1}{nT^2(T-1)(T-2)} \sum_{t\neq s\neq\tau}^{T} \sum_{s\neq\tau}^{T} \sum_{\tau=1}^{T} \sum_{i=1}^{n} \tilde{x}_{is} v_{it}^2 \tilde{x}_{i\tau} = O_p(\frac{1}{T})$;

(2) $t\neq s=\eta\neq\tau$: $\frac{1}{nT^2(T-1)(T-2)} \sum_{t\neq s\neq\tau}^{T} \sum_{s\neq\tau}^{T} \sum_{\tau=1}^{T} \sum_{i=1}^{n} \tilde{x}_{is} v_{it} v_{is} \tilde{x}_{i\tau} = O_p(\frac{1}{T^2\sqrt{n}})$;

(3) $t \neq s \neq \tau = \eta$: $\frac{1}{nT^2(T-1)(T-2)} \sum_{t\neq s\neq \tau}^{T} \sum_{s\neq \tau}^{T} \sum_{\tau=1}^{T} \sum_{i=1}^{n} \tilde{x}_{is} v_{it} v_{i\tau} \tilde{x}_{i\tau} = O_p(\frac{1}{T^2\sqrt{n}})$;

(4) $t \neq s \neq \tau \neq \eta$: $\frac{1}{nT^2(T-1)(T-2)} \sum_{t\neq s\neq \tau\neq\eta}^{T} \sum_{s\neq \tau\neq\eta}^{T} \sum_{\tau\neq\eta}^{T} \sum_{\eta=1}^{T} \sum_{i=1}^{n} \tilde{x}_{is} v_{it} v_{i\eta} \tilde{x}_{i\tau} = O_p(\frac{1}{T\sqrt{n}})$.

$\frac{1}{nT^2(T-1)(T-2)} \sum_{t\neq s\neq \tau}^{T} \sum_{s\neq \tau}^{T} \sum_{\tau=1}^{T} \sum_{\eta=1}^{T} \sum_{i\neq j}^{n} \sum_{j=1}^{n} \tilde{x}_{is} v_{it} v_{j\eta} \tilde{x}_{j\tau}$ can be also expanded similarly:

(1) $t = \eta \neq s \neq \tau$: $\frac{1}{nT^2(T-1)(T-2)} \sum_{t\neq s\neq \tau}^{T} \sum_{s\neq \tau}^{T} \sum_{\tau=1}^{T} \sum_{i\neq j}^{n} \sum_{j=1}^{n} \tilde{x}_{is} v_{it} v_{jt} \tilde{x}_{j\tau} = O_p(\frac{1}{T\sqrt{T}})$;

(2) $t \neq s = \eta \neq \tau$: $\frac{1}{nT^2(T-1)(T-2)} \sum_{t\neq s\neq \tau}^{T} \sum_{s\neq \tau}^{T} \sum_{\tau=1}^{T} \sum_{i\neq j}^{n} \sum_{j=1}^{n} \tilde{x}_{is} v_{it} v_{js} \tilde{x}_{j\tau} = O_p(\frac{1}{T^2})$;

(3) $t \neq s \neq \tau = \eta$: $\frac{1}{nT^2(T-1)(T-2)} \sum_{t\neq s\neq \tau}^{T} \sum_{s\neq \tau}^{T} \sum_{\tau=1}^{T} \sum_{i\neq j}^{n} \sum_{j=1}^{n} \tilde{x}_{is} v_{it} v_{j\tau} \tilde{x}_{j\tau} = O_p(\frac{1}{T^2})$;

(4) $t \neq s \neq \tau \neq \eta$: $\frac{1}{nT^2(T-1)(T-2)} \sum_{t\neq s\neq \tau\neq\eta}^{T} \sum_{s\neq \tau\neq\eta}^{T} \sum_{\tau\neq\eta}^{T} \sum_{\eta=1}^{T} \sum_{i\neq j}^{n} \sum_{j=1}^{n} \tilde{x}_{is} v_{it} v_{j\eta} \tilde{x}_{j\tau} = O_p(\frac{1}{T})$.

Since $\tilde{\beta} - \beta = O_p(\frac{1}{\sqrt{nT}})$, then we have

$$D_5^2 = O_p(\frac{1}{nT^2}).$$

Proof of part (3)

$$\begin{aligned}
D_5^3 =& -\frac{1}{nT(T-1)}(\tilde{\beta}-\beta)' \sum_{s\neq\tau}^{T} \sum_{\tau=1}^{T} \tilde{x}_s'(\frac{1}{T}\sum_{t=1}^{T} v_t) v_s' \tilde{x}_\tau (\tilde{\beta}-\beta) \\
=& -\frac{1}{nT^2(T-1)}(\tilde{\beta}-\beta)' \sum_{t=1}^{T} \sum_{s\neq\tau}^{T} \sum_{\tau=1}^{T} \tilde{x}_s' v_t v_s' \tilde{x}_\tau (\tilde{\beta}-\beta) \\
=& -\frac{1}{nT^2(T-1)}(\tilde{\beta}-\beta)' \sum_{t=1}^{T} \sum_{s\neq\tau}^{T} \sum_{\tau=1}^{T} \sum_{i=1}^{n} \tilde{x}_{is} v_{it} v_{is} \tilde{x}_{i\tau} (\tilde{\beta}-\beta) \\
& -\frac{1}{nT^2(T-1)}(\tilde{\beta}-\beta)' \sum_{t=1}^{T} \sum_{s\neq\tau}^{T} \sum_{\tau=1}^{T} \sum_{i\neq j}^{n} \sum_{j=1}^{n} \tilde{x}_{is} v_{it} v_{js} \tilde{x}_{j\tau} (\tilde{\beta}-\beta).
\end{aligned}$$

$\frac{1}{nT^2(T-1)} \sum_{t=1}^{T} \sum_{s\neq\tau}^{T} \sum_{\tau=1}^{T} \sum_{i=1}^{n} \tilde{x}_{is} v_{it} v_{is} \tilde{x}_{i\tau}$ has three cases for expansion:

(1) $t = s \neq \tau$: $\frac{1}{nT^2(T-1)} \sum_{s\neq\tau}^{T} \sum_{\tau=1}^{T} \sum_{i=1}^{n} \tilde{x}_{is} v_{is}^2 \tilde{x}_{i\tau} = O_p(\frac{1}{T})$;

(2) $s \neq t = \tau$: $\frac{1}{nT^2(T-1)} \sum_{t\neq s}^{T} \sum_{s=1}^{T} \sum_{i=1}^{n} \tilde{x}_{is} v_{it} v_{is} \tilde{x}_{it} = O_p(\frac{1}{T^2\sqrt{n}})$;

(3) $t \neq s \neq \tau$: $\frac{1}{nT^2(T-1)} \sum_{t\neq s\neq\tau} \sum_{s\neq\tau}^{T} \sum_{\tau=1}^{T} \sum_{i=1}^{n} \tilde{x}_{is} v_{it} v_{is} \tilde{x}_{i\tau} = O_p(\frac{1}{T\sqrt{n}})$.

$\frac{1}{nT^2(T-1)} \sum_{t=1}^{T} \sum_{s\neq\tau}^{T} \sum_{\tau=1}^{T} \sum_{i\neq j}^{n} \sum_{j=1}^{n} \tilde{x}_{is} v_{it} v_{js} \tilde{x}_{j\tau}$ also has three cases for expansion:

(1) $t = s \neq \tau : \frac{1}{nT^2(T-1)} \sum_{s\neq\tau}^T \sum_{\tau=1}^T \sum_{i\neq j}^n \sum_{j=1}^n \tilde{x}_{is} v_{is} v_{js} \tilde{x}_{j\tau} = O_p(\frac{1}{T\sqrt{T}})$;

(2) $s \neq t = \tau : \frac{1}{nT^2(T-1)} \sum_{t\neq s}^T \sum_{s=1}^T \sum_{i\neq j}^n \sum_{j=1}^n \tilde{x}_{is} v_{it} v_{js} \tilde{x}_{jt} = O_p(\frac{1}{T^2})$;

(3) $t \neq s \neq \tau : \frac{1}{nT^2(T-1)} \sum_{t\neq s\neq\tau} \sum_{s\neq\tau}^T \sum_{\tau=1}^T \sum_{i\neq j}^n \sum_{j=1} \tilde{x}_{is} v_{it} v_{js} \tilde{x}_{j\tau} = O_p(\frac{1}{T})$.

With $\tilde{\beta} - \beta = O_p(\frac{1}{\sqrt{nT}})$, then we can get

$$D_5^3 = O_p(\frac{1}{nT^2}).$$

Proof of part (4)

$$\begin{aligned}
D_5^4 =& \frac{1}{nT(T-1)}(\tilde{\beta}-\beta)' \sum_{s\neq\tau}^T \sum_{\tau=1}^T \tilde{x}_s' (\frac{1}{T}\sum_{t=1}^T v_t)(\frac{1}{T}\sum_{\eta=1}^T v_\eta)' \tilde{x}_\tau (\tilde{\beta}-\beta) \\
=& \frac{1}{nT^3(T-1)}(\tilde{\beta}-\beta)' \sum_{t=1}^T \sum_{s\neq\tau}^T \sum_{\tau=1}^T \sum_{\eta=1}^T \tilde{x}_s' v_t v_\eta' \tilde{x}_\tau (\tilde{\beta}-\beta) \\
=& \frac{1}{2} C_5^3 = O_p(\frac{1}{nT^2}).
\end{aligned}$$

■

With the results above, we have

$$D_5 = O_p(\frac{1}{nT^2}).$$

H.6 Proof of part d_6

$$\begin{aligned}
D_6 =& \frac{1}{nT(T-1)(T-2)} \sum_{t\neq s\neq\tau}^T \sum_{s\neq\tau}^T \sum_{\tau=1}^T (v_t - \bar{v}_\cdot)' \tilde{x}_s (\tilde{\beta}-\beta)(\tilde{\beta}-\beta)' \tilde{x}_s' (v_\tau - \bar{v}_\cdot) \\
=& \frac{1}{nT(T-1)(T-2)} (\tilde{\beta}-\beta)' \sum_{t\neq s\neq\tau}^T \sum_{s\neq\tau}^T \sum_{\tau=1}^T \tilde{x}_s' (v_\tau - \bar{v}_\cdot)(v_t - \bar{v}_\cdot)' \tilde{x}_s (\tilde{\beta}-\beta) \\
=& \sum_{k=1}^3 D_6^k.
\end{aligned}$$

where

$D_6^1 = \frac{1}{nT(T-1)(T-2)}(\tilde{\beta}-\beta)' \sum_{t\neq s\neq\tau}^T \sum_{s\neq\tau}^T \sum_{\tau=1}^T \tilde{x}_s' v_\tau v_t' \tilde{x}_s(\tilde{\beta}-\beta)$;

$D_6^2 = -\frac{2}{nT(T-1)(T-2)}(\tilde{\beta}-\beta)' \sum_{s\neq\tau}^T \sum_{\tau=1}^T \tilde{x}_s' v_\tau \bar{v}_.' \tilde{x}_s(\tilde{\beta}-\beta)$;

$D_6^3 = \frac{1}{nT(T-1)(T-2)}(\tilde{\beta}-\beta)' \sum_{s=1}^T \tilde{x}_s' \bar{v}_. \bar{v}_.' \tilde{x}_s(\tilde{\beta}-\beta)$.

Lemma 18 *Under Assumption 1, 2, 3 and the null,*

(1) $D_6^1 = O_p(\frac{1}{nT^2})$;

(2) $D_6^2 = O_p(\frac{1}{nT^2})$;

(3) $D_6^3 = O_p(\frac{1}{nT^2})$.

Proof. Proof of part (1)

$$\begin{aligned}
D_6^1 =& \frac{1}{nT(T-1)(T-2)}(\tilde{\beta}-\beta)' \sum_{t\neq s\neq\tau}^T \sum_{s\neq\tau}^T \sum_{\tau=1}^T \sum_{i=1}^n \tilde{x}_{is}^2 v_{i\tau} v_{it}(\tilde{\beta}-\beta) \\
&+ \frac{1}{nT(T-1)(T-2)}(\tilde{\beta}-\beta)' \sum_{t\neq s\neq\tau}^T \sum_{s\neq\tau}^T \sum_{\tau=1}^T \sum_{i\neq j}^n \sum_{j=1}^n \tilde{x}_{is} v_{i\tau} v_{jt} \tilde{x}_{js}(\tilde{\beta}-\beta) \\
=& O_p(\frac{1}{nT^2\sqrt{n}}) + O_p(\frac{1}{nT^2}).
\end{aligned}$$

Proof of part (2)

$$\begin{aligned}
D_6^2 =& -\frac{2}{nT(T-1)}(\tilde{\beta}-\beta)' \sum_{s\neq\tau}^T \sum_{\tau=1}^T \tilde{x}_s' v_\tau (\frac{1}{T}\sum_{t=1}^T v_t)' \tilde{x}_s(\tilde{\beta}-\beta) \\
=& -\frac{2}{nT^2(T-1)}(\tilde{\beta}-\beta)' \sum_{t=1}^T \sum_{s\neq\tau}^T \sum_{\tau=1}^T \tilde{x}_s' v_\tau v_t' \tilde{x}_s(\tilde{\beta}-\beta) \\
=& -\frac{2}{nT^2(T-1)}(\tilde{\beta}-\beta)' \sum_{t=1}^T \sum_{s\neq\tau}^T \sum_{\tau=1}^T \sum_{i=1}^n \tilde{x}_{is}^2 v_{i\tau} v_{it}(\tilde{\beta}-\beta) \\
&- \frac{2}{nT^2(T-1)}(\tilde{\beta}-\beta)' \sum_{t=1}^T \sum_{s\neq\tau}^T \sum_{\tau=1}^T \sum_{i\neq j}^n \sum_{j=1}^n \tilde{x}_{is} v_{i\tau} v_{jt} \tilde{x}_{js}(\tilde{\beta}-\beta).
\end{aligned}$$

$-\frac{1}{nT^2(T-1)}\sum_{t=1}^{T}\sum_{t=1}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n}\tilde{x}_{is}^2 v_{i\tau}v_{it}$ can be expanded as the following terms:

(1) $t=s\neq\tau: -\frac{1}{nT^2(T-1)}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n}\tilde{x}_{is}^2 v_{i\tau}v_{is}=O_p(\frac{1}{T^2\sqrt{n}})$;

(2) $s\neq\tau=t: -\frac{1}{nT^2(T-1)}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n}\tilde{x}_{is}^2 v_{i\tau}^2=O_p(\frac{1}{T})$;

(3) $t\neq s\neq\tau: -\frac{1}{nT^2(T-1)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n}\tilde{x}_{is}^2 v_{i\tau}v_{it}=O_p(\frac{1}{T\sqrt{n}})$.

and $-\frac{1}{nT^2(T-1)}\sum_{t=1}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}\tilde{x}_{is}v_{i\tau}v_{jt}\tilde{x}_{js}$ can be also expanded as the following cases:

(1) $t=s\neq\tau: -\frac{1}{nT^2(T-1)}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}\tilde{x}_{is}v_{i\tau}v_{js}\tilde{x}_{js}=O_p(\frac{1}{T^2})$;

(2) $s\neq\tau=t: -\frac{1}{nT^2(T-1)}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}\tilde{x}_{is}v_{i\tau}v_{j\tau}\tilde{x}_{js}=O_p(\frac{1}{T\sqrt{T}})$;

(3) $t\neq s\neq\tau: -\frac{1}{nT^2(T-1)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}\tilde{x}_{is}v_{i\tau}v_{jt}\tilde{x}_{js}=O_p(\frac{1}{T})$.

Hence, with $\tilde{\beta}-\beta=O_p(\frac{1}{\sqrt{nT}})$, we have

$$D_6^2=O_p(\frac{1}{nT^2}).$$

Proof of part (3)

$$\begin{aligned}
B_6^3=&\frac{1}{nT}(\tilde{\beta}-\beta)'\sum_{s=1}^{T}\tilde{x}_s'(\frac{1}{T}\sum_{t=1}^{T}v_t)(\frac{1}{T}\sum_{\tau=1}^{T}v_\tau)'\tilde{x}_s(\tilde{\beta}-\beta)\\
=&\frac{1}{nT^3}(\tilde{\beta}-\beta)'\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\tilde{x}_s'v_t v_\tau'\tilde{x}_s(\tilde{\beta}-\beta)\\
=&\frac{1}{nT^3}(\tilde{\beta}-\beta)'\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n}\tilde{x}_{is}^2 v_{it}v_{i\tau}(\tilde{\beta}-\beta)\\
&+\frac{1}{nT^3}(\tilde{\beta}-\beta)'\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}\tilde{x}_{is}v_{it}v_{j\tau}\tilde{x}_{js}(\tilde{\beta}-\beta).
\end{aligned}$$

$\frac{1}{nT^3}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n}\tilde{x}_{is}^2 v_{it}v_{i\tau}$ can be expanded as the following terms:

(1) $t=s=\tau: \frac{1}{nT^3}\sum_{t=1}^{T}\sum_{i=1}^{n}\tilde{x}_{it}^2 v_{it}^2=O_p(\frac{1}{T^2})$;

(2) $t=s\neq\tau: \frac{1}{nT^3}\sum_{t\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n}\tilde{x}_{it}^2 v_{it}v_{i\tau}=O_p(\frac{1}{T^2\sqrt{n}})$;

(3) $t=\tau\neq s: \frac{1}{nT^3}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n}\tilde{x}_{is}^2 v_{it}^2=O_p(\frac{1}{T})$;

(4) $t\neq s=\tau: \frac{1}{nT^3}\sum_{t\neq s}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n}\tilde{x}_{is}^2 v_{it}v_{is}=O_p(\frac{1}{T^2\sqrt{n}})$;

(5) $t \neq s \neq \tau$: $\frac{1}{nT^3} \sum_{t \neq s \neq \tau}^{T} \sum_{s \neq \tau}^{T} \sum_{\tau=1}^{T} \sum_{i=1}^{n} \tilde{x}_{is}^2 v_{it} v_{i\tau} = O_p(\frac{1}{T\sqrt{n}})$.

And for the expansions of $\frac{1}{nT^3} \sum_{t=1}^{T} \sum_{s=1}^{T} \sum_{\tau=1}^{T} \sum_{i \neq j}^{n} \sum_{j=1}^{n} \tilde{x}_{is} v_{it} v_{j\tau} \tilde{x}_{js}$, we have the following cases:

(1) $t = s = \tau$: $\frac{1}{nT^3} \sum_{t=1}^{T} \sum_{i \neq j}^{n} \sum_{j=1}^{n} \tilde{x}_{it} v_{it} v_{jt} \tilde{x}_{jt} = O_p(\frac{1}{T^2\sqrt{T}})$;

(2) $t = s \neq \tau$: $\frac{1}{nT^3} \sum_{t \neq \tau}^{T} \sum_{\tau=1}^{T} \sum_{i \neq j}^{n} \sum_{j=1}^{n} \tilde{x}_{it} v_{it} v_{j\tau} \tilde{x}_{jt} = O_p(\frac{1}{T^2})$;

(3) $t = \tau \neq s$: $\frac{1}{nT^3} \sum_{t \neq s}^{T} \sum_{s=1}^{T} \sum_{i \neq j}^{n} \sum_{j=1}^{n} \tilde{x}_{is} v_{it} v_{jt} \tilde{x}_{js} = O_p(\frac{1}{T\sqrt{T}})$;

(4) $t \neq s = \tau$: $\frac{1}{nT^3} \sum_{t \neq s}^{T} \sum_{s=1}^{T} \sum_{i \neq j}^{n} \sum_{j=1}^{n} \tilde{x}_{is} v_{it} v_{js} \tilde{x}_{js} = O_p(\frac{1}{T^2})$;

(5) $t \neq s \neq \tau$: $\frac{1}{nT^3} \sum_{t \neq s \neq \tau}^{T} \sum_{s \neq \tau}^{T} \sum_{\tau=1}^{T} \sum_{i \neq j}^{n} \sum_{j=1}^{n} \tilde{x}_{is} v_{it} v_{j\tau} \tilde{x}_{js} = O_p(\frac{1}{T})$.

Since $\tilde{\beta} - \beta = O_p(\frac{1}{\sqrt{nT}})$, then we have

$$D_6^3 = O_p(\frac{1}{nT^2}).$$

■

With the results of $D_6^k, k = 1, 2, 3$, the order of magnitude of B_6 can be got, which is:

$$D_6 = O_p(\frac{1}{nT^2}).$$

H.7 Proof of d_7

$$\begin{aligned} D_7 = &- \frac{2}{nT(T-1)(T-2)} \sum_{t \neq s \neq \tau}^{T} \sum_{s \neq \tau}^{T} \sum_{\tau=1}^{T} \tilde{v}_t' \tilde{x}_s (\tilde{\beta} - \beta)(\tilde{\beta} - \beta)' \tilde{x}_s' \tilde{x}_\tau (\tilde{\beta} - \beta) \\ = &- \frac{2}{nT(T-1)(T-2)} (\tilde{\beta} - \beta)' (\sum_{t \neq s \neq \tau}^{T} \sum_{s \neq \tau}^{T} \sum_{\tau=1}^{T} \tilde{x}_s' \tilde{v}_t (\tilde{\beta} - \beta) \tilde{x}_s' \tilde{x}_\tau) (\tilde{\beta} - \beta). \end{aligned}$$

We need to calculate the order of magnitude of $\frac{1}{nT(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\tilde{x}_s'\tilde{v}_t\tilde{x}_s'\tilde{x}_\tau$.

$$
\begin{aligned}
&\frac{1}{nT(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\tilde{x}_s'\tilde{v}_t\tilde{x}_s'\tilde{x}_\tau\\
=&\frac{1}{nT(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\tilde{x}_s'(v_t-\bar{v}.)\tilde{x}_s'\tilde{x}_\tau\\
=&\frac{1}{nT(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\tilde{x}_s'v_t\tilde{x}_s'\tilde{x}_\tau-\frac{1}{nT(T-1)}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\tilde{x}_s'\bar{v}.\tilde{x}_s'\tilde{x}_\tau.
\end{aligned}
$$

Where

$$
\begin{aligned}
&\frac{1}{nT(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\tilde{x}_s'v_t\tilde{x}_s'\tilde{x}_\tau\\
=&\frac{1}{nT(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n}\tilde{x}_{is}^2\tilde{x}_{i\tau}v_{it}\\
&+\frac{1}{nT(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}\tilde{x}_{is}v_{it}\tilde{x}_{js}\tilde{x}_{j\tau}\\
=&O_p(\frac{1}{\sqrt{nT}})+O_p(\frac{\sqrt{n}}{\sqrt{T}}).
\end{aligned}
$$

and

$$
\begin{aligned}
&\frac{1}{nT(T-1)}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\tilde{x}_s'\bar{v}.\tilde{x}_s'\tilde{x}_\tau\\
=&\frac{1}{nT(T-1)}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\tilde{x}_s'(\frac{1}{T}\sum_{t=1}^{T}v_t)\tilde{x}_s'\tilde{x}_\tau=\frac{1}{nT^2(T-1)}\sum_{t=1}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\tilde{x}_s'v_t\tilde{x}_s'\tilde{x}_\tau\\
=&\frac{1}{nT^2(T-1)}\sum_{t=1}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n}\tilde{x}_{is}^2\tilde{x}_{i\tau}v_{it}+\frac{1}{nT^2(T-1)}\sum_{t=1}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}\tilde{x}_{is}v_{it}\tilde{x}_{js}\tilde{x}_{j\tau}.
\end{aligned}
$$

To calculate the order of magnitude of the above term, we need to distinguish the cases of $t,s\neq\tau$, the first term then can be expanded as the following 3 expansions:

(1) $t=s\neq\tau:\frac{1}{nT^2(T-1)}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n}\tilde{x}_{is}^2\tilde{x}_{i\tau}v_{is}=O_p(\frac{1}{T\sqrt{nT}})$;

(2) $s \neq \tau = t$: $\frac{1}{nT^2(T-1)} \sum_{s\neq\tau}^T \sum_{\tau=1}^T \sum_{i=1}^n \tilde{x}_{is}^2 \tilde{x}_{i\tau} v_{i\tau} = O_p(\frac{1}{T\sqrt{nT}})$;

(3) $t \neq s \neq \tau$: $\frac{1}{nT^2(T-1)} \sum_{t\neq s\neq\tau}^T \sum_{s\neq\tau}^T \sum_{\tau=1}^T \sum_{i=1}^n \tilde{x}_{is}^2 \tilde{x}_{i\tau} v_{it} = O_p(\frac{1}{\sqrt{nT}})$.

Similarly, we can also get the following three cases for expansions of the second term:

(1) $t = s \neq \tau$: $\frac{1}{nT^2(T-1)} \sum_{s\neq\tau}^T \sum_{\tau=1}^T \sum_{i\neq j}^n \sum_{j=1}^n \tilde{x}_{is} v_{is} \tilde{x}_{js} \tilde{x}_{j\tau} = O_p(\frac{1}{T\sqrt{T}})$;

(2) $s \neq \tau = t$: $\frac{1}{nT^2(T-1)} \sum_{s\neq\tau}^T \sum_{\tau=1}^T \sum_{i\neq j}^n \sum_{j=1}^n \tilde{x}_{is} v_{i\tau} \tilde{x}_{js} \tilde{x}_{j\tau} = O_p(\frac{1}{T\sqrt{T}})$;

(3) $t \neq s \neq \tau$: $\frac{1}{nT^2(T-1)} \sum_{t\neq s\neq\tau}^T \sum_{s\neq\tau}^T \sum_{\tau=1}^T \sum_{i\neq j}^n \sum_{j=1}^n \tilde{x}_{is} v_{it} \tilde{x}_{js} \tilde{x}_{j\tau} = O_p(\frac{\sqrt{n}}{\sqrt{T}})$.

Hence, we get $\frac{1}{nT(T-1)(T-2)} \sum_{t\neq s\neq\tau}^T \sum_{s\neq\tau}^T \sum_{\tau=1}^T \tilde{x}_s' \tilde{v}_t \tilde{x}_s' \tilde{x}_\tau = O_p(\frac{\sqrt{n}}{\sqrt{T}})$ and with the fact that $\tilde{\beta} - \beta = O_p(\frac{1}{\sqrt{nT}})$, we have

$$D_7 = O_p(\frac{1}{nT^2}).$$

H.8 Proof of part d_8

$$\begin{aligned} D_8 =& \frac{1}{nT(T-1)(T-2)} (\tilde{\beta} - \beta)' \sum_{t\neq s\neq\tau}^T \sum_{s\neq\tau}^T \sum_{\tau=1}^T \tilde{x}_t' (v_s - \bar{v}_\cdot)(v_s - \bar{v}_\cdot)' \tilde{x}_\tau (\tilde{\beta} - \beta) \\ =& \sum_{k=1}^3 D_8^k. \end{aligned}$$

where

$D_8^1 = \frac{1}{nT(T-1)(T-2)} (\tilde{\beta} - \beta)' \sum_{t\neq s\neq\tau}^T \sum_{s\neq\tau}^T \sum_{\tau=1}^T \tilde{x}_t' v_s v_s' \tilde{x}_\tau (\tilde{\beta} - \beta)$;

$D_8^2 = -\frac{2}{nT(T-1)(T-2)} (\tilde{\beta} - \beta)' \sum_{t\neq s\neq\tau}^T \sum_{s\neq\tau}^T \sum_{\tau=1}^T \tilde{x}_t' v_s \bar{v}_\cdot' \tilde{x}_\tau (\tilde{\beta} - \beta)$;

$D_8^3 = \frac{1}{nT(T-1)} (\tilde{\beta} - \beta)' \sum_{t\neq\tau}^T \sum_{\tau=1}^T \tilde{x}_t' \bar{v}_\cdot \bar{v}_\cdot' \tilde{x}_\tau (\tilde{\beta} - \beta)$.

Lemma 19 *Under Assumption 1, 2, 3 and the null,*

(1) $D_8^1 = O_p(\frac{1}{nT})$;

(2) $D_8^2 = O_p(\frac{1}{nT^2})$;

(3) $D_8^3 = O_p(\frac{1}{nT^2})$.

Proof. Consider part (1),

$$
\begin{aligned}
D_8^1 =& \frac{1}{nT(T-1)(T-2)}(\tilde{\beta}-\beta)' \sum_{t\neq s\neq \tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i=1}^{n}\tilde{x}_{it}v_{is}^2\tilde{x}_{i\tau}(\tilde{\beta}-\beta)\\
&+\frac{1}{nT(T-1)(T-2)}(\tilde{\beta}-\beta)' \sum_{t\neq s\neq \tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}\tilde{x}_{it}v_{is}v_{js}\tilde{x}_{j\tau}(\tilde{\beta}-\beta)\\
=&O_p(\frac{1}{nT})+O_p(\frac{1}{nT\sqrt{T}})=O_p(\frac{1}{nT}).
\end{aligned}
$$

next consider part (2), $D_8^2=\frac{1}{2}C_5^2=O_p(\frac{1}{nT^2})$; for part (3) $D_8^3=\frac{1}{2}C_5^3=O_p(\frac{1}{nT^2})$. ■

As a result, we have

$$D_8=O_p(\frac{1}{nT}).$$

H.9 Proof of part d_9

$$
\begin{aligned}
D_9=&-\frac{2}{nT(T-1)(T-2)}\sum_{t\neq s\neq \tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}(\tilde{\beta}-\beta)'\tilde{x}_t'(v_s-\bar{v}_.)(\tilde{\beta}-\beta)'\tilde{x}_s'\tilde{x}_\tau(\tilde{\beta}-\beta)\\
=&-\frac{2}{nT(T-1)(T-2)}(\tilde{\beta}-\beta)'\sum_{t\neq s\neq \tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\tilde{x}_s'\tilde{x}_\tau(\tilde{\beta}-\beta)(v_s-\bar{v}_.)'\tilde{x}_t(\tilde{\beta}-\beta)\\
=&-\frac{2}{nT(T-1)(T-2)}(\tilde{\beta}-\beta)'\sum_{t\neq s\neq \tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\tilde{x}_s'\tilde{x}_\tau(\tilde{\beta}-\beta)v_s'\tilde{x}_t(\tilde{\beta}-\beta)\\
&+\frac{2}{nT(T-1)(T-2)}(\tilde{\beta}-\beta)'\sum_{t\neq s\neq \tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T}\tilde{x}_s'\tilde{x}_\tau(\tilde{\beta}-\beta)\bar{v}_.'\tilde{x}_t(\tilde{\beta}-\beta).
\end{aligned}
$$

First, let's consider $\frac{1}{nT(T-1)(T-2)}\sum_{t\neq s\neq\tau}^T\sum_{s\neq\tau}^T\sum_{\tau=1}^T\tilde{x}_s'\tilde{x}_\tau v_s'\tilde{x}_t$,

$$\begin{aligned}
&\frac{1}{nT(T-1)(T-2)}\sum_{t\neq s\neq\tau}^T\sum_{s\neq\tau}^T\sum_{\tau=1}^T\tilde{x}_s'\tilde{x}_\tau v_s'\tilde{x}_t\\
=&\frac{1}{nT(T-1)(T-2)}\sum_{t\neq s\neq\tau}^T\sum_{s\neq\tau}^T\sum_{\tau=1}^T\sum_{i=1}^n\tilde{x}_{it}\tilde{x}_{i\tau}v_{is}\tilde{x}_{it}\\
&+\frac{1}{nT(T-1)(T-2)}\sum_{t\neq s\neq\tau}^T\sum_{s\neq\tau}^T\sum_{\tau=1}^T\sum_{i\neq j}^n\sum_{j=1}^n\tilde{x}_{it}\tilde{x}_{i\tau}v_{js}\tilde{x}_{jt}\\
=&O_p(\frac{1}{\sqrt{nT}})+O_p(\frac{\sqrt{n}}{\sqrt{T}}).
\end{aligned}$$

then, we consider $\frac{1}{nT(T-1)(T-2)}\sum_{t\neq s\neq\tau}^T\sum_{s\neq\tau}^T\sum_{\tau=1}^T\tilde{x}_s'\tilde{x}_\tau\bar{v}_.'\tilde{x}_t$,

$$\begin{aligned}
&\frac{1}{nT(T-1)(T-2)}\sum_{t\neq s\neq\tau}^T\sum_{s\neq\tau}^T\sum_{\tau=1}^T\tilde{x}_s'\tilde{x}_\tau\bar{v}_.'\tilde{x}_t\\
=&\frac{1}{nT(T-1)(T-2)}\sum_{t\neq s\neq\tau}^T\sum_{s\neq\tau}^T\sum_{\tau=1}^T\tilde{x}_s'\tilde{x}_\tau(\frac{1}{T}\sum_{\eta=1}^T v_\eta)\tilde{x}_t\\
=&\frac{1}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^T\sum_{s\neq\tau}^T\sum_{\tau=1}^T\sum_{\eta=1}^T\tilde{x}_s'\tilde{x}_\tau v_\eta'\tilde{x}_t\\
=&\frac{1}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^T\sum_{s\neq\tau}^T\sum_{\tau=1}^T\sum_{\eta=1}^T\sum_{i=1}^n\tilde{x}_{is}\tilde{x}_{i\tau}v_{i\eta}\tilde{x}_{it}\\
&+\frac{1}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^T\sum_{s\neq\tau}^T\sum_{\tau=1}^T\sum_{\eta=1}^T\sum_{i\neq j}^n\sum_{j=1}^n\tilde{x}_{is}\tilde{x}_{i\tau}v_{j\eta}\tilde{x}_{jt}.
\end{aligned}$$

The first term of the above expansions can be calculated by sum of the following cases:

$t=\eta\neq s\neq\tau$: $\frac{1}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^T\sum_{s\neq\tau}^T\sum_{\tau=1}^T\sum_{i=1}^n\tilde{x}_{is}\tilde{x}_{i\tau}v_{it}\tilde{x}_{it}=O_p(\frac{1}{T\sqrt{nT}})$;

$t\neq s=\eta\neq\tau$: $\frac{1}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^T\sum_{s\neq\tau}^T\sum_{\tau=1}^T\sum_{i=1}^n\tilde{x}_{is}\tilde{x}_{i\tau}v_{is}\tilde{x}_{it}=O_p(\frac{1}{T\sqrt{nT}})$;

$t\neq s\neq\tau=\eta$: $\frac{1}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^T\sum_{s\neq\tau}^T\sum_{\tau=1}^T\sum_{i=1}^n\tilde{x}_{is}\tilde{x}_{i\tau}v_{i\tau}\tilde{x}_{it}=O_p(\frac{1}{T\sqrt{nT}})$;

$t\neq s\neq\tau\neq\eta$: $\frac{1}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau\neq\eta}^T\sum_{s\neq\tau\neq\eta}^T\sum_{\tau\neq\eta}^T\sum_{\eta=1}^T\sum_{i=1}^n\tilde{x}_{is}\tilde{x}_{i\tau}v_{i\eta}\tilde{x}_{it}=O_p(\frac{1}{\sqrt{nT}})$.

The second term can be also expanded similarly:

$t=\eta\neq s\neq\tau$: $\frac{1}{nT^2(T-1)(T-2)}\sum_{t\neq s\neq\tau}^T\sum_{s\neq\tau}^T\sum_{\tau=1}^T\sum_{i\neq j}^n\sum_{j=1}^n\tilde{x}_{is}\tilde{x}_{i\tau}v_{jt}\tilde{x}_{jt}=O_p(\frac{\sqrt{n}}{T\sqrt{T}})$;

$t \neq s = \eta \neq \tau : \frac{1}{nT^2(T-1)(T-2)} \sum_{t\neq s\neq \tau}^{T} \sum_{s\neq\tau}^{T} \sum_{\tau=1}^{T} \sum_{i\neq j}^{n} \sum_{j=1}^{n} \tilde{x}_{is}\tilde{x}_{i\tau} v_{js}\tilde{x}_{jt} = O_p(\frac{\sqrt{n}}{T\sqrt{T}})$;

$t \neq s \neq \tau = \eta : \frac{1}{nT^2(T-1)(T-2)} \sum_{t\neq s\neq \tau}^{T} \sum_{s\neq\tau}^{T} \sum_{\tau=1}^{T} \sum_{i\neq j}^{n} \sum_{j=1}^{n} \tilde{x}_{is}\tilde{x}_{i\tau} v_{j\tau}\tilde{x}_{jt} = O_p(\frac{\sqrt{n}}{T\sqrt{T}})$;

$t \neq s \neq \tau \neq \eta : \frac{1}{nT^2(T-1)(T-2)} \sum_{t\neq s\neq \tau\neq\eta}^{T} \sum_{s\neq\tau\neq\eta}^{T} \sum_{\tau\neq\eta}^{T} \sum_{\eta=1}^{T} \sum_{i\neq j}^{n} \sum_{j=1}^{n} \tilde{x}_{is}\tilde{x}_{i\tau} v_{j\eta}\tilde{x}_{jt} = O_p(\frac{\sqrt{n}}{\sqrt{T}})$.

With the fact that $\tilde{\beta} - \beta = O_p(\frac{1}{\sqrt{nT}})$, then we get

$$D_9 = O_p(\frac{1}{nT^2}).$$

H.10 Proof of part d_{10}

To calculate the order of magnitude of D_{10}, we first need to calculate the following term:

$$\begin{aligned}
&\frac{1}{nT(T-1)(T-2)} \sum_{t\neq s\neq\tau}^{T} \sum_{s\neq\tau}^{T} \sum_{\tau=1}^{T} \tilde{x}_t \tilde{x}_s \tilde{x}_s' \tilde{x}_\tau \\
=&\frac{1}{nT(T-1)(T-2)} \sum_{t\neq s\neq\tau}^{T} \sum_{s\neq\tau}^{T} \sum_{\tau=1}^{T} \sum_{i=1}^{n} \tilde{x}_{it} \tilde{x}_{is}^2 \tilde{x}_{i\tau} \\
&+ \frac{1}{nT(T-1)(T-2)} \sum_{t\neq s\neq\tau}^{T} \sum_{s\neq\tau}^{T} \sum_{\tau=1}^{T} \sum_{i\neq j}^{n} \sum_{j=1}^{n} \tilde{x}_{it} \tilde{x}_{is} \tilde{x}_{js} \tilde{x}_{j\tau} \\
=&O_p(1) + O_p(n).
\end{aligned}$$

With $\tilde{\beta} - \beta = O_p(\frac{1}{\sqrt{nT}})$, then we have

$$D_{10} = O_p(\frac{1}{nT^2}).$$

I Proof of part (3) of Proposition 2

$$\begin{aligned}
&\frac{1}{n}\hat{M}_{5,T} = \frac{1}{nT_4}\sum_{t,s,\tau,\eta}^{T}\hat{v}_t'\hat{v}_s\hat{v}_\tau'\hat{v}_\eta \\
=&\frac{1}{nT_4}\sum_{t,s,\tau,\eta}^{T}(\tilde{v}_t-\tilde{x}_t(\tilde{\beta}-\beta))'(\tilde{v}_s-\tilde{x}_s(\tilde{\beta}-\beta))(\tilde{v}_\tau-\tilde{x}_\tau(\tilde{\beta}-\beta))'(\tilde{v}_\eta-\tilde{x}_\eta(\tilde{\beta}-\beta)) \\
=&\sum_{k=1}^{6}E_i.
\end{aligned}$$

Where

$E_1=\frac{1}{nT_4}\sum_{t,s,\tau,\eta}^{T}\tilde{v}_t'\tilde{v}_s\tilde{v}_\tau'\tilde{v}_\eta$; $E_2=-\frac{4}{nT_4}\sum_{t,s,\tau,\eta}^{T}\tilde{v}_t'\tilde{v}_s\tilde{v}_\tau'\tilde{x}_\eta(\tilde{\beta}-\beta)$; $E_3=\frac{2}{nT_4}\sum_{t,s,\tau,\eta}^{T}(\tilde{\beta}-\beta)'\tilde{x}_t'\tilde{v}_s\tilde{v}_\tau'\tilde{x}_\eta(\tilde{\beta}-\beta)$; $E_4=\frac{4}{nT_4}\sum_{t,s,\tau,\eta}^{T}(\tilde{\beta}-\beta)'\tilde{x}_t'\tilde{x}_s(\tilde{\beta}-\beta)\tilde{v}_\tau'\tilde{v}_\eta$; $E_5=-\frac{4}{nT_4}\sum_{t,s,\tau,\eta}^{T}(\tilde{\beta}-\beta)'\tilde{x}_t'\tilde{v}_s(\tilde{\beta}-\beta)'\tilde{x}_\tau'\tilde{x}_\eta(\tilde{\beta}-\beta)$; $E_6=-\frac{1}{nT_4}\sum_{t,s,\tau,\eta}^{T}(\tilde{\beta}-\beta)'\tilde{x}_t'\tilde{x}_s(\tilde{\beta}-\beta)(\tilde{\beta}-\beta)'\tilde{x}_\tau'\tilde{x}_\eta(\tilde{\beta}-\beta)$.

Lemma 20 *Under Assumption 1, 2, 3 and the null,*

(e_1) $E_1=\frac{1}{n}M_{5,T}+4\frac{n-1}{T^2}\sigma_v^4-\frac{3n}{T^2}\sigma_v^4+O_p(\frac{\sqrt{n}}{T^2})$;

(e_2) $E_2=O_p(\frac{1}{T^2})$;

(e_3) $E_3=O_p(\frac{1}{nT^2})$;

(e_4) $E_4=O_p(\frac{1}{T^2})$;

(e_5) $E_5=O_p(\frac{1}{nT^2})$;

(e_6) $E_6=O_p(\frac{1}{nT^2})$.

Proof. Proof of part e_1

$$\begin{aligned}E_1 =&\frac{1}{nT_4}\sum_{t,s,\tau,\eta}^{T}\tilde{v}_t'\tilde{v}_s\tilde{v}_\tau'\tilde{v}_\eta\\ =&\frac{1}{nT_4}\sum_{t,s,\tau,\eta}^{T}(v_t-\bar{v}.)'(v_s-\bar{v}.)(v_\tau-\bar{v}.)'(v_\eta-\bar{v}.)\\ =&\frac{1}{n}M_{5,T}+\sum_{k=1}^{5}C_1^k.\end{aligned}$$

Where

$E_1^1 = -\frac{4}{nT(T-1)(T-2)}\sum_{t\neq s\neq\tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\eta=1}^{T}v_t'v_sv_\tau'\bar{v}.;$

$E_1^2 = \frac{3}{nT(T-1)}\sum_{t\neq s}^{T}\sum_{s=1}^{T}v_t'v_s\bar{v}_.'\bar{v}.;$

$E_1^3 = \frac{3}{nT(T-1)}\sum_{t\neq\tau}^{T}\sum_{\tau=1}^{T}v_t'\bar{v}.v_\tau'\bar{v}.;$

$E_1^4 = -\frac{3}{n}\bar{v}'\,\bar{v}.\bar{v}_.'\bar{v}..$

It is easy to show that $E_1^1 = 2D_1^2 = O_p(\frac{\sqrt{n}}{T^2})$; $E_1^2 = D_1^3 = O_p(\frac{\sqrt{n}}{T^2})$; $E_1^3 = 2C_1^4 = 4\frac{n-1}{T^2}\sigma_v^4 + O_p(\frac{\sqrt{n}}{T^2})$ and $E_1^4 = D_1^6 = -\frac{3n}{T^2}\sigma_v^4 + O_p(\frac{\sqrt{n}}{T^2})$, then we can get

$$E_1 = \frac{1}{n}M_{5,T} + 4\frac{n-1}{T^2}\sigma_v^4 - \frac{3n}{T^2}\sigma_v^4 + O_p(\frac{\sqrt{n}}{T^2}).$$

Proof of part e_2

$$\begin{aligned}E_2 =&-\frac{4}{nT_4}\sum_{t,s,\tau,\eta}^{T}(v_t-\bar{v}.)'(v_s-\bar{v}.)(v_\tau-\bar{v}.)'\tilde{x}_\eta(\tilde{\beta}-\beta)\\ =&\sum_{k=1}^{6}E_2^k.\end{aligned}$$

Where

$E_2^1 = -\frac{4}{nT_4}\sum_{t,s,\tau,\eta}^{T}v_t'v_sv_\tau'\tilde{x}_\eta(\tilde{\beta}-\beta);$

$E_2^2 = \frac{4}{nT(T-1)(T-2)}\sum_{t\neq s\neq\eta}^{T}\sum_{s\neq\eta}^{T}\sum_{\eta=1}^{T}v_t'v_s\bar{v}_.'\tilde{x}_\eta(\tilde{\beta}-\beta);$

$E_2^3 = \frac{8}{nT(T-1)(T-2)} \sum_{t\neq\tau\neq\eta}^T \sum_{s\neq\tau}^T \sum_{\eta=1}^T v_t'\bar{v}.v_\tau'\tilde{x}_\eta(\tilde{\beta}-\beta)$;

$E_2^4 = -\frac{8}{nT(T-1)} \sum_{t\neq\eta}^T \sum_{\eta=1}^T v_t'\bar{v}.\bar{v}_.'\tilde{x}_\eta(\tilde{\beta}-\beta)$;

$E_2^5 = -\frac{4}{nT(T-1)} \sum_{\tau\neq\eta}^T \sum_{\eta=1}^T \bar{v}_.'\bar{v}.v_\tau'\tilde{x}_\eta(\tilde{\beta}-\beta)$;

$E_2^6 = \frac{4}{nT} \sum_{\eta=1}^T \bar{v}_.'\bar{v}.\bar{v}_.'\tilde{x}_\eta(\tilde{\beta}-\beta)$.

To prove it, first we consider E_2^1,

$$\begin{aligned}
E_2^1 = &- \frac{4}{nT_4}\sum_{t,s,\tau,\eta}^T \sum_{i=1}^n v_{it}v_{is}v_{i\tau}\tilde{x}_{i\eta}(\tilde{\beta}-\beta) - \frac{4}{nT_4}\sum_{t,s,\tau,\eta}^T \sum_{i\neq j}^n \sum_{j=1}^n v_{it}v_{is}v_{j\tau}\tilde{x}_{j\eta}(\tilde{\beta}-\beta) \\
=&O_p(\frac{1}{nT^2}) + O_p(\frac{1}{T^2\sqrt{n}}) = O_p(\frac{1}{T^2\sqrt{n}}).
\end{aligned}$$

then we can easily show that $E_2^2 = 4D_2^5 = O_p(\frac{1}{T^2\sqrt{n}})$; $E_2^3 = 8D_2^2 = O_p(\frac{1}{T^2})$; $E_2^4 = 4D_2^6 = O_p(\frac{1}{T^2})$; $E_2^5 = 4D_2^4 = O_p(\frac{1}{T^2})$ and $E_2^6 = 4D_2^7 = O_p(\frac{1}{T^2})$, then we have

$$E_2 = O_p(\frac{1}{T^2}).$$

Proof of part e_3

$$\begin{aligned}
E_3 =&\frac{3}{nT_4}(\tilde{\beta}-\beta)'\sum_{t,s,\tau,\eta}^T \tilde{x}_t'(v_s-\bar{v}.)(v_\tau-\bar{v}.)'\tilde{x}_\eta(\tilde{\beta}-\beta) \\
=&\sum_{k=1}^3 E_3^k.
\end{aligned}$$

Where

$E_3^1 = \frac{3}{nT_4}(\tilde{\beta}-\beta)'\sum_{t,s,\tau,\eta}^T \tilde{x}_t'v_sv_\tau'\tilde{x}_\eta(\tilde{\beta}-\beta)$;

$E_3^2 = -\frac{6}{nT(T-1)(T-2)}(\tilde{\beta}-\beta)'\sum_{t\neq\tau\neq\eta}^T \sum_{\tau\neq\eta}^T \sum_{\eta=1}^T \tilde{x}_t'\bar{v}.v_\tau'\tilde{x}_\eta(\tilde{\beta}-\beta)$;

$E_3^3 = \frac{3}{nT(T-1)}(\tilde{\beta}-\beta)'\sum_{t\neq\tau\neq\eta}^T \sum_{\tau\neq\eta}^T \sum_{\eta=1}^T \tilde{x}_t'\bar{v}.\bar{v}_.'\tilde{x}_\eta(\tilde{\beta}-\beta)$.

We first consider E_3^1

$$E_3^1 = \frac{3}{nT_4}(\tilde{\beta}-\beta)' \sum_{t,s,\tau,\eta}^{T} \sum_{i=1}^{n} \tilde{x}_{it} v_{is} v_{i\tau} \tilde{x}_{i\eta} (\tilde{\beta}-\beta) + \frac{3}{nT_4}(\tilde{\beta}-\beta)' \sum_{t,s,\tau,\eta}^{T} \sum_{i\neq j}^{n} \sum_{j=1}^{n} \tilde{x}_{it} v_{is} v_{j\tau} \tilde{x}_{j\eta} (\tilde{\beta}-\beta)$$
$$= O_p(\frac{1}{nT^2\sqrt{n}}) + O_p(\frac{1}{nT^2}) = O_p(\frac{1}{nT^2}).$$

and with the fact that $E_3^2 = 6D_5^2 = O_p(\frac{1}{nT^2})$; $E_3^3 = 3D_5^4 = O_p(\frac{1}{nT^2})$, we have the result

$$E_3 = O_p(\frac{1}{nT^2}).$$

Proof of part e_4

$$E_4 = \frac{3}{nT_4}(\tilde{\beta}-\beta)' \sum_{t,s,\tau,\eta}^{T} \tilde{x}_t' \tilde{x}_s (v_\tau - \bar{v}.)'(v_\eta - \bar{v}.)(\tilde{\beta}-\beta)$$
$$= \sum_{k=1}^{k} E_4^k.$$

Where

$E_4^1 = \frac{3}{nT_4}(\tilde{\beta}-\beta)' \sum_{t,s,\tau,\eta}^{T} \tilde{x}_t' \tilde{x}_s v_\tau' v_\eta (\tilde{\beta}-\beta)$;

$E_4^2 = -\frac{6}{nT(T-1)(T-2)}(\tilde{\beta}-\beta)' \sum_{t,s,\tau,\eta}^{T} \tilde{x}_t' \tilde{x}_s v_\tau' \bar{v}.(\tilde{\beta}-\beta)$;

$E_4^3 = \frac{3}{nT(T-1)}(\tilde{\beta}-\beta)' \sum_{t,s,\tau,\eta}^{T} \tilde{x}_t' \tilde{x}_s \bar{v}.' \bar{v}.(\tilde{\beta}-\beta)$.

We first consider E_4^1

$$E_4^1 = \frac{3}{nT_4}(\tilde{\beta}-\beta)' \sum_{t,s,\tau,\eta}^{T} \sum_{i=1}^{n} \tilde{x}_{it} \tilde{x}_{is} v_{i\tau} v_{i\eta} (\tilde{\beta}-\beta) + \frac{3}{nT_4}(\tilde{\beta}-\beta)' \sum_{t,s,\tau,\eta}^{T} \sum_{i\neq j}^{n} \sum_{j=1}^{n} \tilde{x}_{it} \tilde{x}_{is} v_{j\tau} v_{j\eta} (\tilde{\beta}-\beta)$$
$$= O_p(\frac{1}{nT^2\sqrt{n}}) + O_p(\frac{1}{T^2\sqrt{n}}) = O_p(\frac{1}{T^2\sqrt{n}}).$$

and then easy to show that $E_4^1 = 6D_4^2 = O_p(\frac{1}{T^2})$; $E_4^2 = 3D_4^4 = O_p(\frac{1}{T^2})$, as a result, we have

$$E_4 = O_p(\frac{1}{T^2}).$$

Proof of part e_5

$$\begin{aligned}
E_5 =& -\frac{4}{nT_4}\sum_{t,s,\tau,\eta}^{T}(\tilde{\beta}-\beta)'\tilde{x}_t'(v_s-\bar{v}.)(\tilde{\beta}-\beta)'\tilde{x}_\tau'\tilde{x}_\eta(\tilde{\beta}-\beta)\\
=& -\frac{4}{nT_4}\sum_{t,s,\tau,\eta}^{T}(\tilde{\beta}-\beta)'\tilde{x}_t'v_s(\tilde{\beta}-\beta)'\tilde{x}_\tau'\tilde{x}_\eta(\tilde{\beta}-\beta)\\
&+\frac{4}{nT(T-1)(T-2)}\sum_{t,s,\tau,\eta}^{T}(\tilde{\beta}-\beta)'\tilde{x}_t'\bar{v}.(\tilde{\beta}-\beta)'\tilde{x}_\tau'\tilde{x}_\eta(\tilde{\beta}-\beta).
\end{aligned}$$

We can write $\frac{4}{nT_4}\sum_{t,s,\tau,\eta}^{T}\tilde{x}_t'v_s\tilde{x}_\tau'\tilde{x}_\eta$ as

$$\begin{aligned}
&\frac{4}{nT_4}\sum_{t,s,\tau,\eta}^{T}\tilde{x}_t'v_s\tilde{x}_\tau'\tilde{x}_\eta\\
=&\frac{4}{nT_4}\sum_{t,s,\tau,\eta}^{T}\sum_{i=1}^{n}\tilde{x}_{it}v_{is}\tilde{x}_{i\tau}\tilde{x}_{i\eta}+\frac{4}{nT_4}\sum_{t,s,\tau,\eta}^{T}\sum_{i\neq j}^{n}\sum_{j=1}^{n}\tilde{x}_{it}v_{is}\tilde{x}_{j\tau}\tilde{x}_{j\eta}\\
=&O_p(\frac{1}{\sqrt{nT}})+O_p(\frac{\sqrt{n}}{\sqrt{T}})=O_p(\frac{\sqrt{n}}{\sqrt{T}}).
\end{aligned}$$

and since the second term $\frac{4}{nT(T-1)(T-2)}\sum_{t,s,\tau,\eta}^{T}\tilde{x}_t'\bar{v}.\tilde{x}_\tau'\tilde{x}_\eta$ equals to 2 times the second term of D_9, then

$$\frac{4}{nT(T-1)(T-2)}\sum_{t,s,\tau,\eta}^{T}\tilde{x}_t'\bar{v}.\tilde{x}_\tau'\tilde{x}_\eta=O_p(\frac{\sqrt{n}}{\sqrt{T}}).$$

With the fact that $\tilde{\beta}-\beta=O_p(\frac{1}{\sqrt{nT}})$, we can get

$$E_5=O_p(\frac{1}{nT^2}).$$

Proof of part e_6

Since $\tilde{\beta} - \beta = O_p(\frac{1}{\sqrt{nT}})$, then we have

$$
\begin{aligned}
E_6 =& - \frac{1}{nT_4} \sum_{t,s,\tau,\eta}^{T} (\tilde{\beta} - \beta)' \tilde{x}_t' \tilde{x}_s (\tilde{\beta} - \beta)(\tilde{\beta} - \beta)' \tilde{x}_\tau' \tilde{x}_\eta (\tilde{\beta} - \beta) \\
=& - O_p(\frac{1}{n^2T^2}) \frac{1}{nT_4} \sum_{t,s,\tau,\eta}^{T} \tilde{x}_t' \tilde{x}_s \tilde{x}_\tau' \tilde{x}_\eta \\
=& - O_p(\frac{1}{n^2T^2}) \frac{1}{nT_4} \sum_{t,s,\tau,\eta}^{T} \sum_{i=1}^{n} \tilde{x}_{it} \tilde{x}_{is} \tilde{x}_{i\tau} \tilde{x}_{i\eta} - O_p(\frac{1}{n^2T^2}) \frac{1}{nT_4} \sum_{t,s,\tau,\eta}^{T} \sum_{i \neq j}^{n} \sum_{j=1}^{n} \tilde{x}_{it} \tilde{x}_{is} \tilde{x}_{j\tau} \tilde{x}_{j\eta} \\
=& O_p(\frac{1}{n^2T^2}) + O_p(\frac{1}{nT^2}) = O_p(\frac{1}{nT^2}).
\end{aligned}
$$

■ ■

Essay II: Testing Cross-Sectional Dependence in Large Panel Data Models with Serial Correlation

1 Introduction

This paper studies testing for cross-sectional dependence in panel data when serial correlation is also present in the disturbances. Cross-sectional dependence could be due to unknown common shocks, spatial effects, or interactions within social networks. Ignoring cross-sectional dependence in panels can have serious consequences. In time series with serial correlation, existing cross-sectional dependence leads to efficiency loss for least squares and invalidates inference. In some cases, it results in inconsistent estimation, see Lee (2002) and Andrews (2005). Testing cross-sectional dependence of panel residuals is therefore important.

One could test for a specific form of dependence in the error like spatial correlation, see Anselin and Bera (1998) for cross-sectional data and Baltagi et al. (2003) for panel data, to mention a few. Alternatively, one could test for dependence without imposing any structure on the form of correlation among the disturbances. The null hypothesis, in that case, is testing the diagonality of the covariance or correlation matrix of the N dimensional disturbance vector $u_t = (u_{1t}, \dots, u_{Nt})'$, which is usually assumed to be independent over time, for $t = 1, \dots, T$. When N is fixed, and T is large, the traditional multivariate statistics techniques, including log-likelihood ratio and Lagrange Multiplier tests are applicable, see for example Breusch and Pagan (1980) who propose a Lagrange Multiplier (LM) test which is based on the average of the squared pair-wise correlation coefficients of the least squares residuals.

However, as N becomes large because of the growing availability of comprehensive databases in macro and finance. This so-called "high dimensional" phenomenon brings challenges to classical statistical inference. As shown in the Random Matrix Theory (RMT) literature, the sample covariance and correlation matrices are ill-conditioned since they are not consistent estimates of their population counterparts, see Johnstone (2001) and Jiang (2004). New approaches have been considered in the statistics literature for testing the diagonality of the sample covariance or correlation matrices, see Ledoit and Wolf (2002), Schott (2005) and Chen et al. (2010), to mention a few.

The above tests for raw data cannot be used directly to test cross-sectional dependence in panel data regressions since the disturbances are not observable. Noise caused by substituting residuals for the actual disturbances may accumulate due to large dimensions and this in turn may lead to biased inference. The bias for cross-sectional dependence tests in large panels depends upon the model specification, the estimation method, and the sample size N and T, among other things. For example, Pesaran et al. (2008) consider an LM test and correct its bias in a large heterogeneous panel data model; Baltagi et al. (2012) extend Schott's (2005) test to a fixed effects panel data model and correct the bias caused by estimating the disturbances with fixed effects residuals in a homogeneous panel data model. Following Ledoit and Wolf (2002), Baltagi et al. (2011) propose a bias-adjusted test for testing the null of sphericity in the fixed effects homogeneous panel data model. But this method does not test cross-sectional dependence directly. Rejection of the null could be due to cross-sectional dependence or heteroscedasticity or both. A general test for cross-sectional dependence was proposed by Pesaran (2004). His test statistic is based on the average of pair-wise correlation coefficients, defined as CD_P. The test is exactly centered at zero, under the null, and does not need bias correction. Pesaran (2015) extends his test statistic to test the null of weak cross-sectional dependence and derives its asymptotic distribution using joint limits. This test is robust to many model specifications and has many applications. Recent surveys for cross-sectional dependence tests in large panels are provided by Moscone and Tosetti (2009), Sarafidis and Wansbeek (2012) and Chudik and Pesaran (2014).

The asymptotics and bias-correction of existing tests for cross-sectional dependence in large panels are carried out under some albeit restrictive assumptions. For instance, the errors are normally distributed; $N/T \to c \in (0, \infty)$ as $(N, T) \to \infty$, and so on. One fundamental restriction is that the errors are independent over time. In fact, the presence of serial correlation in panel data applications is likely to be the rule rather than the exception, especially for macro applications and when T is large. Ignoring serial correlation does not affect the consistency of estimates, but it leads to incorrect inference. In RMT, when $u_1, u_2, \ldots, u_T$

are independent across $t = 1, 2, .., T$, and N is large, the limiting spectral distribution (LSD) of the corresponding sample covariance matrix is the Marcenko-Pastur (M-P) Law, see Bai and Silverstein (2004). Existing correlation among these disturbances may cause a deviation of the LSD from the M-P law. Indeed, Bai and Zhou (2008) show that the LSD of the sample covariance matrix with correlations in columns is different from the M-P law. Gao et al. (2014) show similar results for the sample correlation matrix. Therefore, the cross-sectional dependence tests which heavily depend upon the assumption of independence over time could lead to misleading inference if there is a serial correlation in the disturbances.

To better understand the effects of potential serial correlation on the existing tests of cross-sectional dependence, let us assume that the $T \times 1$ independent random vectors $u_i = (u_{i1}, \ldots, u_{iT})'$, for $i = 1, \ldots, N$ are observable. The correlation coefficients ρ_{ij} of any u_i and u_j $(i \neq j)$ are defined by $u_i'u_j/(\|u_i\| \cdot \|u_j\|)$. Their means are zero vectors. If all the elements of each u_i are independent and identically spherically distributed, Muirhead (1982) shows that $E\left(\rho_{ij}^2\right) = 1/T$. When N is fixed, the summation of all distinct $N(N-1)/2$ terms of ρ_{ij}^2 will be small, as $T \to \infty$. In Section 3, we show that if all the elements of each u_i follow a multiple moving average model of order 1 (MA(1)) with parameter θ, then $E\left(\rho_{ij}^2\right) = \left[1/T + \theta^2/(T + T\theta^2)\right]$. As $N \to \infty$, the extra term $\theta^2/(T + T\theta^2)$ can accumulate and lead to extra bias for the existing LM type tests in panels. Although CD_P is centered at zero, it may still encounter size distortions because serial correlation is ignored.

This paper proposes a modification of the Pesaran CD test of cross-sectional dependence when the error terms are serially correlated in a large panel data models. First, using results from RMT, we study the first two moments of the test statistic and propose an unbiased and consistent estimate of the variance with unknown serial correlation under the null. Second, we derive the limiting distribution of the test under the asymptotic framework with $(N, T) \to \infty$ simultaneously in any order without any distribution assumption. Monte Carlo simulations are conducted to study the performance of our test statistic in finite samples. The results confirm our theoretical findings.

The plan for the paper is as follows. The next section introduces the model and notation, existing LM type tests and he Cross-sectional Dependence (CD) Test. It then presents our assumptions and the proposed modified Pesarn CD test statistic. Section 3 derives the asymptotics of this test statistic. Section 4 reports the results of the Monte Carlo experiments. Section 5 provides some concluding remarks. All the mathematical proofs are provided in the Appendix.

Throughout the paper we adopt the following notation. For a squared matrix B, $\text{tr}(B)$ is the trace of B; $||B|| = (\text{tr}(B'B))^{1/2}$ denotes the Frobenius norm of a matrix or the Euclidean norm of a vector B. $\xrightarrow{d}$ denotes convergence in distribution and $\xrightarrow{p}$ denotes convergence in probability. We use $(N,T) \to \infty$ to denote the joint convergence of N and T when N and T pass to infinity simultaneously. K is a generic positive number not depending on N or T.

2 Model and Tests

Consider the following heterogeneous panel data model

$$y_{it} = \beta_i' x_{it} + u_{it}, \text{ for } i = 1, 2, \ldots, N; \ t = 1, 2, \ldots, T. \tag{2.1}$$

where i and t index the cross section dimension and time dimension respectively; y_{it} is the dependent variable and x_{it} is a $k \times 1$ vector of exogenous regressors. The individual coefficients β_i, are defined on a compact set and allowed to vary across i. The null hypothesis of no cross-sectional dependence is

$$H_0 : cov(u_{it}, u_{jt}) = 0, \text{ for all } t, \ i \neq j.$$

or equivalently as

$$H_0 : \rho_{ij} = 0, \text{ for } i \neq j. \tag{2.2}$$

where ρ_{ij} is the pair-wise correlation coefficients of the disturbances defined by

$$\rho_{ij} = \frac{\sum_{t=1}^{T} u_{it}u_{jt}}{\left(\sum_{t=1}^{T} u_{it}^2\right)^{1/2}\left(\sum_{t=1}^{T} u_{jt}^2\right)^{1/2}}.$$

Under the alternative, there exists at least one $\rho_{ij} \neq 0$, for some $i \neq j$. For the panel regression model (2.1), the residuals are unobservable. In this case, the test statistic is based on the residual-based correlation coefficients $\hat{\rho}_{ij}$. Specifically,

$$\hat{\rho}_{ij} = \frac{\sum_{t=1}^{T} e_{it}e_{jt}}{\left(\sum_{t=1}^{T} e_{it}^2\right)^{1/2}\left(\sum_{t=1}^{T} e_{jt}^2\right)^{1/2}} \tag{2.3}$$

where e_{it} is the Ordinary Least Squares (OLS) estimated using T observations for each $i = 1, 2, .., N$. These OLS residuals are given by

$$e_{it} = y_{it} - x_{it}'\hat{\beta}_i, \tag{2.4}$$

with $\hat{\beta}_i$ being the OLS estimates of β_i from (2.1) for $i = 1, 2, .., N$. Let $M_i = I_T - P_{X_i}$, where $P_{X_i} = X_i\left(X_i'X_i\right)^{-1}X_i'$, and X_i is a $T \times k$ matrix of regressors with the it-th row being the $1 \times k$ vector x_{it}'. We also define $u_i = (u_{i1}, \ldots, u_{iT})'$, $e_i = (e_{i1}, \ldots, e_{iT})'$ and $v_i = e_i/\left\|e_i\right\|$, for $i = 1, \ldots N$. The OLS residuals can be rewritten in vector form as $e_i = M_i u_i$, and the residual-based pair-wise correlation coefficients can be rewritten as $\hat{\rho}_{ij} = v_i'v_j$, for any $1 \leq i \neq j \leq N$.

2.1 LM and CD Tests

For N fixed and $T \to \infty$, Breusch and Pagan (1980) proposed an LM test to test the null of no cross-sectional dependence in (2.2) without imposing any structure on this dependence. This is given by:

$$LM_{BP} = T\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\hat{\rho}_{ij}^2. \tag{2.5}$$

LM_{BP} is asymptotically distributed as chi-squared with $N(N-1)/2$ degrees of freedom under the null. However, for a typical micro-panel data set, N is larger than T; and the Breusch-Pagan LM test statistic is not valid under this "large N, small T" setup. In fact, Pesaran (2004) proposed a scaled version of this LM test as follows:

$$LM_P = \sqrt{\frac{1}{N(N-1)}} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \left(T\hat{\rho}_{ij}^2 - 1\right). \tag{2.6}$$

Pesaran (2004) showed that LM_P is distributed as $N(0,1)$ with $T \to \infty$ first, then $N \to \infty$ under the null. However, $E\left(T\hat{\rho}_{ij}^2 - 1\right)$ is not correctly centered at zero with fixed T and large N. Hence, Pesaran et al. (2008) proposed a bias-adjusted version of this LM test, denoted by LM_{PUY}. They show that the exact mean and variance of $(T-k)\hat{\rho}_{ij}^2$ are given by

$$\mu_{Tij} = \text{E}\left[(T-k)\,\hat{\rho}_{ij}^2\right] = \frac{1}{T-k}\text{tr}\left[\text{E}(M_i M_j)\right]; \tag{2.7}$$

and

$$\nu_{Tij}^2 = \text{var}\left[(T-k)\,\hat{\rho}_{ij}^2\right] = \left\{\text{tr}^2\left[\text{E}(M_i M_j)\right]\right\} a_{1T} + 2\text{tr}\left\{\left[\text{E}(M_i M_j)\right]^2\right\} a_{2T}. \tag{2.8}$$

where $a_{1T} = a_{2T} - \frac{1}{(T-k)^2}$, and $a_{2T} = 3\left[\frac{(T-k-8)(T-k+2)}{(T-k+2)(T-k-2)(T-k-4)}\right]^2$. LM_{PUY} is given by

$$LM_{PUY} = \sqrt{\frac{2}{N(N-1)}}\frac{(T-k)\,\hat{\rho}_{ij}^2 - \mu_{Tij}}{\nu_{Tij}}. \tag{2.9}$$

Pesaran et al. (2008) show that LM_{PUY} is asymptotically distributed as $N(0,1)$ under the null (2.2) and the normality assumption of the disturbances as $T \to \infty$ followed by $N \to \infty$. Alternatively, Pesaran (2004) proposed a test based on the average of pair-wise correlation coefficients rather than their squares. The test statistic is given by

$$CD_P = \sqrt{\frac{2T}{N(N-1)}} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \hat{\rho}_{ij}. \tag{2.10}$$

Pesaran (2015) shows this test is asymptotically distributed as $N(0,1)$ with $(N,T) \to \infty$. He also extends this to test the null of weak cross-sectional dependence.

2.2 Assumptions and the Modified CD Test Statistic

So far, all the methods surveyed above for testing cross-sectional dependence in panel data models assume that the disturbances are independent over time. Ignoring serial correlation usually results in efficiency loss and biased inference. In fact, we show in Section 3, that the existence of serial correlation leads to extra bias in the LM type tests. For the CD_P test in (2.10), it is still centered at zero with serial correlation, but its variance is affected by serial correlation. As a result, we also expect size distortions in CD_P. To correct for this, we consider a modification of this test statistic that accounts for an unknown form of serial correlation in the disturbances. First, we introduce the assumptions needed:

Assumption 1 *Define $\xi_i = (\xi_{i0}, \xi_{i1}, \xi_{i2}, \dots \xi_{iT})'$ and $\varepsilon_i = (\varepsilon_{i0}, \varepsilon_{i1}, \varepsilon_{i2}, \dots \varepsilon_{iT})'$. We also assume that $\xi_i = \sigma_i \varepsilon_i$, for $i = 1, \dots, N$, where ε_i is a random vector with mean vector zero and covariance matrix I_T. Let ε_{it} denote the t-th entry of ε_i, for any $i = 1, \dots, N$. ε_{it} has uniformly bounded 4th moment and there exists a finite constant Δ such that $E(\varepsilon_{it}^4) = 3 + \Delta$. Following Bai and Zhou (2008), the disturbances $u_t = (u_{1t}, u_{2t}, \dots u_{Nt})'$ are generated by:*

$$u_t = \sum_{s=0}^{\infty} d_s \xi_{t-s}, \text{ for } t = 1, 2, \dots, T, \tag{2.11}$$

where $\xi_t = (\xi_{1t}, \xi_{2t}, \dots \xi_{Nt})'$, for $t = 0, 1, \dots, T$, are IID random vectors across time, and $\{d_s\}_{s=0}^{\infty}$ is a sequence of numbers satisfying $\sum_{s=0}^{\infty} |d_s| < K < \infty$.

Assumption 1 allows the error term u_{it} to be correlated over time. The condition $\sum_{k=1}^{\infty} |d_k| < K < \infty$ excludes long memory type strong dependence. We need bounded moment conditions to ensure large (N,T) asymptotics for panel data models with serial correlation. The conditions in Assumption 1 are quite relaxable; they are satisfied by many parametric weak dependence processes, such as stationary and invertible finite-order autoregressive and moving average (ARMA) models. Under Assumption 1, the covariance matrix of each u_i is $\Sigma_i = \sigma_i^2 \Sigma$, where Σ is a $T \times T$ symmetric positive definite matrix. The random vector u_i can be written as $u_i = \sigma_i \Gamma \varepsilon_i$, where $\Gamma\Gamma' = \Sigma$. The generic covariance matrix Σ_i of each u_i captures the serial correlation. Bai and Zhou (2008) use this representation and show that $1/T\mathrm{tr}(\Sigma^\kappa)$ is bounded for any fixed positive integer κ. More specifically, considering a multiple moving average model of order 1 (MA(1)) :

$$u_t = \xi_t + \theta \xi_{t-1}, t = 1, \ldots, T. \tag{2.12}$$

where $|\theta| < 1$ and u_t, ξ_t, u_i and ξ_i are defined in Assumption 1. For this case, $\Sigma^{\mathrm{MA}} = (\delta_{lr})_{T\times T}$, where

$$\delta_{lr} = \begin{cases} (1+\theta^2), & l = r; \\ \theta, & |l-r| = 1; \\ 0, & |l-r| > 1. \end{cases} \tag{2.13}$$

One can also verify that for (2.11), we have the following generic representation,

$$\Sigma = (\varpi_{lr})_{T\times T}, \text{where } \varpi_{lr} = \sum_{s=0}^{\infty} d_s d_{(|l-r|+s)}. \tag{2.14}$$

We use this representation throughout the paper for convenience.

Assumption 2 *The regressors, x_{it}, are strictly exogenous such that*

$$E\left(u_{it}|X_i\right) = 0, \textit{ for all } i = 1, .., N \textit{ and } t = 1, .., T, \tag{2.15}$$

and $X_i'X_i$ is a positive definite matrix.

Assumption 3 *$T > k$ and the OLS residuals, e_{it}, defined by (2.4), are not all zeros.*

Assumptions 2 and 3 are standard for model (2.1), see Pesaran (2004, 2008). We do not impose any restrictions on the distribution of the errors or the relative convergence speed of

(N,T). This framework is quite relaxable while LM type tests usually impose the normality assumption and restrictions on the relative speed of N and T, namely, $N/T \to c \in (0,\infty)$.

Under these assumptions, the OLS estimates for model (2.1) are consistent but inefficient. We focus on the term used in Pesaran's (2004) CD test:

$$T_n = \left(\frac{2}{N(N-1)}\right)^{1/2} \sum_{i=2}^{N} \sum_{j=1}^{i-1} \hat{\rho}_{ij}. \tag{2.16}$$

In the next section, we derive the first two moments of this test statistic and later derive its limiting distribution under this general unknown form of serial correlation over time.

3 Asymptotics

In this section, we study the asymptotics of the test statistic T_n defined in (2.16). To derive its limiting distribution, we first consider its first two moments.

Theorem 1 *Under Assumptions 1-3 and the null given in (2.2),*

$$E(T_n) = 0 \tag{3.1}$$

and

$$\gamma^2 = var(T_n) = \frac{2}{N(N-1)} \sum_{i=2}^{N} \sum_{j=1}^{i-1} E(\hat{\rho}_{ij}^2) = \frac{2}{N(N-1)} \sum_{i=2}^{N} \sum_{j=1}^{i-1} \frac{tr\,(M_j \Sigma M_j M_i \Sigma M_i)}{tr\,(M_i \Sigma)\, tr\,(M_j \Sigma)}. \tag{3.2}$$

where $M_i = I_T - X_i'(X_i'X_i)^{-1}X_i$, and Σ is defined by (2.14).

Theorem 1 shows that the mean of the test statistic is zero. Its variance depends on Σ, which is a generic form containing serial correlation.

In fact, as shown in the proof of Theorem 1 (see the Appendix), $\mathrm{E}\left(\hat{\rho}_{ij}^2\right) = tr\,(M_j \Sigma M_j M_i \Sigma M_i)$ / $[\mathrm{tr}\,(M_i \Sigma)\, \mathrm{tr}\,(M_j \Sigma)]$. In the special case where the error terms are independent over time,

$\Sigma = I_T$, and $\mathrm{E}\left(\hat{\rho}_{ij}^2\right)$ reduces to $\mathrm{tr}(M_j M_i)\,/(T-k)^2$, which yields the results given in equation (2.7) for the LM_{PUY} test statistic with no serial correlation. However, with serial correlation in the errors, an extra bias term is introduced in LM_{PUY} since

$$\frac{\mathrm{tr}\left(M_j \Sigma M_j M_i \Sigma M_i\right)}{\mathrm{tr}\left(M_i \Sigma\right)\mathrm{tr}\left(M_j \Sigma\right)} - \frac{\mathrm{tr}\left(M_j M_i\right)}{(T-k)^2} \neq 0, \text{ if } \Sigma \neq I_T.$$

More specifically, let us assume that u_i, $i = 1, \ldots, N$, are observable, then $\mathrm{E}\left(\rho_{ij}^2\right) = \mathrm{tr}\left(\Sigma^2\right)/\mathrm{tr}^2\left(\Sigma\right).$ For the MA(1) process defined by (2.12), $\mathrm{tr}(\Sigma^2)\,/\mathrm{tr}^2\left(\Sigma\right) = 1/T + \theta^2/(T+T\theta^2)$ and $\mathrm{tr}(\Sigma^2)\,/\mathrm{tr}^2\left(\Sigma\right) = 1/T$, for $\theta = 0$. The extra term $\theta^2/(T+T\theta^2)$ accumulates in the LM test statistic and leads to extra bias as $N \to \infty$. As discussed above, we expect that LM_{PUY} to have serious size distortions when serial correlation is present in the disturbances.

Unlike LM type tests, the test statistic T_n is centered at zero; it does not need bias adjustment. Note that if u_{it} are independent over time, our model reduces to that of Pesaran (2004). Let γ_0^2 be the variance of T_n without serial correlation, it can be written as

$$\gamma_0^2 = \frac{2}{N(N-1)} \sum_{i=2}^{N} \sum_{j=1}^{i-1} \left[\frac{T-2k}{(T-k)^2} + \frac{\mathrm{tr}(P_{X_i} P_{X_j})}{(T-k)^2} \right], \tag{3.3}$$

where $P_{X_i} = X_i \left(X_i' X_i\right)^{-1} X_i'$ and $P_{X_j} = X_j \left(X_j' X_j\right)^{-1} X_j'$. The above result is the exact variance for T_n without serial correlation; it is derived by Pesaran (2015). A modified version of CD_P is also given by Pesaran (2015) using this exact variance. From Theorem 1, γ^2 is different from γ_0^2 if $\Sigma \neq I_T$. Hence, we also expect CD_P to have size distortions when serial correlation is present in the disturbances. Next, we consider the limiting distribution of the proposed test. The result is given in the following Theorem.

Theorem 2 *Under Assumptions 1-3 and the null in (2.2), as $(N,T) \to \infty$, we have*

$$\gamma^{-1} T_n \xrightarrow{d} N\left(0,1\right). \tag{3.4}$$

Theorem 2 shows that T_n appropriately standardized is asymptotically distributed as a standard normal. It is valid for N and T tending to infinity jointly in any order. However, we do not observe Σ in a panel data regression model; and an estimate of the variance γ^2 is needed for practical applications. Following Chen and Qin (2010), an unbiased and consistent estimator of γ^2 under the null, is obtained using the cross-validation approach proposed in the following Theorem:

Theorem 3 *Let* $\hat{\gamma}^2 = \frac{1}{N(N-1)} \sum_{i=2}^{N} \sum_{j=1}^{i-1} v_i'(v_j - \bar{v}_{(i,j)}) v_j'(v_i - \bar{v}_{(i,j)})$, *where* $\bar{v}_{(i,j)} = \frac{1}{N-2} \sum_{1 \leq \tau \neq i,j \leq N} v_\tau$. *Under Assumptions 1-3 and the null in (2.2),* $E\left(\hat{\gamma}^2\right) = \gamma^2$. *As* $(N,T) \to \infty$,

$$\hat{\gamma}^2 \xrightarrow{p} \gamma^2. \tag{3.5}$$

Define $CD_R = \hat{\gamma}^{-1} T_n$. *As* $(N,T) \to \infty$,

$$CD_R \xrightarrow{d} N\left(0,1\right). \tag{3.6}$$

Theorem 3 shows that $\hat{\gamma}^2$ is a good approximation for the variance, and we do not need to specify the structure of Σ. In other words, the test statistic allows the error terms of model (2.1) to be dependent over time. Also, CD_R is a modified version of CD_P, so they are likely to perform very similarly with respect to many model specifications (see Pesaran (2004)).

4 Monte Carlo Simulations

This section conducts Monte Carlo simulations to examine the empirical size and power of the proposed test (CD_R) defined in (3.6) in heterogeneous panel data regression models. We also look at the performance of LM_{PUY} and CD_P defined by (2.9) and (2.10) respectively for comparison purposes. We consider four scenarios: (1) the errors are independent over time, with no serial correlation; (2) the errors follow a moving average model of order 1 (MA(1)) over time; (3) the errors follow an auto-regressive model of order 1 (AR(1)) over time; (4) the errors follow an auto-regressive and moving average of order (1,1) (ARMA(1, 1)) over time. Finally, we provide small sample evidence on the power performance of the modified CD_R test against a factor and spatial auto-regressive model of order one alternatives which are popular in economics for modeling cross-section dependence.

4.1 Experimental Design

Following Pesaran et al. (2008), our experiments use the following data generating process:

$$y_{it} = \alpha_i + \beta_i x_{it} + u_{it},\ i = 1, \ldots, N;\ t = 1, \ldots, T. \tag{4.1}$$

$$x_{it} = \eta x_{it-1} + \upsilon_{it}, \tag{4.2}$$

where α_i ~IID$N(1, 1)$; β_i ~IID$N(1, 0.04)$. x_{it} is a strictly exogenous regressor and we set $\eta = 0.6$ and υ_{it} ~IID$N(\phi_i^2/(1 - 0.6^2))$ with ϕ_i ~IID$\chi^2(6)/6$, for $i = 1, \ldots, N$. The error terms of (4.1) are generated using the following four data generating processes:

$$\text{(1) IID}:\ u_{it} = \xi_{it}; \tag{4.3}$$

$$\text{(2) MA(1)}:\ u_{it} = \xi_{it} + \theta\xi_{it-1}; \tag{4.4}$$

$$\text{(3) AR(1)}:\ u_{it} = \rho u_{it-1} + \xi_{it}; \tag{4.5}$$

$$\text{(4) ARMA(1,1)}:\ u_{it} = \rho u_{it-1} + \xi_{it} + \theta\xi_{it-1}; \tag{4.6}$$

where $\xi_{it} = \sigma_i\varepsilon_{it}$; σ_i^2 ~IID$\chi^2(2)/2$ and ε_{it} ~IID$(0, 1)$. We further set $\theta = 0.8$ and $\rho = 0.6$. To check the robustness of the tests to non-normal distributions, ε_{it} are generated from a Normal$(0, 1)$ and a Chi-squared distribution$(\chi^2(2)/2 - 1)$.

To examine the empirical power of the tests, we consider two different cross-sectional dependence alternatives: factor and spatial models. The factor model is generated by

$$u_{it}^* = \lambda_i f_t + u_{it}, \text{ for } i = 1, \ldots, N; \ t = 1, \ldots, T; \tag{4.7}$$

where f_t ~IID$N(0,1)$ and λ_i ~IID$U[0.1, 0.3]$; In this case, u_{it}^* replaces u_{it} in (4.1) for the power studies. u_{it} is generated by the four scenarios defined by (4.3) − (4.6), respectively. For the spatial model, we consider a first-order spatial auto-correlation model (SAR(1)),

$$u_{it}^* = \delta\left(0.5u_{i-1,t}^* + 0.5u_{i+1,t}^*\right) + u_{it}, \tag{4.8}$$

where $\delta = 0.4$ and u_{it} are defined by (4.3) − (4.6), respectively.

The experiments are conducted for $N = 10, 20, 30, 50, 100, 200$ and $T = 10, 20, 30, 50, 100$. For each pair of (N, T), we run $2,000$ replications. To obtain the empirical size, we conduct the proposed test (CD_R) and CD_P at the two-sided 5% nominal significance level and LM_{PUY} at the positive one-sided 5% nominal significance level.

4.2 Simulation Results

Table 1 reports the empirical size of CD_P, LM_{PUY} and CD_R for normal and chi-squared distributed errors. The error terms are assumed to be independent over time. The results show that all the tests have correct size with different (N, T) combinations under both normal and chi-squared scenarios. Those are consistent with the theoretical findings. The only exceptions are for small N or T equal to 10, especially for LM_{PUY}. Table 2 reports the empirical size of the three tests with MA(1) error terms defined by (4.4). The results show that CD_R has correct size for all (N, T), but CD_P has size distortions for different (N, T) combinations because the disturbances are MA(1) over time. For example, under the normality scenario, the size of CD_P is 9.35% for $N = 10$ and $T = 20$, it becomes 11.1% when T grows to 100. LM_{PUY} suffers serious size distortions, because of the extra bias caused by ignoring serial correlation. From Table 2, the empirical size of LM_{PUY} is 100% as N or T

becomes larger than 30. Tables 3 and 4 report the empirical size of the tests with AR(1) and ARMA(1,1) errors under the two distributions: normal and chi-squared scenarios. Note that CD_R is over-sized in Table 4 for the chi-squared case when $T = 10$. However, it has the correct size as T gets larger than 20. In contrast, LM_{PUY} has serious size issues, rejecting 100% of the time and CD_P is over sized under by as much as 25% . Overall, in comparison with CD_P and LM_{PUY}, the proposed test CD_R controls for size distortions when serial correlation in the disturbances is present, and is not much affected when serial correlation is not present.

Table 5 summarizes the size-adjusted power of CD_R with MA(1), AR(1) and ARMA(1,1) errors under the factor model alternative. Results show that CD_R performs reasonably well under the two distribution scenarios especially for N and $T > 10$. Table 6 confirms the power properties of CD_R for MA(1), AR(1) and ARMA(1,1) errors under the SAR(1) alternative especially for large N and T.

5 Conclusions

In this paper, we find that in the large heterogeneous panel data model, LM_{PUY} exhibit serious size bias when there is serial correlation in the disturbances. While CD_P is centered at zero, it still encounters size distortions caused by ignoring serial correlation. We modify the Pesaran CD_P test to account for serial correlation of an unknown form in the error term and call it CD_R. This paper has several novel aspects: first, an unbiased and consistent estimate of the variance under the assumptions and the null of no cross-section dependence is proposed without knowing the form of serial correlation over time. Second, the limiting distribution of the test is derived as $(N, T) \to \infty$ in any order. Third, it is distribution free. Simulations show that the proposed test CD_R successfully controls for size distortions with serial correlation in the error term. It also has reasonable power under the alternatives of a factor model and a spatial auto-correlation SAR(1) model for different serial correlation specifications.

Acknowledgments

Essay II is based on the working paper of Baltagi, Kao and Peng (2016).

References

[1] Anselin, L. 1998. *Spatial Econometrics: Methods and Models.* Dorddrecht: Kluwer Academic Publishers.

[2] Andrews, D. W. K. 2005. Cross-section Regression with Common Shocks. *Econometrica,* 73, 1551-1585.

[3] Bai, Z. D. and Silverstein, J. W. 2004. CLT for Linear Spectral Statistics of Large-Dimensional Sample Covariance Matrices. *Annals of Probability.* 32, 553-605.

[4] Bai, Z. D., and Zhou, W. 2008. Large Sample Covariance Matrices without Independence Structures in Columns. *Statistica Sinica.* 18, 425-442.

[5] Baltagi, B. H., Song, S.H. and Koh, W. 2003. Testing Panel Data Regression Models with Spatial Error Correlation. *Journal of Econometrics.* 117, 123-150.

[6] Baltagi, B. H., Feng, Q., and Kao, C. 2011. Testing for Sphericity in a Fixed Effects Panel Data Model. *The Econometrics Journal.* 14, 25-47.

[7] Baltagi, B. H., Feng, Q., and Kao, C. 2012. A Lagrange Multiplier Test for Cross-sectional Dependence in a Fixed Effects Panel Data Model. *Journal of Econometrics.* 170, 164-177.

[8] Baltagi, B. H., Kao, C., and Peng, B. 2016. Testing Cross-Sectional Dependence in Large Panel Data Models with Serial Correlation. *Working Paper.*

[9] Billingsley, P. 1995. *Probability and Measure* (3rd ed.). New York: Wiley.

[10] Breusch, T. S. and Pagan, A. R. 1980. The Lagrange Multiplier Test and Its Application to Model Specifications in Econometrics. *Review of Economic Studies*. 47, 239-253.

[11] Bushell, P. J. and Trustrum, G. B. 1990. Trace Inequality for Positive Definite Matrix Power Products. *Linear Algebra and Its Applications*. 132, 173-178.

[12] Chen, S. X., and Qin, Y. L. 2010. A Two-sample Test for High Dimensional Data with Application to Gene-Set Testing. *Annals of Statistics*. 38, 808-835.

[13] Chen, S. X., Zhang, L. X., and Zhong, P. S. 2010. Tests for High Dimensional Covariance Matrices. *Journal of American Statistical Association*. 105, 810-819.

[14] Chudik, A., and Pesaran, M. H. 2014. Large Panel Data Models with Cross-sectional Dependence: A Survey. In Baltagi, B. H., ed. *The Oxford Handbook on Panel Data*. Oxford University Press.

[15] Gao, J. T., Han, X., Pan, G. M. and Yang, Y. R. 2014. High Dimensional Correlation Matrices: CLT and Its Applications. *Working Paper*.

[16] Hsiao, C., Pesaran, M. H., Pick, A. 2012. Diagnostic Tests of Cross-Section Independence for Limited Dependent Variable Panel Data Models. *Oxford Bulletin of Economics and Statistics*. 74(2), 253-277.

[17] Ledoit, O., and Wolf, M. 2002. Some Hypothesis Tests for the Covariance Matrix When the Dimension is Large Compared to the Sample Size. *Annals of Statistics*. 41, 1055-1692.

[18] Lee, L. 2002. Consistency and Efficiency of Least Squares Estimation for Mixed Regressive, Spatial Autoregressive Models. *Econometric Theory*. 18, 252-277.

[19] Lieberman, O. 1994. A Laplace Approximation to the Moments of a Ratio of Quadratic Forms. *Biometrika*. 81, 681-690.

[20] Jiang, T. F. 2004. The Limiting Distributions of Eigenvalues of Sample Correlation Matrices. *Sankhyā*. 66, 35-48.

[21] Johnstone, I. M. 2001. On the Distribution of the Largest Eigenvalue in Principal Components Analysis. *Annals of Statistics*. 29, 295-327.

[22] Moscone, F., Tosetti, E. 2009. A Review and Comparisons of Tests of Cross-section Dependence in Panels. *Journal of Economic Surveys*. 23, 528-561.

[23] Muirhead, R. J. 1982. *Aspects of Multivariate Statistical Theory*. John Wiley & Sons.

[24] Pesaran, M. H. 2004. General Diagnostic Test for Cross Section Dependence in Panels. *CESifo Working Paper Series No.1229, IZA Discussion Paper No. 1240*. Available at SSRN: http://ssrn.com/abstract=572504.

[25] Pesaran, M. H. 2015. Testing Weak Cross-sectional Dependence in Large Panels. *Econometric Reviews*. 34, 1089-1117.

[26] Pesaran, M. H., Ullah, A., and Yamagata, T. 2008. A Bias-adjusted LM Test of Error Cross-section Independence. *The Econometrics Journal*. 11, 105-127.

[27] Sarafidis, V., Wansbeek, T. 2012. Cross-sectional Dependence in Panel Data Analysis. *Econometric Reviews*. 31, 483-531.

[28] Schott, J. R. 2005. Testing for Complete Independence in High Dimensions. *Biometrika*. 92, 951-956.

Table 1: Size of Tests with IID Errors over Time

		Normal					Chi-squared				
Tests	(N,T)	10	20	30	50	100	10	20	30	50	100
CD_R											
	10	5.75	5.90	5.50	4.75	6.45	5.90	4.80	5.55	5.15	6.45
	20	3.85	4.55	5.05	4.70	5.15	4.60	4.50	4.50	5.85	5.40
	30	4.45	4.10	4.70	5.10	4.60	4.40	4.80	4.45	4.50	6.25
	50	4.45	4.75	5.40	5.25	4.50	4.10	3.65	4.75	4.05	4.60
	100	4.65	4.85	4.20	5.65	5.30	4.35	4.80	4.70	4.35	4.95
	200	4.05	4.65	3.90	4.60	5.00	5.65	5.05	4.85	4.65	5.40
CD_P											
	10	5.60	5.50	5.25	4.10	6.00	5.60	4.70	5.05	4.70	5.65
	20	4.05	4.75	5.05	4.90	5.30	4.90	4.70	4.65	5.85	5.30
	30	4.90	4.45	4.85	5.20	5.00	5.20	5.20	4.55	5.00	6.05
	50	4.95	5.20	5.60	5.55	4.45	5.00	4.15	5.00	4.55	4.70
	100	5.65	5.15	4.50	5.95	5.45	5.15	5.65	5.05	4.50	5.05
	200	5.00	5.00	4.45	4.85	5.15	6.35	5.75	5.15	4.70	5.55
LM_{PUY}											
	10	6.75	6.05	6.10	6.00	5.60	6.60	6.85	7.65	7.95	6.60
	20	6.20	5.45	6.75	7.00	5.50	7.05	6.40	6.40	7.15	5.60
	30	6.20	6.25	5.40	6.35	5.95	7.65	5.95	6.35	5.85	7.00
	50	6.55	4.95	5.25	5.60	5.40	7.00	6.85	7.20	5.40	5.85
	100	8.10	5.45	5.40	4.60	4.55	7.00	5.85	6.10	5.85	5.90
	200	8.60	5.75	6.50	5.90	5.35	8.00	7.20	6.30	6.40	6.70

Notes: This table reports the size of CD_P, LM_{PUY} and CD_R with $u_{it} = \xi_{it}$, where $\xi_{it} = \sigma_i \varepsilon_{it}$; $\sigma_i^2 \sim \text{IID}\chi^2(2)/2$. $\varepsilon_{it} \sim \text{IID}(0,1)$ and are generated from Normal and Chi-squared distributions. The tests are conducted at the 5% nominal significance level.

Table 2: Size of Tests with MA(1) Errors

		Normal					Chi-squared				
Tests	(N,T)	10	20	30	50	100	10	20	30	50	100
CD_R											
	10	6.10	6.25	4.45	5.35	6.25	6.30	5.40	5.90	5.85	6.50
	20	5.15	4.80	5.05	4.60	5.30	5.20	5.35	4.70	6.15	4.75
	30	4.50	4.35	4.20	5.35	4.95	5.55	4.75	4.90	5.30	6.15
	50	5.25	4.50	5.30	5.70	4.30	5.00	4.65	4.60	4.35	4.85
	100	4.75	5.35	4.50	5.45	5.60	5.80	4.15	5.45	4.35	4.90
	200	4.35	4.95	3.50	4.50	4.90	6.20	6.30	4.30	4.30	5.50
CD_P											
	10	7.60	9.35	8.40	10.05	11.10	7.80	7.75	10.30	10.25	10.95
	20	6.60	8.30	9.95	9.10	10.90	7.00	8.95	9.30	10.70	10.50
	30	6.45	8.35	8.30	10.50	10.60	7.90	9.65	9.50	10.80	10.60
	50	7.45	7.95	10.75	11.30	9.65	7.55	7.90	9.20	9.70	9.15
	100	6.50	9.35	9.00	10.85	11.55	7.85	8.35	10.60	9.30	10.20
	200	6.65	8.45	8.45	9.70	10.95	9.90	9.50	9.35	9.65	11.20
LM_{PUY}											
	10	37.95	54.40	57.10	59.55	60.70	39.15	53.00	56.50	60.75	61.55
	20	81.55	96.00	96.80	98.25	97.90	83.25	95.45	97.05	97.70	98.20
	30	98.30	100.00	100.00	100.00	100.00	98.45	100.00	100.00	100.00	100.00
	50	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	100	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	200	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Notes: This table reports the size of CD_P, LM_{PUY} and CD_R with $u_{it} = \xi_{it} + \theta\xi_{it-1}$, where $\xi_{it} = \sigma_i\varepsilon_{it}$; $\sigma_i^2 \sim \text{IID}\chi^2(2)/2$. $\varepsilon_{it} \sim \text{IID}(0,1)$ and are generated from Normal and Chi-squared distributions. The tests are conducted at the 5% nominal significance level.

Table 3: Size of Tests with AR(1) Errors

		Normal					Chi-squared				
Tests	(N,T)	10	20	30	50	100	10	20	30	50	100
CD_R											
	10	6.10	6.25	4.90	6.15	6.75	6.05	4.80	6.10	6.00	5.65
	20	4.75	5.65	4.65	4.70	5.00	4.85	5.60	4.50	5.55	4.80
	30	4.15	4.85	4.00	4.55	4.65	5.50	4.25	5.75	5.10	6.65
	50	4.15	4.50	5.20	5.45	4.40	5.25	5.35	4.60	4.40	4.35
	100	4.35	4.80	4.80	5.45	4.80	5.75	4.15	5.30	4.05	5.10
	200	4.85	4.60	4.05	4.55	5.05	7.80	5.35	4.95	4.20	4.55
CD_P											
	10	6.80	9.65	10.20	14.55	16.80	6.55	8.25	12.25	13.90	16.30
	20	5.75	9.50	11.35	13.25	16.85	5.90	9.60	11.50	15.05	15.45
	30	5.65	9.80	10.00	13.30	14.05	7.35	9.65	12.00	15.20	17.15
	50	5.90	8.45	11.95	14.80	14.10	7.10	9.55	9.70	12.40	15.80
	100	6.05	10.00	10.40	14.70	16.55	7.25	8.70	12.25	13.85	15.00
	200	6.65	9.00	10.25	13.30	16.70	9.40	10.3	10.85	13.70	16.10
LM_{PUY}											
	10	37.95	54.40	57.10	59.55	60.70	27.60	66.30	82.45	90.80	95.35
	20	55.50	97.90	99.85	100.00	100.00	59.95	98.40	99.85	100.00	100.00
	30	98.30	99.95	100.00	100.00	100.00	82.75	100.00	100.00	100.00	100.00
	50	97.80	100.00	100.00	100.00	100.00	98.60	100.00	100.00	100.00	100.00
	100	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	200	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Notes: This table reports the size of CD_P, LM_{PUY} and CD_R with $u_{it} = \rho u_{it-1} + \xi_{it}$, where $\xi_{it} = \sigma_i \varepsilon_{it}$; $\sigma_i^2 \sim \text{IID}\chi^2(2)/2$. $\varepsilon_{it} \sim \text{IID}(0, 1)$ and are generated from Normal and Chi-squared distributions. The tests are conducted at the 5% nominal significance level.

Table 4: Size of Tests with ARMA(1,1) Errors

		Normal					Chi-squared				
Tests	(N,T)	10	20	30	50	100	10	20	30	50	100
CD_R											
	10	6.95	6.45	4.90	6.20	5.85	7.20	5.25	6.40	5.40	5.45
	20	5.40	5.55	4.95	4.75	4.95	6.40	5.70	4.95	5.55	4.70
	30	4.65	4.75	4.05	4.80	4.65	7.45	4.60	5.95	5.10	6.50
	50	4.95	4.95	5.25	5.30	4.50	7.50	5.70	4.80	4.35	4.80
	100	5.05	5.15	4.60	5.10	4.90	10.25	5.10	4.65	4.00	4.80
	200	5.75	4.65	4.45	4.85	5.20	17.45	6.60	5.75	4.50	4.25
CD_P											
	10	9.10	15.95	16.35	22.50	24.30	10.95	13.80	19.20	21.70	25.15
	20	8.30	14.40	17.80	20.15	25.05	10.10	14.80	18.90	22.85	23.15
	30	8.30	15.40	17.70	21.55	22.55	10.95	15.25	19.25	23.55	24.25
	50	8.70	14.85	18.80	22.70	23.40	11.75	15.40	17.30	19.15	23.95
	100	9.35	15.90	17.50	22.15	24.20	17.20	14.45	17.95	22.05	22.70
	200	9.50	14.05	18.35	20.00	24.95	25.45	17.00	18.55	21.35	24.65
LM_{PUY}											
	10	83.65	98.45	99.45	99.75	99.80	83.65	98.40	99.70	99.90	100.00
	20	99.85	100.00	100.00	100.00	100.00	99.85	100.00	100.00	100.00	100.00
	30	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	50	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	100	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	200	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Notes: This table reports the size of CD_P, LM_{PUY} and CD_R with $u_{it} = \rho u_{it-1} + \xi_{it} + \theta\xi_{it-1}$, where $\xi_{it} = \sigma_i \varepsilon_{it}$; $\sigma_i^2 \sim \text{IID}\chi^2(2)/2$. $\varepsilon_{it} \sim \text{IID}(0,1)$ and are generated from Normal and Chi-squared distributions. The tests are conducted at the 5% nominal significance level.

Table 5: Size Adjusted Power of CD_R : Factor Model

		Normal					Chi-squared				
DGP	(N,T)	10	20	30	50	100	10	20	30	50	100
$MA(1)$											
	10	14.55	23.95	30.30	45.40	63.05	21.95	30.75	33.65	46.00	66.10
	20	35.70	56.65	68.95	84.05	95.95	47.30	63.25	75.80	86.00	97.40
	30	59.65	81.70	91.75	97.65	99.95	69.75	87.50	92.60	98.00	99.95
	50	83.65	96.60	99.30	100.00	100.00	88.75	98.00	99.55	100.00	100.00
	100	96.75	99.95	100.00	100.00	100.00	98.90	99.90	100.00	100.00	100.00
	200	99.70	100.00	100.00	100.00	100.00	99.70	100.00	100.00	100.00	100.00
$AR(1)$											
	10	18.95	23.95	32.40	38.10	56.75	26.95	35.00	28.90	37.15	61.25
	20	45.60	62.10	69.95	81.45	94.20	55.10	67.45	74.85	85.65	96.60
	30	68.80	83.50	92.30	97.60	99.75	78.15	90.85	92.70	97.40	99.85
	50	88.55	97.45	99.40	100.00	100.00	92.90	98.50	99.65	100.00	100.00
	100	98.80	100.00	100.00	100.00	100.00	99.60	99.95	100.00	100.00	100.00
	200	99.90	100.00	100.00	100.00	100.00	99.85	100.00	100.00	100.00	100.00
$ARMA(1,1)$											
	10	7.70	7.70	10.00	10.80	14.80	9.65	10.35	8.80	9.60	19.60
	20	22.05	18.85	24.25	27.80	39.50	24.85	22.35	23.40	30.60	46.20
	30	37.75	37.45	46.15	48.90	75.00	41.75	47.35	44.15	53.15	71.25
	50	66.50	66.75	71.60	83.10	96.20	66.25	72.35	82.45	88.20	98.00
	100	91.15	96.60	98.75	99.90	100.00	90.45	98.55	99.40	99.95	100.00
	200	98.95	100.00	100.00	100.00	100.00	98.45	99.95	100.00	100.00	100.00

Notes: This table computes the size adjusted power for CD_R with a factor model that allows for cross-sectional dependence in the errors: $u^*_{it} = \lambda_i f_t + u_{it}$. u_{it} are generated by MA(1), AR(1) and ARMA (1,1) defined by (4.4)-(4.6). $\xi_{it} = \sigma_i \varepsilon_{it}$; $\sigma_i^2 \sim \text{IID}\chi^2(2)/2$. $\varepsilon_{it} \sim \text{IID}(0,1)$ and are generated from Normal and Chi-squared distributions.

Table 6: Size Adjusted Power of CD_R: SAR(1) Model

		Normal					Chi-squared				
DGP	(N,T)	10	20	30	50	100	10	20	30	50	100
$MA(1)$											
	10	38.85	60.55	72.20	88.25	97.30	43.05	67.15	72.55	88.45	97.70
	20	37.45	61.70	76.00	92.15	99.05	39.25	61.25	76.80	89.55	99.10
	30	39.60	64.55	78.60	92.00	99.60	40.30	65.65	78.80	91.90	99.35
	50	40.05	66.45	79.15	92.70	99.75	39.95	66.55	78.65	94.65	99.70
	100	33.60	62.70	80.55	92.55	99.65	37.85	64.65	79.20	94.40	99.90
	200	40.65	64.50	80.65	94.70	99.8	37.75	62.50	81.25	95.65	99.80
$AR(1)$											
	10	37.20	53.95	68.20	79.20	92.10	42.85	63.20	61.15	78.00	94.80
	20	38.25	56.50	69.30	82.90	95.85	38.55	55.50	68.65	83.70	97.20
	30	37.90	56.90	71.80	84.65	98.10	38.70	62.00	66.25	85.70	96.90
	50	38.80	59.80	71.40	86.60	98.60	39.70	59.15	71.25	89.00	99.00
	100	38.85	57.85	70.90	86.60	98.75	35.25	59.85	72.55	88.95	98.60
	200	40.75	55.95	74.40	87.75	98.80	33.80	56.00	70.85	90.40	99.10
$ARMA(1,1)$											
	10	29.00	43.40	58.05	70.20	85.90	32.75	49.75	51.30	67.40	88.20
	20	31.05	43.55	56.65	72.10	89.10	28.35	43.45	54.80	71.35	91.35
	30	30.00	45.70	59.35	71.35	94.20	28.10	48.10	54.00	73.05	91.90
	50	33.05	45.30	54.40	71.70	93.30	27.30	43.90	58.00	75.75	94.45
	100	30.60	45.15	55.50	75.40	94.95	21.80	45.45	57.85	77.35	94.75
	200	30.30	42.05	58.15	75.75	95.15	21.05	38.80	55.70	77.50	95.80

Notes: This table computes the size adjusted power for CD_R with a SAR(1) model that allows for cross-sectional dependence in the error: $u_{it}^* = \delta(0.5u_{i-1,t}^* + 0.5u_{i+1,t}^*) + u_{it}$ with $\delta = 0.4$. u_{it} are generated by IID, MA(1), AR(1) and ARMA (1,1) defined by (4.3)-(4.6). $\xi_{it} = \sigma_i \varepsilon_{it}$; $\sigma_i^2 \sim \text{IID}\chi^2(2)/2$. $\varepsilon_{it} \sim \text{IID}(0,1)$ and are generated from Normal and Chi-squared distributions.

Appendix

This appendix includes proofs of the main results in the text. The appendix includes two parts: Part A includes some useful lemmas which are frequently used in the proofs of Theorems; Part B gives the proofs of all the theorems included in the paper.

Let us introduce some notation before proceeding: For two matrices $B = (b_{ij})$ and $C = (c_{ij})$, we define $B \circ C = (b_{ij}c_{ij})$. $\sum$ denotes summation over mutually different indices, e.g., $\sum\limits_{(i_1,i_2,j_1,j_2)}$ means summation over $\{(i_1,i_2,j_1,j_2) : i_1,i_2,j_1,j_2 \text{ are mutually different.}\}$.

A Some Useful Lemmas

Lemma A.1 *Let F and G be non-stochastic $N\times N$ symmetric and positive definite matrices. Define $r = \frac{u_i' F u_i}{u_i' G u_i}$. Under Assumptions 1, we have*

(a) $E(r^k) = \frac{E\left[\left(\varepsilon_i' F\varepsilon_i\right)^k\right]}{\left[E\left(\varepsilon_i' G\varepsilon_i\right)\right]^k}$;

(b) $E(\varepsilon_i' F\varepsilon_i) = tr(F)$;

(c) $E(\varepsilon_i' F\varepsilon_i)^2 = tr(F^2) + 2tr^2(F) + \Delta tr(F\circ F)$;

(d) $tr(F\circ F) \le tr(F^2)$.

The Proof of part (a) is given by Lieberman (1994) and the proof of (b)-(d) are from Proposition 1 of Chen, et al. (2010), hence we omit the proof here.

Lemma A.2 *Define $B_j = M_j\Sigma M_j$, for any j, respectively. Under Assumptions 1-3 and the null in (2.2), we have*

(a) $E\left(\hat\rho_{ij}^2\right) = \frac{tr(B_iB_j)}{tr(B_i)tr(B_j)}$;

(b) $E\left(\hat\rho_{ij}^4\right) \le (3+\Delta)\,\frac{(2+\Delta)tr(B_iB_j)^2 + tr^2(B_iB_j)}{tr^2(B_i)tr^2(B_j)}$;

(c) For any $j_1 \ne j_2$, $E\left(\hat\rho_{ij_1}^2\hat\rho_{ij_2}^2\right) \le \frac{\left((2+\Delta)tr\left(B_iB_{j_1}\right)^2 + tr^2\left(B_iB_{j_1}\right)\right)^{1/2}\left((2+\Delta)tr\left(B_iB_{j_2}\right)^2 + tr^2\left(B_iB_{j_2}\right)\right)^{1/2}}{tr\left(B_{j_1}\right)tr(B_{j_2})tr^2(B_i)}$.

Proof. Recall that the pair-wise correlation coefficients is defined as

$$\hat\rho_{ij} = v_i' v_j = \sum_{t=1}^{T} v_{it}v_{jt},$$

where v_i are the scaled residual vector defined by $v_i = \frac{e_i}{(e_i' e_i)^{1/2}}$. e_i is the OLS residual vector from the individual-specific least squares regression and it is given by

$$e_i = M_i u_i = M_i \sigma_i \Gamma \varepsilon_i, \text{ with } M_i = I_T - P_{X_i} = I_T - X_i \left(X_i' X_i\right)^{-1} X_i'.$$

where M_i is idempotent. Consider part (a),

$$\mathrm{E}\left(\hat{\rho}_{ij}^2\right) = \mathrm{E}\left(v_i' v_j\right)^2 = \mathrm{E}\left(\frac{e_i' e_j}{\left(e_i' e_i\right)^{1/2}\left(e_j' e_j\right)^{1/2}}\right)^2 = \mathrm{E}\left(\frac{e_i' A_j e_i}{e_i' e_i}\right).$$

where $A_j = \frac{e_j e_j'}{e_j' e_j}$. Then

$$\mathrm{E}\left(\hat{\rho}_{ij}^2\right) = \mathrm{E}\left[\mathrm{E}\left(\hat{\rho}_{ij}^2 | \varepsilon_j\right)\right] = \mathrm{E}\left[\mathrm{E}\left(\frac{e_i' A_j e_i}{e_i' e_i} | \varepsilon_j\right)\right].$$

Since $e_i = M_i \sigma_i \Gamma \varepsilon_i$, and using part (a) and (b) of Lemma A.1, we have $\mathrm{E}\left(\frac{e_i' A_j e_i}{e_i' e_i} | \varepsilon_j\right) = \frac{\mathrm{tr}(\Gamma' M_i A_j M_i \Gamma)}{\mathrm{tr}(\Gamma' M_i \Gamma)}$. Moreover,

$$\begin{aligned}
\mathrm{E}\left[\mathrm{tr}\left(\Gamma' M_i A_j M_i \Gamma\right)\right] &= \mathrm{E}\left(\frac{\varepsilon_j' \Gamma' M_j M_i \Gamma \Gamma' M_i M_j \Gamma \varepsilon_j}{\varepsilon_j' \Gamma' M_j \Gamma \varepsilon_j}\right) \\
&= \frac{\mathrm{tr}\left(\Gamma' M_j M_i \Gamma \Gamma' M_i M_j \Gamma\right)}{\mathrm{tr}\left(\Gamma' M_j \Gamma\right)} \\
&= \frac{\mathrm{tr}\left(M_j \Sigma M_j M_i \Sigma M_i\right)}{\mathrm{tr}\left(M_j \Sigma\right)}.
\end{aligned}$$

Together with the above results, we have

$$\mathrm{E}\left(\hat{\rho}_{ij}^2\right) = \frac{\mathrm{tr}\left(M_j \Sigma M_j M_i \Sigma M_i\right)}{\mathrm{tr}\left(M_i \Sigma\right) \mathrm{tr}\left(M_j \Sigma\right)} = \frac{\mathrm{tr}\left(B_i B_j\right)}{\mathrm{tr}\left(B_i\right) \mathrm{tr}\left(B_j\right)}.$$

Consider part (b).

$$\mathrm{E}\left(\rho_{ij}^4\right) = \mathrm{E}\left[\mathrm{E}\left(\rho_{ij}^4 \mid v_j\right)\right] = \mathrm{E}\left(\mathrm{E}\left[\left(\left(\frac{e_i'A_je_i}{e_i'e_i}\right)^2 \mid v_j\right)\right]\right) = \mathrm{E}\left[\frac{\mathrm{E}(\varepsilon_i'\Gamma'M_iA_jM_i\Gamma\varepsilon_i)^2}{\mathrm{tr}^2\left(\Gamma'M_i\Gamma\right)} \mid v_j\right]$$
$$= \mathrm{E}\left[\frac{2\mathrm{tr}\left(\Gamma'M_iA_jM_i\Gamma\right)^2 + \mathrm{tr}^2\left(\Gamma'M_iA_jM_i\Gamma\right) + \Delta\mathrm{tr}\left(\Gamma'M_iA_jM_i\Gamma \circ \Gamma'M_iA_jM_i\Gamma\right)}{\mathrm{tr}^2(B_i)}\right].$$

Using part (a) of Lemma A.1, we have

$$\mathrm{E}\left[\mathrm{tr}^2\left(\Gamma'M_iA_jM_i\Gamma\right)\right] = \mathrm{E}\left(\frac{\varepsilon_j'\Gamma'M_jM_i\Gamma\Gamma'M_iM_j\Gamma\varepsilon_j}{\varepsilon_j'\Gamma'M_j\Gamma\varepsilon_j}\right)^2 = \frac{\mathrm{E}\left(\varepsilon_j'\Gamma'M_jM_i\Gamma\Gamma'M_iM_j\Gamma\varepsilon_j\right)^2}{\left[\mathrm{E}\left(\varepsilon_j'\Gamma'M_j\Gamma\varepsilon_j\right)\right]^2}.$$

Using part (c) of Lemma A.1, we also have

$$\mathrm{E}\left(\varepsilon_j'\Gamma'M_jM_i\Gamma\Gamma'M_iM_j\Gamma\varepsilon_j\right)^2 = 2\mathrm{tr}(\Gamma'M_jM_i\Gamma\Gamma'M_iM_j\Gamma)^2 + \mathrm{tr}^2(\Gamma'M_jM_i\Gamma\Gamma'M_iM_j\Gamma)$$
$$+ \Delta\mathrm{tr}\left(\Gamma'M_jM_i\Gamma\Gamma'M_iM_j\Gamma \circ \Gamma'M_jM_i\Gamma\Gamma'M_iM_j\Gamma\right)$$
$$= 2\mathrm{tr}\left(B_iB_j\right)^2 + \mathrm{tr}^2\left(B_iB_j\right) + \Delta\mathrm{tr}\left(\Gamma'M_jM_i\Gamma\Gamma'M_iM_j\Gamma \circ \Gamma'M_jM_i\Gamma\Gamma'M_iM_j\Gamma\right)$$
$$\leq (2+\Delta)\,\mathrm{tr}\left(B_iB_j\right)^2 + \mathrm{tr}^2\left(B_iB_j\right).$$

With the fact that $\mathrm{E}\big(\varepsilon_j'\Gamma'M_j\Gamma\varepsilon_j\big) = \mathrm{tr}(B_j)$, we obtain

$$\mathrm{E}\left[\mathrm{tr}^2\left(\Gamma'M_iA_jM_i\Gamma\right)\right] \leq \frac{(2+\Delta)\,\mathrm{tr}\left(B_iB_j\right)^2 + \mathrm{tr}^2\left(B_iB_j\right)}{\mathrm{tr}^2(B_j)}.$$

Next we consider $\mathrm{E}\big[\mathrm{tr}\left(\Gamma'M_iA_jM_i\Gamma\right)^2\big]$.

$$\mathrm{E}\left[\mathrm{tr}\left(\Gamma'M_iA_jM_i\Gamma\right)^2\right] = \mathrm{E}\left[\left(\frac{\varepsilon_j'\Gamma'M_jM_i\Gamma\Gamma'M_iM_j\Gamma\varepsilon_j}{\varepsilon_j'\Gamma'M_j\Gamma\varepsilon_j}\right)^2\right]$$
$$\leq \frac{(2+\Delta)\,\mathrm{tr}\left(B_iB_j\right)^2 + \mathrm{tr}^2\left(B_iB_j\right)}{\mathrm{tr}^2(B_j)}.$$

Hence,

$$E\left(\hat{\rho}_{ij}^4\right) \leq (3+\Delta)\,\frac{(2+\Delta)\,\mathrm{tr}\left(B_iB_j\right)^2 + \mathrm{tr}^2\left(B_iB_j\right)}{\mathrm{tr}^2(B_i)\mathrm{tr}^2(B_j)}.$$

Consider part (c), since

$$\begin{aligned}\mathrm{E}\left(\hat{\rho}_{ij_1}^2\hat{\rho}_{ij_2}^2\right) &= \mathrm{EE}\left(\hat{\rho}_{ij_1}^2\hat{\rho}_{ij_2}^2|v_i\right) = \mathrm{E}\left(\mathrm{E}\left(\hat{\rho}_{ij_1}^2|v_i\right)\mathrm{E}\left(\hat{\rho}_{ij_2}^2|v_i\right)\right)\\ &= \frac{\mathrm{E}\left(v_i'B_{j_1}v_iv_i'B_{j_2}v_i\right)}{\mathrm{tr}\left(B_{j_1}\right)\mathrm{tr}(B_{j_2})}.\end{aligned}$$

Note that $|\mathrm{E}\left(v_i'B_{j_1}v_iv_i'B_{j_2}v_i\right)| \leq \left[\mathrm{E}\left(v_i'B_{j_1}v_i\right)^2\right]^{1/2}\left[\mathrm{E}\left(v_i'B_{j_2}v_i\right)^2\right]^{1/2}$ by using Cauchy-Schwarz inequality and

$$\mathrm{E}\left(v_i'B_{j_1}v_i\right)^2 = \mathrm{E}\left(\frac{\varepsilon_i'\Gamma'M_iM_{j_1}\Gamma\Gamma'M_{j_1}M_i\Gamma\varepsilon_i}{\varepsilon_i'\Gamma'M_i\Gamma\varepsilon_i}\right)^2 \leq \frac{(2+\Delta)\,\mathrm{tr}\left(B_iB_{j_1}\right)^2 + \mathrm{tr}^2\left(B_iB_{j_1}\right)}{\mathrm{tr}^2(B_i)}.$$

Hence

$$\mathrm{E}\left(\hat{\rho}_{ij_1}^2\hat{\rho}_{ij_2}^2\right) \leq \frac{\left((2+\Delta)\,\mathrm{tr}\left(B_iB_{j_1}\right)^2 + \mathrm{tr}^2\left(B_iB_{j_1}\right)\right)^{1/2}\left((2+\Delta)\,\mathrm{tr}\left(B_iB_{j_2}\right)^2 + \mathrm{tr}^2\left(B_iB_{j_2}\right)\right)^{1/2}}{\mathrm{tr}\left(B_{j_1}\right)\mathrm{tr}(B_{j_2})\mathrm{tr}^2(B_i)}.$$

■

Lemma A.3 *Under Assumptions 1-3 and the null in (2.2), for any fixed positive number k, we have*

(a) $\frac{1}{T}tr\left(\Sigma^k\right) = O(1)$;

(b) $\frac{1}{T}tr(B_i^k) = O(1)$;

(c) $\frac{1}{T}tr(B_{i_1}B_{i_2}\cdots B_{i_k}) = O(1)$, *for* $i_1 \neq i_2 \neq \cdots \neq i_k$.

Proof. Part (a) is directly from Bai and Zhou (2008), hence we omit it here. Next we consider part (b). Since $I_T - P_{X_i}$ is idempotent, for any $i = 1, 2, \ldots, N$; hence, $\mathrm{tr}\left(B_i^k\right) = \mathrm{tr}[(I_T - P_{X_i})\,\Sigma(I_T - P_{X_i})]^k = \mathrm{tr}\left([(I_T - P_{X_i})\,\Sigma]^k\right)$, By using the inequality that for any positive definite matrices A and B (see Bushell and Trustrum (1990)):

$$\mathrm{tr}\left(AB\right)^k \leq \mathrm{tr}\left(A^kB^k\right).$$

we have

$$\mathrm{tr}\left(B_i^k\right) \leq \mathrm{tr}\left(\left(I_T - P_{X_i}\right)\Sigma^k\right) \leq \mathrm{tr}\left(\Sigma^k\right).$$

Using part (a), then

$$\frac{1}{T}\mathrm{tr}\left(B_i^k\right) \leq \frac{1}{T}\mathrm{tr}\left(\Sigma^k\right) = O(1).$$

For part (c), since for each B_{i_l}, $l = 1, \cdots, k$, it is positive semi-definite. We also have $B_{i_l} \leq \Sigma$, $l = 1, \cdots, k$. By using the facts that for any matrices A, B, with $A \leq B$ and C positive definite, $\mathrm{tr}(AC) \leq \mathrm{tr}(BC)$, we conclude that

$$\frac{1}{T}\mathrm{tr}(B_{i_1}B_{i_2}\cdots B_{i_k}) \leq \frac{1}{T}\mathrm{tr}\left(\Sigma^k\right) = O(1).$$

Part (c) holds. ■

B Proof of Theorems

B.1 Proof of Theorem 1

Proof. Since $\mathrm{E}(e_i|X_i) = 0$, and $\varepsilon_i, i = 1, 2 \ldots, N$, are independent, it is easy to show that

$$\mathrm{E}(\hat{\rho}_{ij}) = 0,$$

which further implies $\mathrm{E}(T_n) = 0$. Next we consider the variance of T_n.

$$\mathrm{var}\left(\sum_{i=1}^{N}\sum_{j=1}^{i-1}\hat{\rho}_{ij}\right) = \mathrm{E}\left(\sum_{i=1}^{N}\sum_{j=1}^{i-1}\hat{\rho}_{ij}\right)^2 = \mathrm{E}\left(\sum_{i_1=1}^{N}\sum_{j_1=1}^{i_1-1}\sum_{i_2=1}^{N}\sum_{j_2=1}^{i_2-1}\hat{\rho}_{i_1j_1}\hat{\rho}_{i_2j_2}\right).$$

To calculate the above term, we have 3 cases to discuss:

(1) i_1, i_2, j_1, j_2 are mutually different. $\mathrm{E}(\hat{\rho}_{i_1j_1}\hat{\rho}_{i_2j_2}) = 0$.

(2) $i_1 = i_2$, $j_1 = j_2$. By using Lemma 3, we have $\mathrm{E}\big(\hat{\rho}_{ij}^2\big) = \frac{\mathrm{tr}(B_iB_j)}{\mathrm{tr}(B_i)\mathrm{tr}(B_j)}$.

(3) $i_1 = i_2$, $i_1 \neq j_1 \neq j_2$. Since $v_{i_1}, v_{j_1}, v_{i_1}$ and v_{j_2} are independent, we have $E\left(\hat{\rho}_{i_1j_1}\hat{\rho}_{i_1j_2}\right) = E\left(v_{i_1}'v_{j_1}v_{i_1}'v_{j_2}\right) = 0$.

Hence, the above results give us the variance of T_n, which is

$$\begin{aligned}\gamma^2 = \mathrm{var}\left(T_n\right) &= \frac{2}{N(N-1)}\sum_{i=1}^{N}\sum_{j=1, j\neq i}^{N}\frac{\mathrm{tr}\left(M_j\Sigma M_jM_i\Sigma M_i\right)}{\mathrm{tr}\left(M_i\Sigma\right)\mathrm{tr}\left(M_j\Sigma\right)} \\ &= \frac{2}{N(N-1)}\sum_{i=2}^{N}\sum_{j=1}^{i-1}\frac{\mathrm{tr}\left(B_iB_j\right)}{\mathrm{tr}\left(B_i\right)\mathrm{tr}\left(B_j\right)}.\end{aligned}$$

and Theorem 1 is proved. ■

B.2 Proof of Theorem 2

Proof. To prove this theorem, we need to employ the martingale central limit theorem (Billingsley (1995)). For that purpose, we define $\mathcal{F}_0 = \{\phi, \Omega\}$, $\mathcal{F}_{Ni}$ as the the σ-field generated by $\{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_i\}$ for $1 \le i \le N$. Let $E_{Nr}(\cdot)$ denote the conditional expectation given filtration $\mathcal{F}_{Nr}[E_0(\cdot) = E(\cdot)]$. Write $L_n = \sum_{i=1}^{N} D_{N,i}$ with $D_{N,1} = 0$. More specifically,

$$D_{N,i} = \left(\frac{1}{N(N-1)}\right)^{1/2} \sum_{j=1}^{i-1} v_i' v_j.$$

For every N, we can further show that

$$\mathrm{E}\left(D_{N,i} \mid \mathcal{F}_{N,i-1}\right) = 0.$$

Hence, $D_{N,i}\ (1 \le i \le N)$ is a martingale difference sequence with respect to $\mathcal{F}_{N,i}\ (1 \le i \le N)$. Let $\delta_{Ni}^2 = \mathrm{E}\left[(D_{Ni})^2 \mid \mathcal{F}_{N,i-1}\right]$. By applying the Martingale Central limit theorem, it is sufficient to show that, as $(N,T) \longrightarrow \infty$,

$$\frac{\sum_{i=1}^{N} \delta_{Ni}^2}{\mathrm{var}\,(T_n)} \xrightarrow{p} 1 \qquad \text{and} \qquad \frac{\sum_{i=1}^{N} \mathrm{E}\left(D_{N,i}^4\right)}{\mathrm{var}^2\,(T_n)} \longrightarrow 0.$$

Lemma B.1 and B.2 prove the above conditions. Hence, we can apply the Martingale Central Limit Theorem and as $(N,T) \longrightarrow \infty$, we have

$$\gamma^{-1} T_n \xrightarrow{d} N(0,1).$$

■

Lemma B.1 *Under Assumptions 1-3 and the null (2.2), as* $(N,T) \to \infty$,

$$\frac{\sum_{i=1}^{N} \delta_{Ni}^2}{var\,(T_n)} \xrightarrow{p} 1,$$

where $\delta_{Ni}^2 = E\left[(D_{Ni})^2 \mid \mathcal{F}_{N,i-1}\right]$.

Proof. To prove Lemma B.1, we first show that $\mathrm{E}\left(\sum_{i=1}^{N} \delta_{Ni}^2\right)$ =var(T_n). Then we will show that as $(N,T) \to \infty$, $\mathrm{var}\left(\sum_{i=1}^{N} \delta_{Ni}^2\right)/\mathrm{var}^2(T_n) \to 0$. It is easy to show that

$$\mathrm{E}\left(\sum_{i=1}^{N} \delta_{Ni}^2\right) = \sum_{i=1}^{N} \mathrm{E}\left\{\mathrm{E}\left[(D_{Ni})^2 \mid \mathcal{F}_{N,i-1}\right]\right\} = \mathrm{var}\left(T_n\right).$$

Next, we only need to show that the second condition is satisfied. We first consider the magnitude of var(T_n). From Lemma A.3, we know that

$$\frac{\mathrm{tr}\left(B_j B_i\right)}{\mathrm{tr}\left(B_i\right)\mathrm{tr}\left(B_j\right)} = O(T^{-1}),$$

which implies $\mathrm{var}^2(T_n) = O(T^{-2})$. Now, consider $\mathrm{var}(\sum_{i=1}^{N} \delta_{Ni}^2)$. Let $Q_j = \sum_{j=1}^{i-1} v_j$, then

$$\begin{aligned}
\delta_{Ni}^2 &= \mathrm{E}\left[(D_{Ni})^2 \mid \mathcal{F}_{N,i-1}\right] \\
&= \frac{2}{N(N-1)}\mathrm{E}\left(v_i' Q_j Q_j' v_i \mid \mathcal{F}_{N,i-1}\right) \\
&= \frac{2}{N(N-1)}\mathrm{E}\left(\frac{\varepsilon_i'\Gamma' M_i Q_j Q_j' M_i \Gamma \varepsilon_i}{(\varepsilon_i' M_i \Gamma'\Gamma M_i \varepsilon_i)} \mid \mathcal{F}_{N,i-1}\right) \\
&= \frac{2}{N(N-1)}\frac{\left(Q_j' M_i \Gamma\Gamma' M_i Q_j\right)}{\mathrm{tr}\left(B_i\right)}.
\end{aligned}$$

Therefore, we need to show the magnitude of $\mathrm{var}\left(\sum_{i=1}^{N} Q_j' M_i \Gamma\Gamma' M_i Q_j\right)$. Rewrite $Q_j' M_i \Gamma\Gamma' M_i Q_j = \sum_{j_1=1}^{i-1}\sum_{j_2=1}^{i-1} v_{j_1}' B_i v_{j_2}$, and

$$\mathrm{E}\left(\sum_{j_1=1}^{i-1}\sum_{j_2=1}^{i-1} v_{j_1}' M_i \Gamma\Gamma' M_i v_{j_2}\right) = \mathrm{E}\left(\sum_{j=1}^{i-1} v_j' B_i v_j\right) = \sum_{j=1}^{i-1} \mathrm{E}\left[\frac{\varepsilon_j'\Gamma' M_j B_i M_j \Gamma \varepsilon_j}{(\varepsilon_j'\Gamma' M_j \Gamma \varepsilon_j)}\right] = \sum_{j=1}^{i-1} \frac{\mathrm{tr}\left(B_j B_i\right)}{\mathrm{tr}\left(B_j\right)}.$$

Next we consider $\mathrm{E}\left(\sum_{j_1=1}^{i-1}\sum_{j_2=1}^{i-1} v_{j_1}' B_i v_{j_2}\right)^2$.

$$\mathrm{E}\left(\sum_{j_1=1}^{i-1}\sum_{j_2=1}^{i-1} v_{j_1}' B_i v_{j_2}\right)^2 = \mathrm{E}\sum_{j_1=1}^{i-1}\sum_{j_2=1}^{i-1}\sum_{j_3=1}^{i-1}\sum_{j_4=1}^{i-1}\left(v_{j_1}' B_i v_{j_2} v_{j_3}' B_i v_{j_4}\right).$$

To calculate magnitude order of the above term, we have 3 cases to discuss.

(1) $j_1 = j_2 = j_3 = j_4 = j$.

$$\begin{aligned}\mathrm{E}\left(v_j' B_i v_j\right)^2 &= \mathrm{E}\frac{\left(\varepsilon_j'\Gamma' M_j B_i M_j \Gamma\varepsilon_j\right)^2}{(\varepsilon_j'\Gamma' M_j\Gamma\varepsilon_j)^2} = \frac{\mathrm{E}\left(\varepsilon_j'\Gamma' M_j B_i M_j \Gamma\varepsilon_j\right)^2}{\left[\mathrm{E}\left(\varepsilon_j'\Gamma' M_j\Gamma\varepsilon_j\right)\right]^2} \\ &= \frac{\mathrm{tr}^2\left(B_j B_i\right) + 2\mathrm{tr}\left(B_j B_i\right)^2 + \Delta \mathrm{tr}\left(B_j B_i \circ B_j B_i\right)}{\mathrm{tr}^2\left(B_j\right)} \leq \frac{(3+\Delta)\,\mathrm{tr}^2\left(B_j B_i\right)}{\mathrm{tr}^2\left(B_j\right)}.\end{aligned}$$

(2) $j_1 = j_2 \neq j_3 = j_4$.

$$\mathrm{E}\left(v_{j_1}' B_i v_{j_1}\right)\left(v_{j_3}' B_i v_{j_3}\right) = \mathrm{E}\left(v_{j_1}' B_i v_{j_1}\right)\mathrm{E}\left(v_{j_3}' B_i v_{j_3}\right) = \frac{\mathrm{tr}\left(B_{j_1} B_i\right)}{\mathrm{tr}\left(B_{j_1}\right)}\frac{\mathrm{tr}\left(B_{j_3} B_i\right)}{\mathrm{tr}\left(B_{j_3}\right)}.$$

(3) $j_1 = j_3 \neq j_2 = j_4$.

$$\begin{aligned}&\mathrm{E}\left(v_{j_1}' B_i v_{j_2}\right)\left(v_{j_1}' B_i v_{j_2}\right) = \mathrm{EE}\left(v_{j_1}' B_i v_{j_2} v_{j_2}' B_i v_{j_1} \mid v_{j_2}\right) \\ &= \mathrm{E}\left[\frac{\mathrm{tr}\left(\Gamma' M_{j_1} B_i M_{j_2}\Gamma\varepsilon_{j_2}\varepsilon_{j_2}'\Gamma' M_{j_2} B_i M_{j_1}\Gamma\right)}{\mathrm{tr}\left(M_{j_1}\Sigma\right)\varepsilon_{j_2}'\Gamma' M_{j_2}\Gamma\varepsilon_{j_2}}\right] \\ &= \frac{\mathrm{tr}\left(B_{j_2} B_i B_{j_1} B_i\right)}{\mathrm{tr}\left(B_{j_1}\right)\mathrm{tr}\left(B_{j_2}\right)}.\end{aligned}$$

Hence

$$\begin{aligned}&\mathrm{var}(Q_j'\Gamma M_i\Gamma' Q_j) = \mathrm{E}(Q_j'\Gamma M_i\Gamma' Q_j)^2 - \left[\mathrm{E}\left(Q_j'\Gamma M_i\Gamma' Q_j\right)\right]^2 \\ &\leq \sum_{j_1=1}^{i-1}\sum_{j_2=1, j_2\neq j_1}^{i-1}\frac{\mathrm{tr}\left(B_{j_1} B_i\right)\mathrm{tr}\left(B_{j_2} B_i\right)}{\mathrm{tr}\left(B_{j_1}\right)\mathrm{tr}\left(B_{j_2}\right)} + (3+\Delta)\sum_{j=1}^{i-1}\frac{\mathrm{tr}^2\left(B_j B_i\right)}{\mathrm{tr}^2\left(B_j\right)} \\ &+ 2\sum_{j_1=1}^{i-1}\sum_{j_2=1, j_2\neq j_1}^{i-1}\frac{\mathrm{tr}\left(B_{j_2} B_i B_{j_1} B_i\right)}{\mathrm{tr}\left(B_{j_1}\right)\mathrm{tr}\left(B_{j_2}\right)} - \left(\sum_{j=1}^{i-1}\frac{\mathrm{tr}\left(B_j B_i\right)}{\mathrm{tr}\left(B_j\right)}\right)^2 \\ &= 2\sum_{j_1=1}^{i-1}\sum_{j_2=1, j_2\neq j_1}^{i-1}\frac{\mathrm{tr}\left(B_{j_2} B_i B_{j_1} B_i\right)}{\mathrm{tr}\left(B_{j_1}\right)\mathrm{tr}\left(B_{j_2}\right)} + (2+\Delta)\sum_{j=1}^{i-1}\frac{\mathrm{tr}^2\left(B_j B_i\right)}{\mathrm{tr}^2\left(B_j\right)}.\end{aligned}$$

It further leads to

$$\begin{aligned}\operatorname{var}\left(\sum_{i=1}^{N}\delta_{Ni}^{2}\right) &\leq \frac{4}{N^{2}(N-1)^{2}}N\sum_{i=1}^{N}\operatorname{var}\left(\delta_{Ni}^{2}\right)\\ &\leq \frac{8}{N(N-1)^{2}}\sum_{i=1}^{N}\sum_{j_1=1}^{i-1}\sum_{j_2=1,j_2\neq j_1}^{i-1}\frac{\operatorname{tr}\left(B_{j_2}B_iB_{j_1}B_i\right)}{\operatorname{tr}^2\left(B_i\right)\operatorname{tr}\left(B_{j_1}\right)\operatorname{tr}\left(B_{j_2}\right)}\\ &+\frac{4(2+\Delta)}{N(N-1)^{2}}\sum_{i=1}^{N}\sum_{j=1}^{i-1}\frac{\operatorname{tr}^2\left(B_jB_i\right)}{\operatorname{tr}^2\left(B_i\right)\operatorname{tr}^2\left(B_j\right)}.\end{aligned}$$

By using Lemma A.3, we have

$$\operatorname{var}\left(\sum_{i=1}^{N}\delta_{Ni}^{2}\right) \leq K\left[O\left(\frac{1}{T^3}\right)+O(\frac{1}{NT^2})\right].$$

As $(N,T)\to\infty$, $\operatorname{var}\left(\sum_{i=1}^{N}\delta_{Ni}^{2}\right)/\operatorname{var}^2(T_n)\to 0$. Lemma B.1 is proved.

■

Lemma B.2 *Under Assumptions 1-3 and the null (2.2), as* $(N,T)\to\infty$*,*

$$\frac{\sum_{i=1}^{N}E\left(D_{N,i}^{4}\right)}{var^2\left(T_n\right)}\longrightarrow 0.$$

Proof. Rewrite

$$\begin{aligned}&\mathrm{E}\left(D_{N,i}^{4}\right)=\mathrm{EE}\left(D_{N,i}^{4}|\mathcal{F}_{N,i-1}\right)=\mathrm{EE}\left[\left(v_i'Q_jQ_j'v_i\right)^2\mid\mathcal{F}_{N,i-1}\right]\\ &=\mathrm{E}\left[\frac{\operatorname{tr}^2\left(\Gamma'M_iQ_jQ_j'M_i\Gamma\right)+2\operatorname{tr}(\Gamma'M_iQ_jQ_j'M_i\Gamma)^2+\Delta\operatorname{tr}\left(\Gamma'M_iQ_jQ_j'M_i\Gamma\circ\Gamma'M_iQ_jQ_j'M_i\Gamma\right)}{\operatorname{tr}^2\left(B_i\right)}\right].\end{aligned}$$

By using the results from Lemma B.1, we have

$$\begin{aligned}
&\mathrm{E}\left[\mathrm{tr}^2\left[\Gamma' M_i Q_j Q_j' M_i \Gamma\right]\right] = \mathrm{E}\left(Q_j' B_i Q_j\right)^2 \\
&\leq \sum_{j_1=1}^{i-1} \sum_{j_3=1, j_3\neq j_1}^{i-1} \frac{\mathrm{tr}\left(B_{j_1} B_i\right) \mathrm{tr}\left(B_{j_3} B_i\right)}{\mathrm{tr}\left(B_{j_1}\right) \mathrm{tr}\left(B_{j_3}\right)} + (3+\Delta) \sum_{j=1}^{i-1} \frac{\mathrm{tr}^2\left(B_j B_i\right)}{\mathrm{tr}^2\left(B_j\right)} \\
&+ 2 \sum_{j_1=1}^{i-1} \sum_{j_2=1, j_2\neq j_1}^{i-1} \frac{\mathrm{tr}\left(B_{j_2} B_i B_{j_1} B_i\right)}{\mathrm{tr}\left(B_{j_1}\right) \mathrm{tr}\left(B_{j_2}\right)}.
\end{aligned}$$

Since

$$\mathrm{tr}(\Gamma' M_i Q_j Q_j' M_i \Gamma)^2 \leq \mathrm{tr}^2\left(\Gamma' M_i Q_j Q_j' M_i \Gamma\right)$$

and

$$\mathrm{tr}\left(\Gamma' M_i Q_j Q_j' M_i \Gamma \circ \Gamma' M_i Q_j Q_j' M_i \Gamma\right) \leq \mathrm{tr}^2\left(\Gamma' M_i Q_j Q_j' M_i \Gamma\right),$$

thus

$$\begin{aligned}
&\sum_{i=1}^{N} \mathrm{E}\left(D_{N,i}^4\right) \leq \frac{K}{N^2(N-1)^2} \sum_{i=1}^{N} \sum_{j_1=1}^{i-1} \sum_{j_2=1, j_2\neq j_1}^{i-1} \frac{\mathrm{tr}\left(B_{j_1} B_i\right) \mathrm{tr}\left(B_{j_2} B_i\right)}{\mathrm{tr}^2\left(B_i\right) \mathrm{tr}\left(B_{j_1}\right) \mathrm{tr}\left(B_{j_3}\right)} \\
&+ \frac{K}{N^2(N-1)^2} \sum_{i=1}^{N} \sum_{j=1}^{i-1} \frac{\mathrm{tr}^2\left(B_j B_i\right)}{\mathrm{tr}^2\left(B_i\right) \mathrm{tr}^2\left(B_j\right)} \\
&+ \frac{K}{N^2(N-1)^2} \sum_{i=1}^{N} \sum_{j_1=1}^{i-1} \sum_{j_2=1, j_2\neq j_1}^{i-1} \frac{\mathrm{tr}\left(B_{j_2} B_i B_{j_1} B_i\right)}{\mathrm{tr}^2\left(B_i\right) \mathrm{tr}\left(B_{j_1}\right) \mathrm{tr}\left(B_{j_2}\right)} \\
&\leq \frac{K^2}{NT^2} = O\left(\frac{1}{NT^2}\right).
\end{aligned}$$

Hence $\frac{\sum_{i=1}^{N} \mathrm{E}\left(D_{N,i}^4\right)}{\mathrm{var}^2(T_n)} \longrightarrow 0$, as $(N,T) \to \infty$. Lemma B.2 is proved. ■

B.3 Proof of Theorem 3

Proof. we want to show

$$E(\hat{\gamma}^2) = \gamma^2 \text{ and } \hat{\gamma}^2 - \gamma^2 = o_p(1).$$

Note that

$$
\begin{aligned}
\hat{\gamma}^2 &= \frac{1}{2N(N-1)} \sum_{(i,j)}^{N} v_i' \left(v_j - \bar{v}_{(i,j)}\right) v_j' \left(v_i - \bar{v}_{(i,j)}\right) \\
&= \frac{1}{2N(N-1)} \left[\sum_{(i,j)}^{N} \left(v_i' v_j\right)^2 - v_i' v_j v_j' \bar{v}_{(i,j)} - v_i' \bar{v}_{(i,j)} v_j' v_i + v_i' \bar{v}_{(i,j)} v_j' \bar{v}_{(i,j)} \right] \\
&= a_1 + a_2 + a_3 + a_4, \text{ say.}
\end{aligned}
$$

It is easy to show that the first term $\mathrm{E}(a_1) = \gamma^2$, and $\mathrm{E}(a_i) = 0, i = 2, 3, 4$. So we prove the first part. By using Lemma 3 and Theorem 1, we have $\gamma^2 = O(T^{-1})$. Hence, to prove $\hat{\gamma}^2 - \gamma^2 = o_p(1)$, we only need to show that var(a_1)= $o_p(T^{-2})$ and $a_i = o_p\left(\gamma^2\right),$ for $i = 2, 3, 4$. Let us consider var(a_1),

$$
\begin{aligned}
\mathrm{var}(a_1) &= E(a_1^2) - \gamma^4 \\
&= \frac{4}{N^2(N-1)^2} \mathrm{E} \left(\sum_{i=1}^{N} \sum_{j=1}^{i-1} \hat{\rho}_{ij}^2 \right)^2 - \frac{4}{N^2(N-1)^2} \left(\sum_{i=2}^{N} \sum_{j=1}^{i-1} \frac{\mathrm{tr}\left(B_j B_i\right)}{\mathrm{tr}\left(B_i\right) \mathrm{tr}\left(B_j\right)} \right)^2 . \\
&= \frac{4}{N^2(N-1)^2} \mathrm{E} \left(\sum_{i_1=2}^{N} \sum_{j_1=1}^{i-1} \sum_{i_2=2}^{N} \sum_{j_2=1}^{i_2-1} \rho_{i_1 j_1}^2 \rho_{i_2 j_2}^2 \right) - \frac{4}{N^2(N-1)^2} \left(\sum_{i=2}^{N} \sum_{j=1}^{i-1} \frac{\mathrm{tr}\left(B_j B_i\right)}{\mathrm{tr}\left(B_i\right) \mathrm{tr}\left(B_j\right)} \right)^2
\end{aligned}
$$

Now we only consider the term $\mathrm{E}\Big(\sum_{i_1=2}^{N} \sum_{j_1=1}^{i-1} \sum_{i_2=2}^{N} \sum_{j_2=1}^{i-1} \rho_{i_1 j_1}^2 \rho_{i_2 j_2}^2 \Big)$. There are 3 cases for this term and Lemma 2 is used frequently:

(1) i_1, i_2, j_1 and j_2 are mutually different.

$$
\mathrm{E}\left(\rho_{i_1 j_1}^2 \rho_{i_2 j_2}^2\right) = \frac{\mathrm{tr}\left(B_{i_1} B_{j_1}\right) \mathrm{tr}\left(B_{i_2} B_{j_2}\right)}{\mathrm{tr}\left(B_{i_1}\right) \mathrm{tr}\left(B_{i_1}\right) \mathrm{tr}\left(B_{i_2}\right) \mathrm{tr}\left(B_{i_2}\right)} = O_p\left(\frac{1}{T^2}\right).
$$

(2) $i_1 = i_2$, $j_1 = j_2$ and $i_1 \neq j_1$.

$$
\mathrm{E}\left(\rho_{ij}^4\right) \leq (3+\Delta) \frac{(2+\Delta)\, \mathrm{tr}\left(B_i B_j\right)^2 + \mathrm{tr}^2\left(B_i B_j\right)}{\mathrm{tr}^2(B_i) \mathrm{tr}^2(B_j)} = O_p\left(\frac{1}{T^2}\right).
$$

(3) $i_1 = i_2$, $i_1 \neq j_1 \neq j_2$.

$$\begin{aligned}
\mathrm{E}\left(\rho_{ij_1}^2 \rho_{ij_2}^2\right) &\leq \frac{\left((2+\Delta)\,\mathrm{tr}\left(B_i B_{j_1}\right)^2 + \mathrm{tr}^2\left(B_i B_{j_1}\right)\right)^{1/2} \left((2+\Delta)\,\mathrm{tr}\left(B_i B_{j_2}\right)^2 + \mathrm{tr}^2\left(B_i B_{j_2}\right)\right)^{1/2}}{\mathrm{tr}\left(B_{j_1}\right)\mathrm{tr}(B_{j_2})\mathrm{tr}^2(B_i)} \\
&= O_p\left(\frac{1}{T^2}\right).
\end{aligned}$$

From above results, we have

$$\mathrm{var}(a_1) = O_p\left(\frac{1}{N^2 T^2}\right).$$

Hence $a_1 \xrightarrow{p} \gamma^2$. Consider the second term a_2, which is equal to $\frac{1}{2N(N-1)(N-2)} \sum_{(i,j,\tau)}^{N} v_i' v_j v_j' v_\tau$. The first term of $\mathrm{E}\left(\sum_{(i,j,\tau)}^{N} v_i' v_j v_j' v_\tau\right)^2$ is

$$\begin{aligned}
&\sum_{(i,j_1,j_2,\tau)}^{N} E\left(v_i' v_{j_1} v_{j_1}' v_\tau v_i' v_{j_2} v_{j_2}' v_\tau\right) \\
&= \sum_{(i,j_1,j_2,\tau)}^{N} \frac{tr(M_{j_2} M_\tau \Sigma M_\tau M_i \Sigma M_i M_{j_1} \Sigma M_{j_1} M_{j_2} \Sigma)}{tr\left(B_\tau\right) tr\left(B_{j_2}\right) tr(B_{j_1}) tr\left(B_i\right)} \\
&= O\left(N^4 T^{-3}\right),
\end{aligned}$$

by using Lemmas A.2-A.3. By using part (c) of Lemma A.3, the second term of $\mathrm{E}\left(\sum_{(i,j,s)}^{N} v_i' v_j v_j' v_\tau\right)^2$ is

$$E\left[\sum_{(i,j,\tau)}^{N} \left(v_i' v_j v_j' v_\tau\right)^2\right] = O_p(N^3 T^{-2}).$$

Hence $a_2 = O_p\left(N^{-1} T^{-3/2}\right) + O_p\left(N^{-3/2} T^{-1}\right)$, which further implies $a_2 = o_p(\gamma^2)$. Since $a_2 = a_3$, $a_3 = o_p(\gamma^2)$. Consider a_4, it can be divided into two terms:

$$\frac{1}{2N(N-1)\left(N-2\right)^2} \sum_{(i,j,\tau)}^{N} \left(v_i' v_s v_j' v_\tau\right) \text{ and } \frac{1}{2N(N-1)\left(N-2\right)^2} \sum_{(i,j,\tau_1,\tau_2)}^{N} \left(v_i' v_{\tau_1} v_j' v_{\tau_2}\right).$$

It is easy to show that the former term is $O_p\left(N^{-1}a_2\right)$, then it is $o_p(\gamma^2)$. We only need to consider the latter term $\mathrm{E}\left(\sum_{(i,j,\tau_1,\tau_2)}^{N}\left(v_i'v_{\tau_1}v_j'v_{\tau_2}\right)\right)^2$

$$\mathrm{E}\left[\sum_{(i,j,\tau_1,\tau_2)}^{N}\left(v_i'v_{\tau_1}v_j'v_{\tau_2}\right)^2\right] = \sum_{(i,j,\tau_1,\tau_2)}^{N}\mathrm{E}\left[\left(v_i'v_{\tau_1}\right)^2\left(v_j'v_{\tau_2}\right)^2\right] = O\left(N^4T^{-2}\right),$$

by using Lemma A.2-A.3. Hence the latter term is $O_p(N^{-2}T^{-1})$. The above results together lead to $a_4 = o_p(\gamma^2)$. The first part of Theorem 3 holds, the second part of Theorem 3 is directly derived by using Theorem 2 and the first part of Theorem 3. ■

Essay III: Tests of Specification for Large Dynamic Panel Data Models

1 Introduction

Dynamic panel data models have been of interest for years in a wide range of economic empirical studies, including labor participation, economic growth and finance, among many others. The generalized method of moments (GMM) estimation using lags as instruments is the most popular approach to estimate dynamic panel data models, see Eakin et al. (1988), Arellano and Bond (1991), Arellano and Bover (1995), Blundel and Bond (1998). However, existence of some serial correlation in errors invalidates moment conditions; besides, the existence of slope heterogeneity or cross-sectional dependence in errors leads to persistent serial correlation, which further invalidates moment conditions. Under such misspecifications, the GMM estimators become inconsistent. It is therefore essential to check the validity of moment conditions via diagnostic tests. One could use the overidentifying test, which serves as a general misspecification test, see Sargan (1958), Hansen (1982), Arellano and Bond (1991); or test the null of no serial correlation in first-difference errors. Arellano and Bond (1991) propose a test $\left(m_2^{(2)}\right)$ on testing the null hypothesis of no second-order serial correlation in the first-difference errors. However, this test may lack power for finite samples against more general high-order or persistent serial correlation. A more powerful joint test $\left(m_{(2,p)}^{(2)2}\right)$ is proposed later by Yamagata (2008) on testing the null of no second to pth-order serial correlation in the first-difference errors.

It is well known that the asymptotic framework used for the existing tests is standard: the dimension of time periods (T) is fixed and the number of cross-sectional units (N) is large. With the growing availability of panel data sets, T grows to be large and not negligible relative to N. Long periods of time provide many valid moment conditions for the GMM estimators. The Arellano and Bond's (1991) GMM (AB-GMM) estimator uses all available instruments in first-difference equations, and thus the number of orthogonality conditions grows at a rate of order T^2. When T is large, the number of instruments can easily become large relative to the sample size N, making some asymptotic results of GMM estimates and related tests misleading for dynamic panel models. If the number of instruments is larger than N, the estimated weight matrix of GMM is not invertible. Even if the number is smaller than N, the estimated weight matrix may not be a good approximation to the population counterpart. Voluminous statistic studies show that the sample covariance matrix of random vectors is a poor estimator for its population covariance matrix when its dimension is relatively as large as its sample size. Moreover, numerous instruments can over-fit instrumented variables and bias coefficient estimates in dynamic panel data models, see Alvarez and Arellano (2003), Bun and Kiviet (2006), Hayakawa (2015), Hsiao and Zhang (2015), Hsiao and Zhou (2015). The bias may accumulate and invalidate test statistics.

It turns out that the conventional Sargan's test proposed by Arellano and Bond (1991) for dynamic panels tends to be undersized and has virtually no power when the number of moment conditions is proportional to sample size, see simulation studies in Bowsher (2002) and Windmeijer (2005). With T fixed, the Sargan's test is asymptotically distributed as χ^2 distribution with degrees of freedom of order T^2. As $T \to \infty$, the tests diverges. Donald et al. (2003) suggest re-centering and re-scaling the χ^2 distribution to correct overidentifying test in cross-sectional regression models with moderately many instruments, i.e., the number of instruments grows asymptotically at a slower rate than the sample size. However, Anatolyev and Gospodinov (2010) show that when the number of instruments increases at the same rate as the sample size, this corrected test of overidentifying restrictions is asymptotically incorrect, a similar result can also be found in Hayakawa (2015). They propose a modification of the test. More studies of tests of overidentifying restrictions with many instruments in cross-sectional regression models can also be found in Lee and Okui (2012), and Chao et al. (2014). Despite the literature mentioned above, correction of the conventional Sargan's test for large dynamic panels has not been well studied.

It is worth pointing out that reducing the number of instruments used in GMM leads to less bias in parameter estimation in dynamic panels. Therefore, the literature suggests using less instruments for GMM estimates and related tests. Two different instrument matrices with less instruments for GMM estimates in dynamic panels are investigated by Bun and

Kiviet (2006). The first one is the block-diagonal matrix which includes only a subset of all available instruments; the second one is the matrix which includes a linear transformation of a subset of all available instruments. The latter one is also called collapsed instrument matrix and recommended by the Roodman (2009)[1]. The number of instruments used in this matrix is fixed, and not increasing with T. They show that the bias of GMM estimates with those two instrument matrices are smaller in comparison to the instrument matrix used in AB-GMM. However, limited literature provides either theoretical or empirical guidance for the usage of tests of specification which built upon the GMM estimates with different instrument matrices for large dynamic panels. More importantly, large N and large T asymptotics of the tests have not been well examined.

This paper considers the tests of specification, including the tests for serial correlation (Arellano (2003), Yamagata (2008)) and the tests of overidentifying restrictions (Arellano and Bond (1991)), for large dynamic panel data models. The test statistics are built upon the two-step GMM estimations using three different instrument matrices: a block-diagonal matrix with a full set of all available instruments, a block-diagonal matrix with a subset of all available instruments and the collapsed instrument matrix. We first extend the tests of serial correlation to the large N and large T framework; later we propose an accurate correction for the Sargan's tests when the number of instruments is growing with T. The limiting distributions of all the tests are derived as N and T go to infinity simultaneously. The asymptotic local power under $MA(q)$ and $AR(q)$ alternatives of the tests are investigated, which yields several findings. First, the tests for serial correlation are powerful against $MA(q)$ and $AR(q)$ alternatives; and the joint tests for the second to pth serial correlation are more powerful than the tests for sth-order serial correlation, which is consistent with the result of Yamagata (2008). Second, under the local $MA(q)$ and $AR(q)$ alternatives, the power of the corrected Sargan's tests decreases when the number of instruments increases. Therefore, their power only increases as N increases and the corrected Sargan's test with collapsed

[1]Roodman (2009) recommends using collapsed instruments matrix for estimation and test in large dynamic panel models, he also provides two empirical examples for verification.

instrument matrix are more powerful. Besides, the power properties are also discussed under the misspecifications such as heterogeneous slopes and cross-sectional dependence. The Monte Carlo simulations are conducted for checking the small sample properties. Simulation results confirm our theoretical findings. The results show that the corrected Sargan's tests have good size while the conventional Sargan's test suffers size distortions. Given the same type of tests, the test with collapsed instrument matrix often outperforms else others.

The rest of this paper is organized as follows. Section 2 introduces the model, the GMM estimators and the Assumptions; Section 3 presents the existing tests; Section 4 proposes the test statistics and derives their limiting distributions; Section 5 discusses the power properties; Section 6 reports the results of the Monte Carlo experiments; and Section 7 provides some concluding remarks. All the mathematical proofs are provided in the Appendix.

Notations: Denote $T_s = T - s - 2$. For a squared matrix B, $\mathrm{tr}(B)$ is the trace of B; $||B|| = (\mathrm{tr}(B'B))^{1/2}$ denotes the Frobenius norm of a matrix or the Euclidean norm of a vector B. $\xrightarrow{d}$ denotes convergence in distribution and $\xrightarrow{p}$ denotes convergence in probability. We use $(N,T) \to \infty$ to denote the joint convergence of N and T when N and T pass to infinity simultaneously. K is a generic positive number not depending on N or T.

2 The model and the estimators

Consider a dynamic panel data model

$$y_{it} = \rho y_{i,t-1} + \beta' x_{it} + \eta_i + u_{it},\ i = 1, \ldots, N,\ t = 2, \ldots, T, \tag{1}$$

where $|\rho| < 1$ and η_i are the unobserved individual effects with finite mean and variance. x_{it} are $(k-1) \times 1$ vector of predetermined regressors such that $\mathrm{E}(x_{is}u_{it}) \neq 0$ for $s > t$ and zero otherwise. u_{it} are residuals which are independent across time and individuals. For the presentation of the estimators below, it is convenient to stack model (1) for each i, which yields

$$y_i = \rho y_{i,-1} + x_i\beta + \eta_i\iota_T + u_i,\ i = 1, \ldots, N, \tag{2}$$

where $y_i = (y_{i1}, \ldots, y_{iT})'$, $y_{i,-1} = (y_{i0}, \ldots, y_{i,T-1})'$, $x_i = (x_{i1}, \ldots, x_{iT})'$, ι_T is a $(T \times 1)$ vector of ones, and $u_i = (u_{i1}, \ldots, u_{iT})'$. To eliminate the fixed effects η_i, taking the first difference of model (1) yields

$$\Delta y_{it} = \rho\Delta y_{i,t-1} + \beta'\Delta x_{it} + \Delta u_{it},\ i = 1, \ldots, N,\ t = 3, \ldots, T, \tag{3}$$

where $\Delta y_{it} = y_{it} - y_{i,t-1}$, $\Delta x_{it} = x_{it} - x_{i,t-1}$, and $\Delta u_{it} = u_{it} - u_{i,t-1}$. Similarly, stacking equation (3) for each i gives

$$\Delta y_i = \rho\Delta y_{i,-1} + \Delta x_i\beta + \Delta u_i,\ i = 1, \ldots, N, \tag{4}$$

where $\Delta y_i = (\Delta y_{i2}, \ldots, \Delta y_{iT})'$, $\Delta y_{i,-1} = (\Delta y_{i1}, \ldots, \Delta y_{i,T-1})'$, $\Delta x_i = (\Delta x_{i2}, \ldots, \Delta x_{iT})'$ and $\Delta u_i = (\Delta u_{i2}, \ldots, \Delta u_{iT})'$. To estimate the parameters ρ and β, Arellano and Bond (1991) suggest using the following moment conditions

$$E\left(y_{i(t-s)}\Delta u_{it}\right) = 0;\ \text{for } t = 3, \ldots, T \text{ and } 2 \leq s \leq t-1, \tag{5}$$

$$E\left(x_{i(t-s)}\Delta u_{it}\right) = 0;\ \text{for } t = 3, \ldots, T \text{ and } 1 \leq s \leq t-1. \tag{6}$$

Define

$$W_{li}^{(2)} = \begin{bmatrix} [y_{i1}, x_{i1}', x_{i2}'] & 0 & \cdots & 0 \\ 0 & [y_{i,1}, y_{i,2}, x_{i1}', x_{i2}', x_{i3}'] & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & [y_{i,1}, \ldots, y_{i,T-2}, x_{i1}', \ldots, x_{i,T-1}'] \end{bmatrix}.$$

Then the matrix of instruments is $W^{(2)} = [W_{l1}^{(2)\prime}, \ldots, W_{lN}^{(2)}{}']'$ and the moments conditions described above are given by $\mathrm{E}\Big(W_{li}^{(2)\prime}\Delta u_i\Big) = 0$. Pre-multiplying the difference equation (3) in vector form by $W^{(2)\prime}$ gives

$$W^{(2)\prime}\Delta y = W^{(2)\prime}\left(\Delta y_{-1}\right)\rho + W^{(2)\prime}\left(\Delta X\right)\beta + W^{(2)\prime}\Delta u, \tag{7}$$

where ΔX is the stacked $N(T-2)\times(k-1)$ matrix. Define $Z = [\Delta y_{-1}, \Delta X]$ with $Z_{it} = [\Delta y_{it-1}, \Delta x_{it}']$ and $\theta = (\rho, \beta')'$. The Arellano-Bond preliminary one-step GMM estimator is defined as

$$\hat{\theta}_{FDGMM1}^{(2)} = \left[Z'W^{(2)}\left(W^{(2)\prime}AW^{(2)}\right)^{-1}W^{(2)\prime}Z\right]^{-1}\left[Z'W^{(2)}\left(W^{(2)\prime}AW^{(2)}\right)^{-1}W^{(2)\prime}\Delta y\right], \tag{8}$$

where $A = I_N \otimes D$ and D is a $(T-2)\times(T-2)$ matrix with $2s$ on the main diagonal, $-1s$ on the first sub-diagonal, and zeros otherwise. The Arellano-Bond two-step GMM (AB-GMM) estimator is given by

$$\hat{\theta}_{FDGMM2}^{(2)} = \left[Z'W^{(2)}\hat{\Omega}_{(2)}^{-1}W^{(2)\prime}Z\right]^{-1}\left[Z'W^{(2)}\hat{\Omega}_{(2)}^{-1}W^{(2)\prime}\left(\Delta y\right)\right], \tag{9}$$

where $\hat{\Omega}_{(2)} = \frac{1}{N}\sum_{i=1}^{N} W_{li}^{(2)\prime}\left(\Delta\breve{u}_i^{(2)}\right)\left(\Delta\breve{u}_i^{(2)}\right)' W_{li}^{(2)}$ and $\Delta\breve{u}_i^{(2)}$ are the residuals estimated by the one-step GMM defined in (8).

The AB-GMM estimator uses $q_2 = (T-1)(T-2)/2 + (k-1)(T+1)(T-2)/2 = O(T^2)$ instruments given in (5) and (6). When T is fixed, the AB-GMM estimator is well performed and easily computed. But as T grows, the number of instruments q_2 becomes very large; for example, assume $k=1$ and $T=5$, $q_2 = 6$; when T is 50, $q_2 = 1176$. Hence, inverting a large dimension matrix is computationally difficult. If $q_2 > N$, we cannot compute $\hat{\Omega}_{(2)}^{-1}$ since $1/N\sum_{i=1}^{N} W_{li}^{(2)\prime}\left(\Delta\breve{u}_i^{(2)}\right)\left(\Delta\breve{u}_i^{(2)}\right)' W_{li}^{(2)}$ is singular. Even if $q_2 < N$, $\hat{\Omega}_{(2)}$ is not a good approximation for $\mathrm{E}\Big(W_{li}^{(2)\prime}\Delta u_i \Delta u_i' W_{li}^{(2)}\Big)$. In fact, if $T/N \to c \in (0,\infty)$, $\hat{\Omega}_{(2)}$ is not consistent to $\mathrm{E}\Big(W_{li}^{(2)\prime}\Delta u_i \Delta u_i' W_{li}^{(2)}\Big)$, see detailed discussion in Section 3. Consequently, the estimations and tests built upon this approach may perform poorly with large T.

Recently, several literature suggest using less instruments can reduce the bias of estimation. Bun and Kiviet (2006) study the performances of two different instrument matrices with less instruments. The first one is defined by

$$W_{li}^{(1)} = \begin{bmatrix} y_{i1} & x_{i2}' & 0 & 0' & \cdots & 0 & 0' \\ 0 & 0' & y_{i2} & x_{i3}' & \cdots & 0 & 0' \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0' & \cdots & \cdots & \cdots & y_{i,T-2} & x_{i,T-1}' \end{bmatrix} \tag{10}$$

This instrument matrix includes only $q_1 = k(T-2)$ instruments, which is a subset of all the instruments. The second one includes a linear transformation of the instruments in $W_{li}^{(2)}$, which is also called collapsed instrument matrix by Roodman (2009). If we fix the number of lags as instruments, say κ ($\kappa \geq 2$), then the corresponding instrument matrix is defined by

$$W_{li}^{(0)} = \begin{bmatrix} y_{i1} & x_{i2}' & 0 & 0' & \cdots & 0 & 0' \\ y_{i2} & x_{i3}' & y_{i1} & x_{i2}' & \cdots & 0 & 0' \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ y_{i,T-2} & x_{i,T-1}' & y_{i,T-3} & x_{i,T-2}' & \cdots & y_{i,T-\kappa} & x_{i,T-\kappa+1}' \end{bmatrix} \tag{11}$$

The above matrix only includes $q_0 = k \times \kappa$ instruments. The one- and two-step GMM estimators using above two instrument matrices are given by,

$$\hat{\theta}_{FDGMM1}^{(j)} = \left[Z'W^{(j)}\left(W^{(j)\prime}AW^{(j)}\right)^{-1}W^{(j)\prime}Z\right]^{-1}\left[Z'W^{(j)}\left(W^{(j)\prime}AW^{(j)}\right)^{-1}W^{(j)\prime}\Delta y\right], \tag{12}$$

and

$$\hat{\theta}_{FDGMM2}^{(j)} = \left[Z'W^{(j)}\hat{\Omega}_{(j)}^{-1}W^{(j)\prime}Z\right]^{-1}\left[Z'W^{(j)}\hat{\Omega}_{(j)}^{-1}W^{(j)\prime}\Delta y\right],\ j = 0, 1; \tag{13}$$

respectively, where

$$\hat{\Omega}_{(1)} = \frac{1}{N}\sum_{i=1}^{N} W_{li}^{(1)}\left(\Delta\breve{u}_i^{(1)}\right)\left(\Delta\breve{u}_i^{(1)}\right)' W_{li}^{(1)},$$

and

$$\hat{\Omega}_{(0)} = \frac{1}{(NT_0)}\sum_{i=1}^{N} W_{li}^{(0)}\left(\Delta\breve{u}_i^{(0)}\right)\left(\Delta\breve{u}_i^{(0)}\right)' W_{li}^{(0)}.$$

$\Delta\breve{u}_i^{(j)}$ are the one-step GMM residuals, and $W^{(j)} = \left[W_{l1}^{(j)}, \cdots, W_{lN}^{(j)\prime}\right]'$, $j = 0, 1$. The consistency of the above estimators can be obtained directly from Bun and Kiviet (2006), more specifically,

$$\left\|\hat{\theta}_{FDGMM2}^{(j)} - \theta\right\| = O_p\left(\frac{1}{\sqrt{NT}}\right), \; j = 0, 1, 2. \tag{14}$$

To facilitate our analysis, we need the following Assumptions.

Assumption 1 *$\{y_i\}_{i=1}^N$ and $\{x_i\}_{i=1}^N$ are independently and identically distributed (i.i.d) random matrices. For each y_{it}, x_{it}, $i = 1, \ldots, N$, $t = 1, \ldots, T$, they have finite fourth moments: $E(y_{it}^4) \leq K < \infty$ and $E\left[\|x_{it}\|^4\right] \leq K < \infty$.*

Assumption 2 *(i) $\{u_{it}\}$ $(t = 1, \ldots, T;\; i = 1, \ldots, N)$ are i.i.d across time and individuals, with mean zero and a finite positive variance σ_u^2, and finite moments up to the fourth order. (ii) $E(u_{it}|y_{i,t-1}, \ldots, y_{i1}, x_{it}, \ldots, x_{i1}, \eta_i) = 0$, $t = 2, \ldots, T$.*

Assumption 3 *(i) $\Sigma_1^{(0)} = E\left(Q_i^{(0)} Q_i^{(0)\prime}\right)$, $\Sigma_{2s}^{(j)} = E\left(Q_{iTs}^{(0)} Q_{iTs}^{(0)\prime}\right)$ and $\Sigma_3^{(0)} = E\left(w_{iTs}^{(0)} w_{iTs}^{(0)\prime}\right)$ are symmetric positive definite matrices, where $Q_i^{(0)} = \left(v_{ip}, w_{iT2}^{(0)}\right)'$ and $Q_{iTs}^{(0)} = \left(v_{iTs}, w_{iTs}^{(0)}\right)'$ with $v_{ip} = (v_{iT2}, \ldots, v_{iTp})'$, where $v_{iTs} = T_s^{-1/2}\sum_{t=3}^{T-s}\Delta u_t \Delta u_{t+s}$; $w_{iTs}^{(0)} = T_s^{-1/2} W_{li}^{(0)\prime}\Delta u_i$, $s = 2, \ldots, p$. (ii) $\liminf_T \lambda_{\min}\left(\Sigma_1^{(0)}\right) > 0$; $\liminf_T \lambda_{\min}\left(\Sigma_{2s}^{(0)}\right) > 0$ and $\liminf_T \lambda_{\min}\left(\Sigma_3^{(0)}\right) > 0$, where $\lambda_{\min}$ denotes the minimum eigenvalue of a matrix.*

Assumptions 1-2 follow the literature, see Arellano (2003), Yamagata (2008). Assumption 1 requires all the variables are i.i.d across i and have finite fourth moments in that they are needed to apply the Central Limit Theorem in the paper. Part (i) of Assumption 2 excludes heteroscedastic time series; part (ii) of Assumption 2 ensures the validity of the moment conditions (5) and (6). Assumption 3 provides the requirements for applying the joint limit CLT for scaled variates of Phillips and Moon (1999).

3 Existing tests of specification

The standard tests for checking the validity of instruments in dynamic panel models are Sargan's test and tests for serial correlation. The test for second-order serial correlation, the Sargan's test and Sargan's difference test for overidentifying restrictions proposed by Arellano and Bond (1991) are most widely used in empirical applications. In this section, more generic versions of tests for serial correlation and tests of overidentifying restrictions are discussed. The joint test for the second to pth-order serial correlation proposed by Yamagata (2008) is also introduced. Besides, we also address the issues raised by large T.

3.1 Tests for serial correlation

Arellano and Bond (1991) propose a test $\left(m_2^{(2)}\right)$ for testing the hypothesis that there is no second-order serial correlation for the disturbance of first-difference equation (3). Arellano (2003) extends this test to test the null of no sth-order serial correlation, $s = 2, \ldots, p$. The hypotheses of this test are:

$$H_0 : \mathrm{E}(\Delta u_{it} \Delta u_{i,t+s}) = 0 \text{ against } H_1 : \mathrm{E}(\Delta u_{it} \Delta u_{i,t+s}) \neq 0. \tag{15}$$

The test statistic is given by

$$m_s^{(2)} = \frac{1}{\sqrt{N \hat{\gamma}_{(2)}^2}} \sum_{i=1}^{N} \hat{v}_{iTs}^{(2)}, \tag{16}$$

where $\hat{v}_{iTs}^{(2)} = \sum_{t=3}^{T-s} \Delta\hat{u}_{it}^{(2)} \Delta\hat{u}_{i,t+s}^{(2)}$, with $\Delta\hat{u}_{i,t}^{(2)} = \Delta y_{it} - Z_{it}'\hat{\theta}_{FDGMM2}^{(2)}$, and

$$\hat{\gamma}_{(2)}^2 = \frac{1}{N} \sum_{i=1}^{N} \hat{v}_{iTs}^{(2)2} + \hat{\varpi}_{Ns}^{(2)\prime} \hat{\Psi}_N^{(2)-1} \hat{\varpi}_{Ns}^{(2)} - 2\hat{\varpi}_{Ns}^{(2)\prime} \hat{\Psi}_N^{(2)-1} \left(\frac{Z' W^{(2)} \hat{\Omega}_{(2)}^{-1}}{N} \right) \left(\frac{1}{N} \sum_{i=1}^{N} W_{li}^{(2)\prime} \Delta\check{u}_i^{(2)} \hat{v}_{iTs}^{(2)} \right),$$

with

$$\hat{\varpi}_{Ns}^{(2)} = \frac{1}{N}\sum_{i=1}^{N}\sum_{t=3}^{T-s}\Delta\hat{u}_{it}^{(2)}Z_{i,t+s}$$

and

$$\hat{\Psi}_N^{(2)} = \left(\frac{1}{N}\sum_{i=1}^{N}W_{li}^{(2)\prime}Z_i\right)^{\prime}\hat{\Omega}_{(2)}^{-1}\left(\frac{1}{N}\sum_{i=1}^{N}W_{li}^{(2)\prime}Z_i\right).$$

Under the null hypothesis, as $N \to \infty$ with T fixed, $m_s^{(2)}$ converges to the standard normal distribution. Yamagata (2008) points out that it may not have enough power against more general higher serial correlation. He considers a joint test on testing the null of no second to pth-order first difference error serial correlation. The hypotheses of this joint test is

$$H_0 : \mathrm{E}(\Delta u_{it}\Delta u_{i,t+s}) = 0 \text{ jointly for } s = 2, \ldots, p \ \ (\leq T-3)\,, \tag{17}$$

against

$$H_1 : \mathrm{E}(\Delta u_{it}\Delta u_{i,t+s}) \neq 0; \text{ for some } s. \tag{18}$$

The joint test statistic is defined as

$$m_{(2,p)}^{(2)2} = \iota_N^{\prime}\hat{V}^{(2)}\left(\hat{G}^{(2)\prime}\hat{G}^{(2)}\right)^{-1}\hat{V}^{(2)\prime}\iota_N \tag{19}$$

where ι_N is a $N\times1$ vector of ones; $\hat{V}^{(2)} = \left(\hat{v}_1^{(2)}, \ldots, \hat{v}_N^{(2)}\right)^{\prime}$ with $\hat{v}_i^{(2)} = \left(\hat{v}_{iT2}^{(2)}, \ldots, \hat{v}_{iTp}^{(2)}\right)^{\prime}$; $\hat{G}^{(2)} = \left(\hat{g}_1^{(2)}, \ldots, \hat{g}_N^{(2)}\right)^{\prime}$ with $\hat{g}_1^{(2)} = \left(\hat{g}_{i2}^{(2)}, \ldots,\ \hat{g}_{ip}^{(2)}\right)$ and $\tilde{g}_{is} = \hat{v}_{iTs}^{(2)} - \hat{\varpi}_{Ns}^{(2)\prime}\hat{\Psi}_N^{(2)-1}\left[N^{-1}Z^{\prime}W^{(2)}\hat{\Omega}_{(2)}^{-1}\right]$ $W_{li}^{(2)\prime}\Delta\check{u}_i^{(2)}$, $s = 2, \ldots, p$. Under the null hypothesis (17), as $N \to \infty$ with fixed T, $m_{(2,p)}^{(2)2} \xrightarrow{d} \chi_{p-1}^2$.

Similar tests $m_s^{(1)}$ and $m_{(2,p)}^{(1)2}$, which are built upon the two-step GMM estimation with instrument matrix $W^{(1)}$, can be defined by modifying all the notations "(2)" to "(1)" respectively.

3.2 Tests of overidentifying restrictions

Arellano and Bond (1991) propose a Sargan's test of overidentifying restrictions for dynamic panel data models. The test statistic is defined as

$$S_{(2)} = \left(\sum_{i=1}^{N} \Delta\hat{u}_i^{(2)\prime} W_{li}^{(2)}\right) \left[\sum_{i=1}^{N} \left(W_{li}^{(2)\prime} \Delta\check{u}_i^{(2)} \Delta\check{u}_i^{(2)\prime} W_{li}^{(2)}\right)\right]^{-1} \left(\sum_{i=1}^{N} W_{li}^{(2)\prime} \Delta\hat{u}_i^{(2)}\right) \tag{20}$$

Under the null hypothesis $\mathrm{E}\left(W_{li}^{(2)\prime} \Delta u_i\right) = 0$, as $N \to \infty$, and T fixed, $S_{(2)}$ converges to χ^2 distribution with degree of freedom $q_2 - k$. To check the validity of subsets of moment restrictions, they suggest a Sargan's difference test. More specifically, suppose $W_{li}^{(2)}$ can be decomposed into two subsets of instruments as follows,

$$W_{li}^{(2)} = \left[W_{li1}^{(2)}, W_{li2}^{(2)}\right], \tag{21}$$

where $W_{li1}^{(2)}$ is $(T-2) \times q_{21}$ matrix with valid instrument sets, $W_{li2}^{(2)}$ is $(T-2) \times q_{22}$ matrix with invalid instrument sets under the alternative, and $q_{21} + q_{22} = q_2$. For example, if the errors in levels are $MA(1)$, $W_{li2}^{(2)}$ contains y_{it-2}, $x_{i.t-1}$ and x_{it-2}, while $W_{li1}^{(1)}$ contains the other lags. The Sargan's difference test for dynamic panels is defined as

$$DS_{(2)} = S_{(2)} - S_{(2)I}, \tag{22}$$

where

$$S_{(2)I} = \left(\sum_{i=1}^{N} \Delta\hat{u}_{i1}^{(2)\prime} W_{li1}^{(2)}\right) \left[\sum_{i=1}^{N} \left(W_{li1}^{(2)\prime} \Delta\check{u}_{i1}^{(2)} \Delta\check{u}_{i1}^{(2)\prime} W_{li1}^{(2)}\right)\right]^{-1} \left(\sum_{i=1}^{N} W_{li1}^{(2)\prime} \Delta\hat{u}_{i1}^{(2)}\right), \tag{23}$$

where $\check{u}_{li1}^{(2)}$ and $\Delta\hat{u}_{i1}^{(2)}$ are the one- and two-step AB-GMM residuals by using instruments matrix $W_{li1}^{(2)}$, respectively. Under the null hypothesis of $\mathrm{E}\left(W_{li}^{(2)\prime} \Delta u_i\right) = 0$, as $N \to \infty$, T fixed, $DS_{(2)} \xrightarrow{d} \chi^2_{q_{22}}$.

We define the Sargan's test based upon the two-step GMM estimation with instrument matrix $W^{(1)}$ to be $S_{(1)}$ by replacing all the notations "(2)" by "(1)" respectively. Under the null hypothesis of $\mathrm{E}\left(W_{li}^{(1)\prime}\Delta u_i\right) = 0$, as $N \to \infty$, T fixed, $S_{(1)}$ converges to χ^2 with degree of freedom $q_1 - k$. We do not define the Sargan's difference test here since $W^{(1)}$ only includes the instruments of one lag.

3.3 Issues raised by large T

It is well known that the asymptotics of the existing tests are derived with fixed T. As T being large, some issues are existing in the tests of specification. First, the limiting distributions of existing tests are only derived with fixed T. For example, as T fixed, $N \to \infty$, $S_{(2)}$ is distributed as χ^2 with degree of freedom $q_2 - k$; However, as $T \to \infty$, the test statistic diverges. How to derive the limiting distributions of the tests under large (N, T) framework is not studied. Second, several simulations have shown that if T is relatively large, the tests of overidentifying restrictions perform poorly. Bowsher (2002) uses Monte Carlo experiments and shows fixing N at 100, and letting T increase over the range $(5, 7, 9, 11, 13, 15)$, the Sargan's test $\left(S_{(2)}\right)$ using the full set of moment conditions tends to be undersized. He shows that the Monte Carlo variance of the test is much smaller than the chi-square approximation. For example, when $T = 15$, it is 13.7 when it should be 180. Two reasons are leading to this results. The first one is the estimation bias. The second one is the inaccurate chi-squared approximation, when the number of instruments used is relatively as large as N.

Consider the effect of estimation bias first. Hsiao and Zhou (2015) show that the GMM estimate is asymptotically biased of order $\sqrt{c}$ using all lags or one lag as instruments where $c = T/N \in (0, \infty)$ as $(N, T) \to \infty$. More specifically, $\mathrm{E}\left[\sqrt{NT}\left(\hat{\theta}^{(j)}_{FDGMM2} - \theta\right)\right] = b_j/N \neq 0$, for $j = 1, 2$, where b_j is some certain non-zero constant. $\mathrm{E}\left(W_{li}^{(2)}\Delta\hat{u}_i^{(2)}\right)$ is not zero or not asymptotic zero due to the accumulation of the bias when T is large. To better understand the issue, we first consider the estimated moment restrictions,

$$\mathrm{E}\left(W_{li}^{(2)\prime}\Delta\hat{u}_i^{(2)}\right) = \mathrm{E}\left(W_{li}^{(2)\prime}\Delta u_i\right) - \mathrm{E}\left[W_{li}^{(2)}\Delta Z_i'\left(\hat{\theta}^{(2)}_{FDGMM2} - \theta\right)\right].$$

Together with the fact that $\left\| W_{li}^{(2)} \Delta Z_{it}' \right\| = O_p\left(T^2\right)$, we have

$$\left\| \mathrm{E}\left[W_{li}^{(2)} \Delta Z_i' \left(\hat{\theta}_{FDGMM2}^{(2)} - \theta \right) \right] \right\| = \left(\frac{T}{N} \right)^{3/2} C. \tag{24}$$

where C is some non-zero constant. As $(N, T) \to \infty$ and $T/N \to c \in (0, \infty)$, $\mathrm{E}\left(W_{li}^{(2)\prime} \Delta \hat{u}_i^{(2)} \right) \neq \mathrm{E}\left(W_{li}^{(2)\prime} \Delta u_i \right) = 0$ under the null hypothesis. According to the above results, as $(N, T) \to \infty$ and $T/N \to c \in (0, \infty)$, we have $\left\| W_{li}^{(2)\prime} \Delta \hat{u}_i^{(2)} - W_{li}^{(2)\prime} \Delta u_i \right\| = O_p(1)$; It further leads to inaccuracy in estimation of the variance. As a consequence, the Sargan's test using the full set of instruments performs poorly when T is growing relatively as large as N. The accumulation of the bias in the conventional Sargan's test is because the dimension of instrument matrix is growing with T. On the contrary, the bias does not accumulate in the tests for serial correlation.

Recent literature also points out that the poor property of conventional Sargan's tests is due to the inaccurate approximation for the limiting distribution when the number of instruments is becoming relatively large as the sample size. More specifically, consider a general test statistic of overidentifying restrictions, which is defined by $\hat{T} = N\bar{g}(\hat{\theta})'\hat{\Omega}_g^{-1}\bar{g}(\hat{\theta})$, where $\bar{g}(\hat{\theta}) = \sum_{i=1}^{N} g_i\left(\hat{\theta}\right)$, $\hat{\Omega}_g = N^{-1}\sum_{i=1}^{N} g_i\left(\hat{\theta}\right) g_i\left(\hat{\theta}\right)'$ and $g_i\left(\hat{\theta}\right)$ is $p_1 \times 1$ vector of estimated moment restrictions. Let the number of regressors is k. Hence, under some assumptions, $\hat{T}$ is asymptotically distributed as χ^2 with degrees of freedom $p_1 - k$. However, when p_1 increases, $\hat{T}$ diverges. Donald et al. (2003) suggest a correction for $\hat{T}$ when its degrees of freedom increases, which yields

$$\hat{J} = \frac{\hat{T} - p_1 + k}{\sqrt{2\left(p_1 - k\right)}} \xrightarrow{d} N\left(0, 1\right). \tag{25}$$

However, from part (i) Theorem 1 in Anatolyev and Gospodinov (2010), it shows when $\lim_{(N,T)\to\infty} \left(p_1 - k\right)/N \in (0, 1)$,

$$\frac{\hat{J}}{\sqrt{1 - \frac{p_1 - k}{N}}} \xrightarrow{d} N\left(0, 1\right). \tag{26}$$

Similar to above result, Hayakawa (2015) also shows that $\hat{T}/N$ approximates to a Beta distribution with parameter $(p_1 - k)/2$ and $(N - p_1 + k)/2$. Hence, the mean and variance of $\hat{T}$ are given by $p_1 - k$ and $2(p_1 - k)(N - p_1 + k)/(N + 2)$. Note that the variance

$$2(p_1 - k)\left(\frac{N - p_1 + k}{N + 2}\right) \approx 2(p_1 - k)\left(1 - \frac{p_1 - k}{N}\right).$$

If p_1 is relatively as large as N, then $2(p_1 - k)\left(1 - \frac{p_1 - k}{N}\right)$ is much smaller than $2(p_1 - k)$ when we use chi-square approximation with degrees of freedom $(p_1 - k)$. Hence, this explains why the variance of $S_{(2)}$ is much smaller than the chi-square variance in Bowsher's (2002) simulation.

Overall, when the dimension of the instrument matrix used in dynamic panels increases with growing T, it may lead to bias accumulation in test statistics. Especially, when the number of instruments becomes relatively as large as N, the chi-squared approximation is not valid.

4 Test statistics and asymptotics

In this Section, we introduce the tests for serial correlation and the tests of overidentifying restrictions built upon the two-step GMM estimation with instrument matrix $W_{li}^{(0)}$. We present the statistics here since they are different from those introduced in Section 3. Using this collapsed instrument matrix $W_{li}^{(0)}$, the test statistics are double-index summations with some correlations across their elements. The instrument matrix $W_{li}^{(0)}$ has several advantages. First, the collapsed form is a linear transformation of a subset of all valid instruments, and it keeps more information without dropping too many lags. Second, the number of instruments is reduced dramatically and is much smaller compared to the sample size. The corrections of $S_{(j)}$, $j = 0, 1, 2$, are proposed for large q_j. The limiting distributions of the tests with different instruments $W_{li}^{(j)}$, $j = 0, 1, 2$, are derived as $(N, T) \to \infty$ in each subsection.

4.1 Asymptotics of the tests for serial correlation

Similar to (16), we define the test for serial correlation based upon the two step GMM estimation with instrument matrix $W_{li}^{(0)}$,

$$m_s^{(0)} = \frac{1}{\sqrt{N\hat{\gamma}_{(0)}^2}} \sum_{i=1}^{N} \hat{v}_{iTs}^{(0)}, \tag{27}$$

where $\hat{v}_{iTs}^{(0)} = \frac{1}{\sqrt{T_s}} \sum_{t=3}^{T-s} \Delta\hat{u}_{it}^{(0)} \Delta\hat{u}_{i,t+s}^{(0)}$, with $\Delta\hat{u}_{it}^{(0)} = \Delta y_{it} - Z_{it}'\hat{\theta}_{FDGMM2}^{(0)}$, and

$$\hat{\gamma}_{(0)}^2 = \frac{1}{N}\sum_{i=1}^{N} \hat{v}_{iTs}^{(0)2} + \hat{\varpi}_{Ns}^{(0)\prime}\hat{\Psi}_N^{(0)-1}\hat{\varpi}_{Ns}^{(0)} - 2\hat{\varpi}_{Ns}^{(0)\prime}\hat{\Psi}_N^{(0)-1}\left(\frac{Z'W^{(0)}\hat{\Omega}_{(0)}^{-1}}{NT_s}\right)\left(\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T_s^{1/2}}W_{li}^{(0)\prime}\Delta\check{u}_i^{(0)}\hat{v}_{iTs}^{(0)}\right),$$

where

$$\hat{\varpi}_{Ns}^{(0)} = \frac{1}{NT_s}^{-1} \sum_{i=1}^{N}\sum_{t=3}^{T-s} \Delta\hat{u}_{it}^{(0)}\Delta Z_{i,t+s}$$

and

$$\hat{\Psi}_N^{(0)} = \left[\frac{1}{(NT_s)^2}\left(\Delta y_{-1}\right)' W^{(0)}\hat{\Omega}_{(0)}^{-1}W^{(0)\prime}\left(\Delta y_{-1}\right)\right].$$

It should be noted that as $T \to \infty$, deriving the asymptotics of the serial correlation tests is not straightforward. For the case with large N and fixed T, the limiting theory is natural and has been well studied. However, the test defined by (27) is a double-index summation with some correlations across its elements. Hence, its limit cannot be obtained straightforwardly by using classical Multivariate Central Limit Theory. To cover this, we derive the limiting distribution of $m_s^{(0)}$ by applying the joint limit CLT for scaled variates given by Phillips and Moon (1999). The following Theorem gives the result.

Theorem 1 *Under the Assumptions 1-3 and the null hypothesis of (15), as* $(N, T) \to \infty$,

$$m_s^{(0)} \xrightarrow{d} N(0,1). \tag{28}$$

Similar to Yamagata (2008), we can also extend this result to define a new joint test of no second to pth-order serial correlation in first-difference errors, which is given by

$$m_{(2,p)}^{(0)2} = \iota_N' \hat{V}^{(0)} \left(\hat{G}^{(0)\prime} \hat{G}^{(0)} \right)^{-1} \hat{V}^{(0)\prime} \iota_N \tag{29}$$

where ι_N is a $N \times 1$ vector of ones; $\hat{V}^{(0)} = \left(\hat{v}_1^{(0)}, \ldots, \hat{v}_N^{(0)} \right)'$ with $\hat{v}_i^{(0)} = \left(\hat{v}_{iT2}^{(0)}, \ldots, \hat{v}_{iTp}^{(0)} \right)'$; $\hat{G}^{(0)} = \left(\hat{g}_1^{(0)}, \ldots, \hat{g}_N^{(0)} \right)'$ with $\hat{g}_i^{(0)} = \left(\hat{g}_{i2}^{(0)}, \ldots, \hat{g}_{ip}^{(0)} \right)$, where $\hat{g}_{is}^{(0)} = \hat{v}_{iTs}^{(0)} - \hat{\varpi}_{Ns}^{(0)\prime} \hat{\Psi}_N^{(0)-1} \left[(NT_s)^{-1} Z' W^{(0)} \hat{\Omega}_{(0)}^{-1} \right] \left(T_s^{-1/2} W_{li}^{(0)\prime} \Delta \hat{u}_i^{(0)} \right)$. The following Theorem gives the limiting distribution of $m_{(2,p)}^{(0)2}$.

Theorem 2 *Under the Assumptions 1-3 and the null hypothesis of (17), as* $(N, T) \to \infty$,

$$m_{(2,p)}^{(0)2} \xrightarrow{d} \chi_{p-1}^2. \tag{30}$$

Remark 1 *We derive the limiting distributions of the above two tests under the framework with large* T*. The results also hold for fixed* T*. Besides, Yamagata (2008) shows that the* $m_2^{(2)}$ *tests are equivalent to* $m_{(2,2)}^{(2)2}$*; similarly,* $m_2^{(j)}$ *tests are equivalent to* $m_{(2,2)}^{(j)2}$, $j = 0, 1$.

Since $m_s^{(j)}$ and $m_{(2,p)}^{(j)}$ are similar to $m_s^{(0)}$ and $m_{(2,p)}^{(j)}$, $j = 1, 2$; respectively, we can easily extend the $m_s^{(j)}$ and $m_{(2,p)}^{(j)}$ to the large (N, T) asymptotics. The following Corollary gives the results.

Corollary 1 *With Assumptions 1-3; (1) under the null hypothesis of (15), as* $(N, T) \to \infty$, $\frac{q_2 - k}{N} \in [0, 1)$, $m_s^{(2)} \xrightarrow{d} N(0, 1)$; (2) *Under the null hypothesis of (15), as* $(N, T) \to \infty$, $\frac{q_1 - k}{N} \in [0, 1)$, $m_s^{(1)} \xrightarrow{d} N(0, 1)$; *(3) Under the null hypothesis of (17), as* $(N, T) \to \infty$, $\frac{q_2}{N} \in [0, 1)$, $m_{(2,p)}^{(2)2} \xrightarrow{d} \chi_{p-1}^2$; *(4) Under the null hypothesis of (17), as* $(N, T) \to \infty$, $\frac{q_1 - k}{N} \in [0, 1)$, $m_{(2,p)}^{(1)2} \xrightarrow{d} \chi_{p-1}^2$.

Corollary 1 shows that the tests for serial correlation $m_s^{(j)}$ and $m_{(2,p)}^{(j)2}$, $j = 1, 2$; are also valid with large T.

4.2 Corrected tests of overidentifying restrictions and their asymptotics

The Sargan's test which is built upon the two-step GMM with instrument matrix $W_{li}^{(0)}$ is defined as

$$S_{(0)} = \left(\sum_{i=1}^{N} \Delta\hat{u}_i^{(0)\prime} W_{li}^{(0)}\right) \left[\sum_{i=1}^{N} \left(W_{li}^{(0)\prime} \Delta\check{u}_i^{(0)} \Delta\check{u}_i^{(0)\prime} W_{li}^{(0)}\right)\right]^{-1} \left(\sum_{i=1}^{N} W_{li}^{(0)\prime} \Delta\hat{u}_i^{(0)}\right) \tag{31}$$

Note that $\sum_{i=1}^{N} W_{li}^{(0)\prime} \Delta u_i$ can be rewritten as $\sum_{i=1}^{N} \sum_{t=3}^{T-s} W_{lit}^{(0)} \Delta u_{it}$, which is a double index summation of dependent multivariate variates and their dimensions are fixed. The limiting distribution of $S_{(0)}$ is given by the Theorem as below.

Theorem 3 *Under the Assumptions 1-3 and the null hypothesis* $E\left(W_{li}^{(0)\prime} \Delta u_i\right) = 0$, $i = 1, \ldots, N;$ *as* $(N, T) \to \infty$,

$$S_{(0)} \xrightarrow{d} \chi^2_{q_0 - k}. \tag{32}$$

Similarly, we can also define a Sargan's difference test to check the validity of a subset of instruments. Decompose the matrix of instruments $W_{li}^{(0)}$ into two subsets $W_{li}^{(0)} = \left[W_{li1}^{(0)}, W_{li2}^{(0)}\right]$, where $W_{li1}^{(0)}$ is $(T-2) \times q_{02}$ matrix with valid instruments set and $W_{li2}^{(0)}$ is $(T-2) \times (q_0 - q_{02})$ matrix with invalid instruments set. The Sargan's difference test is defined by $DS_{(0)} = S_{(0)} - S_{(0)I}$, with

$$S_{(0)I} = \left(\sum_{i=1}^{N} \Delta\hat{u}_{i1}^{(0)\prime} W_{li1}^{(0)}\right) \left[\sum_{i=1}^{N} \left(W_{li1}^{(0)\prime} \Delta\check{u}_{i1}^{(0)} \Delta\check{u}_{i1}^{(0)\prime} W_{li1}^{(0)}\right)\right]^{-1} \left(\sum_{i=1}^{N} W_{li1}^{(0)\prime} \Delta\hat{u}_{i1}^{(0)}\right).$$

The limiting distribution of $DS_{(0)}$ is given below.

Theorem 4 *Under the Assumptions 1-3 and the null hypothesis* $E\left(W_{li}^{(0)\prime} \Delta u_i\right) = 0$, $i = 1, \ldots, N;,$ *as* $(N, T) \to \infty$,

$$DS_{(0)} \xrightarrow{d} \chi^2_{q_0 - q_{02}}. \tag{33}$$

Remark 2 *The Sargan's difference test uses precise information about the alternative, but we cannot observe the information in practice. Hence, we neglect the power discussion and simulations in this paper.*

With fixed T, as $N \to \infty$, $S_{(j)}$ converge to χ^2 distribution with degrees of freedom $q_j - k$, for $j = 0, 1, 2$. As $T \to \infty$, q_1 and q_2 diverge, hereafter the test statistics. Following Donald et al. (2003), we use the asymptotic normal approximation to the chi-square for large degrees of freedom. Define

$$\hat{J}_{(j)} = \frac{S_{(j)} - q_j + k}{\sqrt{2\,(q_j - k)}}, \ j = 0, 1, 2. \tag{34}$$

As discussed in Section 3.3, the above normal approximation is not accurate if $\lim\limits_{(N,T)\to\infty} (q_j - k) / N = c_j \in (0, 1)$. Following Anatolyev and Gospodinov (2010), we define the corrected Sargan's tests as below,

$$\tilde{J}_{(j)} = \frac{S_{(j)} - q_j + k}{\sqrt{2(q_j - k)\,(1 - (q_j - k)\,/N)}}, \ j = 0, 1, 2. \tag{35}$$

The following Theorem gives the limiting distributions of the above corrected Sargan's tests.

Theorem 5 *Under the Assumptions 1-3 and the null of* $E\Big(W_{li}^{(j)\prime}\Delta u_i\Big) = 0$, *as* $(N, T) \to \infty$, $q_j \to \infty$, $(q_j - k)\,/N \to c_j \in [0, 1)$, $\tilde{J}_{(j)} \xrightarrow{d} N(0, 1)$, *for* $j = 0, 1, 2$.

Theorem 5 can be obtained directly from Theorem 1 of Anatolyev and Gospodinov (2010). Hence, we do not provide detailed proof in this paper. By Theorem 5, the corrected Sargan's tests are well defined with large T. When $(q_j - k)\,/N \to 0$, $\tilde{J}_{(j)}$ and $\hat{J}_{(j)}$ are equivalent.

5 Power properties

In this Section, we discuss the power properties of the tests for serial correlation and the tests of overidentifying restrictions. $m_s^{(j)}$ and $m_{(2,p)}^{(j)}$, $j = 0, 1, 2$, can be used to test the null of no error serial correlation. However, rejecting the null, they may not help to indicate whether the errors are following any forms of serial correlation. Moreover, the rejection of the null of no error serial correlation does not necessarily mean any particular alternatives. Hence, tests for serial correlation can be regarded as misspecification tests. In this Section, we consider three important three misspecifications: serial correlation, slope heterogeneity, and cross-sectional dependence. Particularly, we consider the alternatives of qth-order serial correlation ($MA(q)$ and $AR(q)$) for serial correlation and factor model for cross-sectional dependence.

For simplicity and without loss of generality, we focus on the following AR(1) model without regressors,

$$y_{it} = \rho y_{i,t-1} + \eta_i + u_{it},\ i = 1, \ldots, N,\ t = 2, \ldots, T, \tag{36}$$

where $\eta_i \sim i.i.d.(0, \sigma_\eta^2)$ and $u_{it} \sim i.i.d.(0, \sigma_u^2)$ without misspecification. An asymptotic expansion of $N^{-1/2} \sum_{i=1}^N \hat{v}_{ip}^{(0)}$ around $\hat{\rho} = \rho$, with all cross section average replacing averages of expectations, yields

$$\frac{1}{\sqrt{N}} \sum_{i=1}^N \hat{v}_{ip}^{(0)} = \frac{1}{\sqrt{N}} \sum_{i=1}^N v_{ip} - \bar{\varpi} \bar{\Psi}_N^{(j)-1} \bar{a}_{(0)} \bar{\Omega}_{(0)}^{-1} \frac{1}{\sqrt{NT}} \sum_{i=1}^N W_{li}^{(0)\prime} \Delta u_i + o_p(1),\ j = 0, 1, 2, \tag{37}$$

where $\bar{\varpi} = \left[(T/T_2)^{1/2} \bar{\varpi}_2, \ldots, (T/T_p)^{1/2} \bar{\varpi}_p \right]$ with $\bar{\varpi}_s = (NT_s)^{-1} \sum_{i=1}^N \sum_{t=3}^{T-s} \mathrm{E}(\Delta u_{it} \Delta Z_{i,t+s})$, $s = 2, \ldots, p$; $\bar{\Psi}_N^{(0)-1} = \bar{a}_{(0)}' \bar{\Omega}_{(0)}^{-1} \bar{a}_{(0)}$; $\bar{a}_{(0)} = (NT_s)^{-1} \sum_{i=1}^N \mathrm{E}(W_{li}^{(0)\prime} \Delta y_{i,-1})$; and $\bar{\Omega}_{(0)} = (NT_s)^{-1}$ $\sum_{i=1}^N \mathrm{E}\left(W_{li}^{(0)\prime} \Delta u_i \Delta u_i W_{li}^{(0)} \right)$.

For the Sargan's test, we consider the term $(NT_0)^{-1/2} \sum_{i=1}^N W_{li}^{(0)\prime} \Delta u_i$.

5.1 Power analysis under the local AR(q) and MA(q) errors

Consider an alternative of $MA(q)$ Process,

$$u_{it} = \sum_{l=1}^q \rho_l \varepsilon_{i,t-l} + \varepsilon_{it}, \tag{38}$$

where $|\rho_l| < \infty$. Or an $AR(q)$ process,

$$u_{it} = \sum_{l=1}^{q} \psi_l u_{i,t-l} + \varepsilon_{it}, \tag{39}$$

$\varepsilon_{it} \sim i.i.d.(0, \sigma_\varepsilon^2)$. Consider local versions of $MA(q)$ and $AR(q)$ errors[2], which are $\rho_l = (NT)^{-1/2}\,\delta_l$ in (38) and $\psi_l = (NT)^{-1/2}\,\delta_l$ in (39), $l = 1, 2, \ldots, q$. It is assuming $0 < |\delta_l| < \infty$, and it satisfies stationary condition of u_{it} given N. The sth-order error auto-covariance can be expressed by the parameters σ_ε^2 and δ_l, $l = 1, 2, \ldots, q$, (see Hamilton (1994))

$$r_s = \begin{cases} \sigma_u^2 + o((NT)^{-1/2}), & \text{for } s = 0; \\ \sigma_u^2 \delta_l / (NT)^{-1/2} + o((NT)^{-1/2}), & \text{for } s = 1, 2, \ldots, q; \\ o((NT)^{-1/2}), & \text{for } s > q. \end{cases} \tag{40}$$

Denote the non-central chi-square distribution with n degree of freedom with non-centrality parameter ζ be $\chi^2(n, \zeta)$. Under these local alternatives,

$$m_{(2,p)}^{(0)2} \xrightarrow{d} \chi^2\left(p - 1, \varphi_p^{(0)\prime} V_p^{(0)-1} \varphi_p^{(0)}\right), \tag{41}$$

where $V_p^{(0)}$ is $\underset{(n,T)\to\infty}{\text{plim}} \left(G_p^{(0)\prime} G_p^{(0)\prime}/NT\right)$; $\varphi_p^{(0)} = \underset{(n,T)\to\infty}{\text{plim}} N^{-1/2} \sum_{i=1}^{N} \hat{v}_{ip}^{(0)}$ and φ_p is expressed as

$$\varphi_p^{(0)} = c_p + d_p^{(0)}, \tag{42}$$

[2]Yamagata (2008) shows that MA(q) and AR(q) errors are locally equivalent alternatives.

where

$$c_p = \sigma_u^2 \begin{pmatrix} (2\delta_2 - \delta_3 - \delta_1) \\ \vdots \\ (2\delta_{q-1} - \delta_q - \delta_{q-2}) \\ 2\delta_{q-2} - \delta_{q-1} \\ -\delta_q \\ 0_{p-q-1} \end{pmatrix} \tag{43}$$

and $d_p^{(0)} = \sigma_u^2 \varpi^* \bar{\Psi}_N^{(0)-1} \bar{a}_{(0)} \bar{\Omega}_{(0)}^{-1} e^{(0)}$, where $\varpi^* = (\varpi_2, \ldots, \varpi_p)$ and $e^{(0)} = \underset{(N,T)\to\infty}{\text{plim}} \frac{1}{\sqrt{NT_0}} \sum_{i=1}^N W_{li}^{(0)\prime} \Delta u_i = \underset{(N,T)\to\infty}{\text{plim}} \left(e_1^{(0)}, \ldots, e_\kappa^{(0)}\right)'$, with

$$e_i^{(0)} = \begin{cases} \frac{1}{T} \left[\sum_{h=i}^q \sum_{l=0}^{h-i} \rho^l (\delta_{1+l+i} - \delta_{i+l}) + (T-1-i-q) \left(\sum_{l=0}^{q-i} \rho^l \left(\delta_{i+l+1} - \delta_{i+l}\right) \right) \right] & , \ i \leq q; \\ 0 & , \ i > q. \end{cases}$$

Similar to $m_{(2,p)}^{(0)2}$, we can also consider the power property of $m_{(2,p)}^{(j)2}$, $j = 1, 2$, under the $MA(q)$ or $AR(q)$ local alternative, we have

$$m_{(2,p)}^{(j)2} \xrightarrow{d} \chi^2 \left(p-1, \varphi_p^{(j)\prime} V_p^{(j)-1} \varphi_p^{(j)}\right), \ j = 1, 2, \tag{44}$$

where $V_p^{(j)} = \underset{(n,T)\to\infty}{\text{plim}} \left(G_p^{(j)\prime} G_p^{(j)\prime}/N\right)$, and $\varphi_p^{(j)} = c_p + d_p^{(j)}$ with $d_p^{(j)} = \sigma_u^2 \varpi^* \bar{\Psi}_N^{(j)-1} \bar{a}_{(j)} \bar{\Omega}_{(j)}^{-1} e^{(j)}$, in which $\bar{\Psi}_N^{(j)-1} = \bar{a}_{(j)}' \bar{\Omega}_{(j)}^{-1} \bar{a}_{(j)}$ with $\bar{a}_{(j)} = N^{-1} \sum_{i=1}^N \mathrm{E}(W_{li}^{(j)\prime} \Delta y_{i,-1})$ and $\bar{\Omega}_{(j)} = N^{-1} \sum_{i=1}^N \mathrm{E}\left(W_{li}^{(j)\prime} \Delta u_i \Delta u_i W_{li}^{(j)}\right)$, $j = 1, 2$. Define $q_0' = \kappa$, $q_1' = T - 2$ and $q_2' = \left[(T-1) \times (T-2)\right]/2$. $e^{(1)}$ is a $q_1' \times 1$ vector and $e^{(2)}$ is a $q_2' \times 1$ vector. The ith element in $e^{(1)}$ is defined as

$$e_i^{(1)} = \begin{cases} \frac{1}{\sqrt{T}} \sum_{l=0}^{i-1} \rho^l (\delta_{2+l} - \delta_{1+l}), & i \leq q; \\ \frac{1}{\sqrt{T}} \sum_{l=0}^{q-1} \rho^l \left(\delta_{2+l} - \delta_{1+l}\right), & i > q. \end{cases}$$

For $e^{(2)}$, the number of non-zero elements of order T, each non-zero element in $e^{(2)}$ can be express as $\vartheta_i \left(\rho, \delta_s\right)/\sqrt{T}$, where $\vartheta_i \leq K < \infty$.

Above results show that $m_{(2,p)}^{(j)2}$ are powerful against both $MA(q)$ and $AR(q)$ alternatives. Note that the power properties of $m_{(2,p)}^{(j)2}$ depends on the magnitude of c_p and $d_p^{(j)}$. c_p is due to the asymptotic bias of $N^{-1/2}\sum_{i=1}^{N} v_{iT}$. Its first q elements are non-zeros. $d_p^{(j)}$ is due to the non-zero $e^{(j)}$. Although they have the same non-zero part c_p, $d_p^{(j)}$, $j = 0, 1, 2$, are different. There is no obvious power rank for the three tests.

Next, we consider the power property of $S_{(0)}$. Under the local alternatives, we have

$$S_{(0)} \xrightarrow{d} \chi^2\left(p-k, e^{(0)\prime}V^{(0)-1}e^{(0)}\right), \tag{45}$$

where $V^{(0)} = \text{plim}_{(N,T)\to\infty}\left(\sum_{i=1}^{N}\left(W_{li}^{(0)\prime}\Delta\breve{u}_i^{(0)}\Delta\breve{u}_i^{(0)\prime}W_{li}^{(0)}\right)/NT\right)$. Since $e^{(0)}$ is q_0' dimensional vector with q non-zero elements, then $S_{(0)}$ is powerful against both $MA(q)$ and $AR(q)$ alternatives. It should also be noted that q_0' affect the power property. Since $e^{(0)}$'s first q elements are non-zeros and the left are zeros, $S_{(0)}$ is more powerful when $q_0' = q$ comparing to those when $q_0' \neq q$. Next, we consider the power properties of the corrected Sargan's tests with large q_j, under the local alternatives, we have

$$\tilde{J}_{(j)} \xrightarrow{d} N\left(\lim_{(N,T)\to\infty}\frac{\varsigma_{(j)}}{\sqrt{2\left(q_j'-k\right)\left(1-\left(q_j'-k\right)/N\right)}}, 1\right),\ j = 0, 1, 2, \tag{46}$$

where $\varsigma_{(0)} = e^{(0)\prime}V^{(0)-1}e^{(0)}$ and $\varsigma_{(j)} = e^{(j)\prime}V^{(j)-1}e^{(j)}$ with

$$V^{(j)} = \underset{(N,T)\to\infty}{\text{plim}}\frac{1}{N}\sum_{i=1}^{N}\left(W_{li}^{(j)\prime}\Delta\breve{u}_i^{(j)}\Delta\breve{u}_i^{(j)\prime}W_{li}^{(j)}\right), j = 1, 2.$$

From above results, we have $\varsigma_{(0)} = O_p(1)$; $\varsigma_{(1)} = O_p(1)$ and $\varsigma_{(2)} = O_p(1)$. As $q_j \to \infty$, $\varsigma_{(j)}/\sqrt{2\left(q_j'-k\right)\left(1-\left(q_j'-k\right)/N\right)} \to 0$. Under the local $MA(q)$ and $AR(q)$ alternatives, as $T \to \infty$, $S_{(1)}$ and $S_{(2)}$ have no power asymptotically. Hencer, their power under $MA(q)$ and $AR(q)$ alternatives only increases with N.

It should be worth to notice that under the local alternatives, the tests for serial correlation have non-trivial power. However, the power of the corrected Sargan's tests decrease as the dimension of instrument matrix increases. Hence, test statistics using the collapsed matrix with fixed number of instruments should be more powerful than those using the block-diagonal instrument matrix. Reducing the dimension of instrument matrix can increase the power under the qth-order serial correlation alternative. Therefore, the tests of specification built upon two-step GMM estimation with $W_{li}^{(0)}$ may be more powerful than the tests with $W_{li}^{(1)}$ or $W_{li}^{(2)}$.

5.2 Slope heterogeneity and error cross-sectional dependence

Pesaran and Smith (1995) show that ignoring slope heterogeneity in dynamic panel model may lead to persistent serial correlation in errors. Consider a model (37) with slope heterogeneity

$$y_{it} = \rho_i y_{i,t-1} + \eta_i + u_{it},\ i = 1, \ldots, N,\ t = 2, \ldots, T, \tag{47}$$

where $\rho_i = \rho + \upsilon_i$, $\upsilon_i \sim i.i.d.\left(0, \sigma_\upsilon^2\right)$. If ignoring the slope heterogeneity, the error term of (36) can be rewritten as

$$u_{it} = \upsilon_i y_{i,t-1} + \varepsilon_{it}. \tag{48}$$

From above expression, it is clear that the error term exists serial correlation and this correlation is similar to $AR(1)$ process, hence the power property under this alternative is similar to $AR(1)$ alternative.

Ignorance of cross-sectional dependence in error terms also results in serial correlation. Consider a factor model in error terms,

$$u_{it} = \lambda_i' f_t + \varepsilon_{it}, \tag{49}$$

where $\lambda_i \sim i.i.d.(0, \Sigma_\lambda)$, λ_i and ε_{jt} are not correlated for any i, j, t, and $f_t \sim i.i.d.(0, \Sigma_f)$.

Condition on f_t,

$$\mathrm{E}\left(u_{it}u_{i,t+s}\right) = \mathrm{E}\left[\left(\Delta f_t'\lambda_i + \Delta\varepsilon_{it}\right)\left(\lambda_i'\Delta f_{t+s} + \Delta\varepsilon_{i,t+s}\right)\right] = \Delta f_t'\Sigma_\lambda\Delta f_{t+s}. \tag{50}$$

The magnitude of $\mathrm{E}\left(u_{it}u_{i,t+s}\right)$ does not necessarily decrease as s increases with given t. Consequently, the power of the joint serial correlation test is likely to increase as p increases; and the Sargan's test $S_{(0)}$ is likely to be more powerful when κ is larger.

5.3 Choice of p and κ

In practice, we cannot observe the true alternatives. Since the power properties of the proposed tests depend on the specifications of the alternatives, there is no clear theoretical guidance for selecting the value of p and κ. For the choice of p, one can follow Yamagata's (2008) suggestions as the follows. Choosing p to be equal or slightly larger than $q+1$. If there is reasonable to doubt there is $MA(q)$ or $AR(q)$ error serial correlation; choose p to be its maximum value or close to it on testing for general misspecifications when T is relatively small. For the choice of κ, one can just set $\kappa = p$. Given p and κ, the power of both tests grows as T increases. Therefore, when T is relatively large, we may just fix the values of p and κ comparing to the value of T.

6 Monte Carlo simulations

This section conducts Monte Carlo simulations to examine the empirical size and power of the tests for second-order serial correlation $\left(m_2^{(j)}\right)$, the joint tests for the second to pth serial correlation $\left(m_{(2,p)}^{(j)2}\right)$, the Sargan's tests $\left(S_{(j)}\right)$ and the corrected Sargan's test $\left(\tilde{J}_{(j)}\right)$ in a dynamic panel data model, for $j = 0, 1, 2$. Following Yamagata (2008), we consider six scenarios of misspecifications: MA(1) errors; MA(2) errors; AR(1) errors; AR(2) errors; heterogeneous slopes; and errors with cross-sectional dependence.

6.1 Experimental design

The data generating process (DGP) is given by

$$y_{it} = \rho y_{i,t-1} + \beta x_{it} + \eta_i + u_{it},\ |\rho| < 1;\ i = 1, \ldots, N,\ t = -48, -47, \ldots, T, \tag{51}$$

with

$$x_{it} = \rho_x x_{i,t-1} + \tau u_{i,t-1} + v_{it},\ |\rho_x| < 1;\ i = 1, \ldots, N,\ t = -48, -47, \ldots, T. \tag{52}$$

Let $y_{i,49} = 0$ and $x_{i,49} = 0$. The first 50 observations of both DGPs are discarded. We set $\rho = \rho_x = \beta = \tau = 0.5$. Under the null, $u_{it} = \varepsilon_{it}$, where $\varepsilon_{it} \sim i.i.d.N(0, \sigma_\varepsilon^2)$ with $\sigma_\varepsilon^2 = 1$. $v_{it} \sim i.i.d.N(0, \sigma_v^2)$ and $\eta_i \sim i.i.d.N(0, \sigma_\eta^2)$. Following Kiviet (1995) and Bun and Kiviet (2006), we control the signal-to-noise ratio under the null, $u_{it} = \varepsilon_{it}$ through σ_v^2. The signal-to-noise ratio is defined as $\omega = \sigma_s^2/\sigma_\varepsilon^2$, where σ_s^2 =var$(y_{it}^* - \varepsilon_{it})$ with $y_{it}^* = y_{it} - \eta_i/\left(1-\rho\right)$. Specifically, using the derivations of parameters in Sarafidis et al. (2009), we have

$$\sigma_v^2 = \frac{\left[\sigma_\varepsilon^2(1+\omega)\right]/a_1 - b_1}{\beta^2}, \tag{53}$$

where

$$a_1 = \frac{1+\rho\rho_x}{\left(1-\rho_x^2\right)\left(1-\rho^2\right)\left(1-\rho\rho_x\right)}, \tag{54}$$

$$b_1 = 1 + \left(\beta\tau - \rho_x\right)^2 + \frac{2\left(\beta\tau - \rho_x\right)\left(\rho + \rho_x\right)}{1+\rho\rho_x}. \tag{55}$$

Set $\omega = 3$ and choose $\sigma_\eta^2 = (1-\rho^2)a_1 b_1$.

For power property, we consider six different error specifications, denoted by (a)-(f). All the parameters are controlled to set the variance of u_{it} to be 1.

(a) MA(1) error model,

$$u_{it} = \sigma_\varepsilon(\varepsilon_{it} + \psi_1\varepsilon_{i,t-1}), \tag{56}$$

where $\psi_1 = 0.2$ and $\sigma_\varepsilon^2 = 1/(1+\psi_1^2)$; so that $r_0 = 1$ and $r_1 = 0.2$.

(b) MA(2) error model:

$$u_{it} = \sigma_\varepsilon(\varepsilon_{it} + \psi_1\varepsilon_{i,t-1} + \psi_2\varepsilon_{i,t-2}), \tag{57}$$

where $\psi_1 = 20/103$, $\psi_2 = 13/90$. $\sigma_\varepsilon^2 = 1/(1 + \psi_1^2 + \psi_2^2)$; so that $r_0 = 1$, $r_1 = 2/9$ and $r_2 = 13/90$.

(c) AR(1) error model:

$$u_{it} = \rho_1 u_{i,t-1} + \sigma_\varepsilon\varepsilon_{it}, \tag{58}$$

where $\rho_1 = 0.2$ and $\sigma_\varepsilon^2 = (1 - \rho_1^2)$; so that $r_0 = 1$ and $r_1 = 0.2$.

(d) AR(2) error model:

$$u_{it} = \rho_1 u_{i,t-1} + \rho_2 u_{i,t-2} + \sigma_\varepsilon\varepsilon_{it}, \tag{59}$$

where $\rho_1 = 0.2$, $\rho_2 = 0.1$ and $\sigma_\varepsilon^2 = (1+\rho_2)\left[(1 - \rho_2)^2 - \rho_1^2\right] / (1 - \rho_2)$; so that $r_0 = 1$, $r_1 = 2/9$ and $r_2 = 13/90$.

(e) Heterogeneous slopes: equation (51) is replaced by

$$y_{it} = \rho y_{i,t-1} + \beta_i x_{it} + \eta_i + u_{it}, \tag{60}$$

where $\beta_i \sim i.i.d.N(0.5, 1)$.

(f) Error with cross-sectional dependence:

$$u_{it} = 0.75\left(\lambda_i' f_t + \sigma_\varepsilon\varepsilon_{it}\right), \tag{61}$$

where $\lambda_i \sim i.i.d.U[-1, 1]$, $f_t \sim i.i.d.N(0, 1)$ and $\sigma_\varepsilon^2 = 1$.

The experiments are conducted for $N = 100, 150, 200$ and $T = 7, 11, 20, 30$. For all pair of (N, T), we run $2,000$ replications. To obtain the empirical size, we conduct $m_2^{(j)}$ and $\tilde{J}_{(j)}$, $j = 0, 1, 2$, at the two-sided 5% nominal significance level, conduct $m_{(2,p)}^{(j)2}$ and $S_{(j)}$, $j = 0, 1, 2$;

at the positive one-side 5% nominal significance level. p is set to be $3, 4, 5, 6, 7$ and κ is set to be $2, 3, 4, 5, 6, 7$.

6.2 Simulation results

Table 1 reports the empirical size of each test. The results show that all $m_2^{(j)}$ and $m_{(2,p)}^{(j)2}$, $j = 0, 1, 2$, have correct size with different (N, T) combinations. Using the χ^2 approximation, the Sargan's test $S_{(0)}$ perform correct size, while $S_{(j)}$, $j = 1, 2$, tend to reject the null infrequently. When T is moderately large, for example, $T = 11$ and $N = 100$, those tests are both extremely undersized. The size of $\tilde{J}_{(2)}$ and $\tilde{J}_{(1)}$ are 0 and 0.4, respectively. It is consistent with the simulation results in previous literature. $S_{(1)}$ and $S_{(2)}$ have more size distortions when T increases. On the contrary, the corrected Sargan's tests $\tilde{J}_{(j)}$ perform much better size compared to $S_{(j)}$, for $j = 1, 2$. For example, when $N = 100, T = 11$, the size of $\tilde{J}_{(2)}$ and $\tilde{J}_{(1)}$ are 3 and 3.4, which are slightly undersized but close to 5.

The rest of tables contain the results of power and size-adjusted power in parentheses for each tests under varieties of alternatives. Table 2 and 3 give the power results under the MA(1) alternative. In terms of size-adjusted power, the results show that $m_{(2,p)}^{(j)2}$ are often superior to $m_2^{(j)}$, $j = 0, 1, 2$. Among $m_{(2,p)}^{(j)2}$, $j = 0, 1, 2$, $m_{(2,p)}^{(0)2}$ has the largest power, but their differences are not significant. $\tilde{J}_{(0)}$ with specified κ performs good power properties. As (N, T) being large, the power of $m_2^{(j)}$, $m_{(2,p)}^{(j)2}$, $j = 0, 1, 2$, and $\tilde{J}_{(0)}$ are all close to 1. On the other hand, $\tilde{J}_{(2)}$ has very low power for small N and large T and $\tilde{J}_{(1)}$ has extremely low power for all cases. Even for relatively large N given T fixed, the power of $\tilde{J}_{(1)}$ does not increase significantly. Both $\tilde{J}_{(2)}$ and $\tilde{J}_{(1)}$ have much smaller power than $m_{(2,p)}^{(j)2}$ and $\tilde{J}_{(0)}$. $m_2^{(0)}, m_{(2,p)}^{(0)2}$ and $\tilde{J}_{(0)}$ have less power as κ increases, which also confirms our theoretical analysis.

Table 4 and 5 report the power under the MA(2) alternative. $m_2^{(j)}$ have much smaller power compared to the $m_{(2,p)}^{(j)2}$, $j = 0, 1, 2$, which is consistent to the results of Yamagata (2008). $\tilde{J}_{(2)}$ has low power when N is small and T is moderately large. $\tilde{J}_{(1)}$ still has extremely low power for all combinations. $\tilde{J}_{(0)}$ has reasonable power, but there is no obvious rank compared to $m_{(2,p)}^{(j)2}$, $j = 0, 1, 2$.

Table 6 and 7 report the power results under the AR(1) alternative. The results are similar to the case of MA(1). Note that all the tests have more power than the case of MA(1) given the same N and T. This result is because, in the case of AR(1) errors, all the elements of the bias terms are non-zeros, whereas, in the cases of MA(1), only the first two parts of the bias terms are non-zeros. Table 8 and 9 report the results of power under the AR(2) scenario. $m_2^{(j)}$, $j = 0, 1, 2$, have very low power. It is caused by the choice of parameters in auto-regressive errors. More specifically, in this case, $\mathrm{E}(\Delta u_{it} \Delta u_{i,t+2}) = 0.07$ and $\mathrm{E}(\Delta u_{it} \Delta u_{i,t+3}) = -0.14$; therefore, they tends to be less powerful than $m_{(2,p)}^{(j)2}$, also see Yamagata (2008). Among $\tilde{J}_{(0)}$ with different κ, the test with $\kappa = 3$ dominates all the others. Moreover, $\tilde{J}_{(1)}$ still have low power for all the pairs of (N, T).

The power results under slope heterogeneity alternative are reported in table 10 and 11. $\tilde{J}_{(0)}$ with $\kappa = 2$ dominates all the other tests for all the combinations of N and T. $m_{(2,p)}^{(j)2}$, are superior to $m_2^{(j)}$, $j = 0, 1, 2$; for most of the (N, T) combinations. The power results under the cross sectional dependence alternative are specified in table 12 and 13. The power of all the tests monotonically increases as p and κ increase. From the results, we can observe that given a relatively large p and small κ, $m_{(2,p)}^{(j)2}$, $j = 0, 1, 2$, have more power than $\tilde{J}_{(0)}$ for most of the (N, T) combinations. However, it flips over with small p and large κ in some cases. $\tilde{J}_{(2)}$ has more power than $\tilde{J}_{(0)}$ for small T; when T increases, it becomes less powerful than $\tilde{J}_{(0)}$ with slightly large κ. $\tilde{J}_{(1)}$ has much more power compared to itself under the other alternatives, but it is still less powerful than the other tests for most of the cases.

Overall, several interesting findings are obtained by the simulation studies: first, the corrected Sargan's tests $\left(\tilde{J}_{(j)},\ j = 1, 2\right)$ have reasonable size compared to the uncorrected Sargan's tests $\left(S_{(j)},\ j = 1, 2\right)$; second, using the collapsed instruments matrices, the performance of each test dominates those using the block-diagonal instrument matrices for most of the (N, T) combinations; the exception is only for very small T; third, there is no absolute ranking between the $\tilde{J}_{(0)}$ and $m_{(2,p)}^{(j)2}$, $j = 0, 1, 2$, but both of them are often superior to $m_2^{(j)}$.

7 Concluding remarks

In this paper, we consider the tests of specification in dynamic panel data models with large N and large T. All the test statistics are built upon the two-step GMM estimation using three different instrument matrices: the block-diagonal matrix with the full set of all available instruments, the block-diagonal matrix with a subset of all available instruments and the collapsed instrument matrix. This paper shows the conventional Sargan's test ($S_{(2)}$) does not approximate to the chi-square distribution when the number of instruments used is relatively as large as N. Therefore, it proposes corrected Sargan's tests $\left(\tilde{J}_{(j)},\ j = 0, 1, 2\right)$ with the three instrument matrices. It extends the conventional tests for serial correlation to large N and large T framework. The asymptotics of all the tests are well established as N and T go to infinity jointly. This paper also adds power analysis under different alternatives. Under local $MA(q)$ and $AR(q)$ alternatives, it shows that $m^{(j)2}_{(2,p)}$, $j = 0, 1, 2$, have good power properties and the power of $\tilde{J}_{(j)}$, $j = 1, 2$, decreases as the number of instrument increases, which means the power of $\tilde{J}_{(j)}$, $j = 1, 2$, only increases with N, while the power of $\tilde{J}_{(0)}$ with fixed number of instruments increases when either N, T or both increase.

The Monte Carlo simulations are conducted for studying the small sample properties of all the tests with three different instrument matrices. The simulation results show several interesting facts: in the first place, the corrected Sargan's tests have correct size, while the conventional Sargan's tests suffer serious size distortions; second, the corrected Sargan's test $\tilde{J}_{(0)}$ has reasonable power under varieties of alternatives; especially, $\tilde{J}_{(0)}$ with just specified lag instruments ($\kappa = q + 1$) matrix mostly outperform the other tests of specification; $\tilde{J}_{(1)}$ has very lower power for finite samples for all the alternatives except cross-sectional dependence; and last but not the least, there is no clear rank between $m^{(j)2}_{(2,p)}$, $j = 0, 1, 2$, and $\tilde{J}_{(0)}$, but both of them dominate $m^{(j)2}_{2}$, $j = 0, 1, 2$, for most of the cases. Among $m^{(j)2}_{(2,p)}$, $j = 0, 1, 2$, $m^{(0)2}_{(2,p)}$ has the largest power, but their differences are not significant.

For practical purpose, we suggest using the collapsed instrument matrix $\left(W^{(0)}_{li}\right)$ for testing since they have several advantages. First, they do not have the invert-ability issue of large covariance matrix when T large and are computationally easy to implement; second, they have much more power compared to the tests using block-diagonal instrument matrix $\left(W_{li}^{(1)} \text{ and } W_{li}^{(2)}\right)$. It is worth to remark that there is no technical standards to choose the number of p for the joint serial correlation tests and κ for the collapsed instruments matrix. The power properties imply they depend on the actual misspecifications.

Acknowledgments

Essay III is based on the working paper of Lee and Peng (2016).

References

[1] Ahn, S.C., Schmidt, P., 1995. Efficient estimation of models for dynamic panel data. *Journal of Econometrics* 68, 5-27.

[2] Anatolyev, S., Gospodinov, N., 2011. Specification Testing in models with many instruments. *Econometric Theory* 27, 427-441

[3] Alvarez, J., Arellano, M., 2003, The time series and cross-section asymptotics of dynamic panel data estimators. *Econometrica* 71(4), 1121-1159.

[4] Arellano, M., 2003. *Panel Data Econometrics*. Oxford University Press.

[5] Arellano, M., Bond, S., 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies* 24, 1-37.

[6] Arellano, M., Bover, O., 1995. Another look at the instrumental variable estimation of error components models. *Journal of Econometrics* 68, 29-51.

[7] Baltagi, B.H., 2008. *Econometric Analysis of Panel Data*, New York: Wiley.

[8] Blundell, R.W., Bond, S.R., 1998. Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics* 87, 115-143.

[9] Bowsher, C.G., 2002. On Testing overidentifying restrictions in dynamic panel data models. *Economics Letters* 77, 211-220.

[10] Bun, M.J.G, Kiviet, J.F., 2006. The effects of dynamic feedbacks on LS and MM estimator accuracy in panel data models. *Journal of Econometrics* 132, 409-444.

[11] Chao, J.C., Hausman, J.A., Newey, W.K., Swanson, N.R., Woutersen, T. 2014. Testing overidentifying restrictions with many instruments and heteroskedasticity. *Journal of Econometrics* 178, 15-21.

[12] Donald, S.G., Imbens, G.W., Newey, W.K. 2003. Empirical likelihood estimation and consistent tests with conditional moment restrictions. *Journal of Econometrics* 117, 55-93.

[13] Hamilton, J.D., 1994. *Time Series Analysis*. Princeton University Press, Princeton.

[14] Hansen, L. 1982. Large sample properties of generalized method of moments estimators. *Econometrica* 50, 1371-1395.

[15] Hausman, J.A., 1978. Specification tests in econometrics. *Econometrica* 46, 1251-1272.

[16] Hayakawa, K., 2015. The asymptotic properties of the system GMM estimator in dynamic panel data models when both N and T are large. *Econometric Theory* 31(3), 647-667.

[17] Hayakawa, K., 2015. On the effect of weighting matrix in GMM specification test. *Working paper.*

[18] Holtz-Eakin, D., Newey, W.K., Rosen, H.S., 1988. Estimating vector autoregressions with panel data. *Econometrica* 56, 1371-1395.

[19] Hsiao, C., Zhang, J., 2015. IV, GMM or likelihood approach to estimate dynamic panel models when either N or T or both are large. *Journal of Econometrics*,

[20] Hsiao, C., Zhou, Q., 2015. First difference or forward demeaning: implications for the method of moment estimators. *Working paper.*

[21] Kiviet, J.F., 1995. On bias, inconsistency, and efficiency of various estimators in dynamic panel data models. *Journal of Econometrics* 68, 53-78.

[22] Lee, Y., Okui, R., 2012. Hahn-Hausman test as a specification test. *Journal of Econometrics* 167, 133-139.

[23] Lee, Y., Peng, B., 2016. Tests of Specification for Large Dynamic Panel Data Models. *Working Paper.*

[24] Newey, W.K., 1985. Generalized method of moments specification testing. *Journal of Econometrics* 29, 229-256.

[25] Pesaran, M.H., Smith, R., 1995. Estimating long-run relationships from dynamic heterogeneous panels. *Journal of Econometrics* 29, 229-256.

[26] Pesaran, M.H., Yamagata, T., 2008. Testing slope homogeneity in large panels. *Journal of Econometrics* 142, 50-93.

[27] Phillips, P.C.B, Moon, H.R., 1999. Linear regression limit theory for nonstationary panel data. *Econometrica* 67, 1057-1110.

[28] Roodman, D., 2009. A note on the theme of too many instruments. *Oxford Bulletin of Economics and Statistics* 71, 135-158.

[29] Sarafidis, V., Yamagata, T., Robertson, D., 2009. A test of cross section dependence for a linear dynamic panel data with regressors. *Journal of Econometrics* 148, 149-161.

[30] Sargan D., 1958. The estimation of economic relationships using instrumental variables. *Econometrica* 26, 393-415.

[31] Windmeijer, F., 2005. A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of Econometrics* 126, 25-51.

[32] Wood, A.T.A., 1999. Rosenthal's inequality for point process martingales. *Stochastic Processes and their Applications* 81, 231-246.

[33] Yamagata, T. 2008. A joint serial correlation test for linear panel data models. Journal of Econometrics 146, 135-145.

Table 1: Size of Tests

		Type of Instrument Matrix/Number of Instruments									Type of Instrument Matrix/Number of Instruments						
		$j=2$	$j=1$	$j=0$							$j=1$	$j=0$					
Tests	(N,T)	T(T-2)	2(T-2)	4	6	8	10	12	14	(N,T)	2(T-2)	4	6	8	10	12	14
$m_2^{(j)}$	(100,7)	6.25	5.95	5.90	6.00	5.90	6.15			(100,20)	4.65	5.05	4.95	5.00	4.90	4.95	4.95
$m_{(2,3)}^{(j)2}$	(100,7)	5.40	5.70	5.30	5.30	5.25	5.20			(100,20)	5.25	5.20	5.20	5.30	5.30	5.20	5.25
$m_{(2,4)}^{(j)2}$	(100,7)	5.55	4.80	4.55	4.25	4.40	4.45			(100,20)	4.80	4.85	4.75	4.80	4.75	4.65	4.80
$m_{(2,5)}^{(j)2}$	(100,7)									(100,20)	4.45	4.95	4.85	4.85	4.90	4.90	4.85
$m_{(2,6)}^{(j)2}$	(100,7)									(100,20)	4.30	4.55	4.45	4.50	4.50	4.45	4.45
$m_{(2,7)}^{(j)2}$	(100,7)									(100,20)	3.95	4.20	4.25	4.15	4.20	4.25	4.25
$S_{(j)}$	(100,7)	2.70	4.15	4.85	4.40	4.25	4.30			(100,20)	1.30	4.50	4.20	4.20	4.20	4.30	4.20
$\tilde{J}_{(j)}$	(100,7)	4.70	4.20	5.05	4.50	4.30	4.30			(100,20)	3.25	4.65	4.20	4.20	4.15	4.45	4.45
$m_2^{(j)}$	(150,7)	4.55	4.35	4.35	4.25	4.30	4.35			(150,20)	5.25	5.10	5.15	5.15	5.15	5.10	5.10
$m_{(2,3)}^{(j)2}$	(150,7)	4.40	4.45	4.25	4.25	4.30	4.10			(150,20)	4.85	4.90	4.80	4.80	4.80	4.80	4.80
$m_{(2,4)}^{(j)2}$	(150,7)	4.95	4.50	4.20	4.20	4.30	4.50			(150,20)	4.85	4.70	4.65	4.70	4.65	4.65	4.70
$m_{(2,5)}^{(j)2}$	(150,7)									(150,20)	4.30	4.80	4.65	4.65	4.60	4.60	4.55
$m_{(2,6)}^{(j)2}$	(150,7)									(150,20)	5.20	5.65	5.85	5.85	5.85	6.00	5.85
$m_{(2,7)}^{(j)2}$	(150,7)									(150,20)	5.45	5.50	5.25	5.25	5.20	5.25	5.20
$S_{(j)}$	(150,7)	2.80	4.95	4.65	4.80	5.90	4.75			(150,20)	2.75	5.45	4.55	4.75	4.15	4.25	4.65
$\tilde{J}_{(j)}$	(150,7)	4.05	4.80	4.90	4.80	5.75	4.45			(150,20)	4.10	5.65	4.65	4.55	3.85	4.00	4.50
$m_2^{(j)}$	(200,7)	5.30	5.35	5.50	5.45	5.40	5.40			(200,20)	5.60	5.70	5.70	5.75	5.75	5.75	5.75
$m_{(2,3)}^{(j)2}$	(200,7)	5.20	5.30	5.00	4.85	5.00	4.90			(200,20)	5.85	6.25	6.00	6.15	6.15	6.15	6.15
$m_{(2,4)}^{(j)2}$	(200,7)	5.45	4.75	4.80	4.75	4.60	4.75			(200,20)	5.95	5.80	5.85	5.75	5.80	5.65	6.00
$m_{(2,5)}^{(j)2}$	(200,7)									(200,20)	5.65	5.80	6.05	6.05	6.05	6.05	5.80
$m_{(2,6)}^{(j)2}$	(200,7)									(200,20)	4.95	5.45	5.55	5.40	5.40	5.40	5.40
$m_{(2,7)}^{(j)2}$	(200,7)									(200,20)	5.40	6.35	5.90	5.80	5.75	5.75	5.85
$S_{(j)}$	(200,7)	4.50	5.25	4.95	4.85	4.85	4.30			(200,20)	2.65	4.35	4.05	4.60	5.05	4.75	6.00
$\tilde{J}_{(j)}$	(200,7)	5.55	4.95	5.20	4.85	4.50	4.10			(200,20)	3.35	4.45	4.05	3.90	4.75	4.45	5.60
$m_2^{(j)}$	(100,11)	4.85	5.05	5.00	5.00	4.95	5.00	5.00	5.00	(100,30)	4.50	4.65	4.70	4.70	4.65	4.75	4.75
$m_{(2,3)}^{(j)2}$	(100,11)	4.35	4.65	4.50	4.40	4.45	4.50	4.55	4.55	(100,30)	4.55	4.90	4.85	4.95	4.95	4.95	4.90
$m_{(2,4)}^{(j)2}$	(100,11)	5.20	5.45	5.00	5.00	5.10	5.10	5.10	5.00	(100,30)	4.25	4.55	4.40	4.35	4.45	4.45	4.35
$m_{(2,5)}^{(j)2}$	(100,11)	5.05	4.70	5.05	5.05	5.05	5.00	5.05	5.00	(100,30)	4.20	4.40	4.35	4.35	4.30	4.45	4.40
$m_{(2,6)}^{(j)2}$	(100,11)	5.00	4.80	4.80	4.85	4.75	4.80	4.70	4.80	(100,30)	4.10	4.60	4.60	4.75	4.70	4.75	4.90
$m_{(2,7)}^{(j)2}$	(100,11)	5.95	5.20	5.50	5.45	5.55	5.55	5.60	5.45	(100,30)	4.50	5.65	5.55	5.50	5.35	5.45	5.50
$S_{(j)}$	(100,11)	0.00	0.40	5.10	4.50	3.90	3.85	4.25	4.80	(100,30)	0.10	5.20	4.20	4.00	4.25	4.40	4.50
$\tilde{J}_{(j)}$	(100,11)	3.00	3.70	5.65	4.65	3.90	3.80	4.25	5.15	(100,30)	3.45	5.65	4.40	4.00	4.20	4.65	5.20
$m_2^{(j)}$	(150,11)	5.55	5.20	5.45	5.45	5.45	5.40	5.35	5.35	(150,30)	4.55	4.65	4.70	4.60	4.55	4.55	4.55
$m_{(2,3)}^{(j)2}$	(150,11)	4.55	4.55	4.60	4.45	4.50	4.55	4.60	4.65	(150,30)	4.70	4.90	4.90	4.90	5.10	5.10	5.15
$m_{(2,4)}^{(j)2}$	(150,11)	5.15	5.00	5.40	5.25	5.20	5.20	5.20	5.20	(150,30)	4.05	4.55	4.50	4.35	4.40	4.30	4.30
$m_{(2,5)}^{(j)2}$	(150,11)	4.20	4.20	4.05	3.95	3.95	3.95	4.15	4.25	(150,30)	4.10	4.50	4.40	4.50	4.40	4.70	4.55
$m_{(2,6)}^{(j)2}$	(150,11)	4.45	4.40	4.35	4.40	4.40	4.40	4.40	4.55	(150,30)	4.00	4.40	4.35	4.25	4.30	4.35	4.25
$m_{(2,7)}^{(j)2}$	(150,11)	4.85	4.55	4.45	4.50	4.45	4.35	4.60	4.60	(150,30)	4.05	4.55	4.60	4.35	4.25	4.10	4.20
$S_{(j)}$	(150,11)	0.05	0.70	4.70	4.65	4.35	4.75	4.40	4.95	(150,30)	1.75	4.10	4.90	3.80	3.85	3.75	4.40
$\tilde{J}_{(j)}$	(150,11)	4.50	4.10	4.95	4.70	4.30	4.45	4.15	4.85	(150,30)	4.45	4.15	4.90	3.75	3.55	3.70	4.30
$m_2^{(j)}$	(200,11)	4.75	4.95	5.00	4.95	4.95	4.95	4.90	4.90	(200,30)	5.15	5.20	5.05	5.05	5.15	5.15	5.15
$m_{(2,3)}^{(j)2}$	(200,11)	4.55	4.90	4.70	4.80	4.75	4.75	4.70	4.65	(200,30)	4.85	4.85	4.80	4.75	4.80	4.95	4.85
$m_{(2,4)}^{(j)2}$	(200,11)	5.10	5.15	5.35	5.15	5.10	5.10	5.10	5.20	(200,30)	4.75	5.00	4.85	4.80	4.80	4.75	4.75
$m_{(2,5)}^{(j)2}$	(200,11)	4.95	5.05	5.00	4.90	4.95	4.85	4.90	4.95	(200,30)	4.60	4.85	4.60	4.70	4.75	4.65	4.80
$m_{(2,6)}^{(j)2}$	(200,11)	4.40	4.15	4.20	4.25	4.25	4.15	4.15	4.25	(200,30)	4.90	5.05	4.95	5.00	5.05	5.05	5.00
$m_{(2,7)}^{(j)2}$	(200,11)	4.50	4.25	4.35	4.20	4.10	4.00	4.05	4.05	(200,30)	4.45	4.60	4.60	4.40	4.45	4.60	4.50
$S_{(j)}$	(200,11)	1.05	0.80	5.80	4.30	5.15	5.40	5.05	4.85	(200,30)	2.35	5.20	5.30	4.90	5.35	4.85	5.00
$\tilde{J}_{(j)}$	(200,11)	5.45	4.20	6.25	4.30	5.05	4.85	4.50	4.70	(200,30)	4.05	5.30	5.30	4.75	5.00	4.75	5.25

Notes: This table reports the size of $m_s^{(j)}$, $m_{(2,p)}^{(j)2}$, $S_{(j)}$ and $\tilde{J}_{(j)}$; for $j=0,1,2$; $p=3,4,5,6,7$. For $j=0$, we choose $\kappa=2,3,4,5,6,7$. The data is generated as $y_{it}=\alpha_i+0.5y_{i,t-1}+0.5x_{it}+u_{it}$, where $u_{it}=\varepsilon_{it}\sim i.i.d.N(0,1)$; $x_{it}=0.5x_{i,t-1}+0.5u_{i,t-1}+v_{it}$, $v_{it}\sim i.i.d.N(0,\sigma_v^2)$, $i=1,\dots,N$; $t=-48,\dots,T$, with $y_{i,-49}=x_{i,-49}=0$. The first 50 observations are discarded. The signal-to-noise ratio is fixed to be 3 through σ_v^2. The tests are conducted at the 5% nominal significance level.

Table 2: Power (Size-Adjusted Power) of Tests: MA(1)

		Type of Instrument Matrix/Number of Instruments							
		$j=2$	$j=1$	$j=0$					
Tests	(N,T)	T(T-2)	2(T-2)	4	6	8	10	12	14
$m_2^{(j)}$	(100,7)	45.90 (36.05)	44.30 (35.75)	46.35 (36.90)	46.35 (36.80)	46.30 (36.90)	46.15 (37.10)		
$m_{(2,3)}^{(j)2}$	(100,7)	45.05 (42.15)	36.50 (34.00)	42.05 (40.90)	42.65 (41.65)	42.75 (41.80)	42.65 (42.15)		
$m_{(2,4)}^{(j)2}$	(100,7)	42.95 (39.70)	33.10 (34.10)	37.65 (40.55)	38.20 (40.80)	38.45 (40.70)	38.55 (40.10)		
$S_{(j)}$	(100,7)	11.40 (19.70)	4.70 (5.40)	42.70 (43.05)	32.90 (35.40)	27.55 (29.55)	23.30 (26.45)		
$\tilde{J}_{(j)}$	(100,7)	15.80 (10.80)	4.70 (5.15)	43.55 (32.00)	33.50 (25.30)	27.55 (20.85)	23.25 (17.65)		
$m_2^{(j)}$	(150,7)	62.95 (61.75)	60.55 (60.75)	62.60 (62.80)	62.50 (62.45)	62.50 (62.50)	62.55 (62.25)		
$m_{(2,3)}^{(j)2}$	(150,7)	63.75 (65.50)	57.50 (60.25)	61.55 (64.25)	61.95 (64.40)	61.85 (65.40)	61.95 (64.85)		
$m_{(2,4)}^{(j)2}$	(150,7)	63.60 (64.35)	55.75 (58.25)	59.65 (62.40)	60.00 (63.40)	60.05 (63.35)	59.95 (63.15)		
$S_{(j)}$	(150,7)	30.90 (37.40)	4.85 (9.95)	58.90 (60.20)	50.85 (51.80)	44.45 (40.90)	37.50 (38.20)		
$\tilde{J}_{(j)}$	(150,7)	32.10 (27.90)	4.80 (4.70)	59.70 (49.25)	51.00 (38.50)	44.10 (32.15)	36.95 (29.10)		
$m_2^{(j)}$	(200,7)	73.60 (65.40)	71.55 (65.45)	72.80 (66.75)	73.10 (66.40)	73.05 (66.50)	73.10 (66.70)		
$m_{(2,3)}^{(j)2}$	(200,7)	77.95 (77.75)	73.05 (72.20)	76.10 (76.15)	76.50 (76.85)	76.70 (76.70)	76.60 (77.10)		
$m_{(2,4)}^{(j)2}$	(200,7)	79.15 (77.70)	72.25 (73.35)	76.15 (76.80)	76.35 (76.80)	76.45 (76.75)	76.30 (77.05)		
$S_{(j)}$	(200,7)	48.20 (49.50)	4.55 (9.45)	71.85 (71.85)	63.10 (63.10)	57.05 (57.60)	50.55 (53.95)		
$\tilde{J}_{(j)}$	(200,7)	47.50 (39.80)	4.35 (5.35)	72.35 (62.00)	63.10 (52.30)	56.40 (49.95)	49.50 (42.65)		
$m_2^{(j)}$	(100,11)	77.85 (73.50)	76.15 (73.60)	78.05 (74.70)	77.80 (74.35)	77.90 (74.10)	78.00 (74.50)	78.10 (73.80)	78.10 (73.80)
$m_{(2,3)}^{(j)2}$	(100,11)	83.25 (85.15)	79.75 (80.50)	82.65 (83.95)	83.20 (84.40)	83.00 (84.15)	83.05 (84.55)	83.15 (84.55)	83.10 (84.20)
$m_{(2,4)}^{(j)2}$	(100,11)	84.40 (83.50)	80.05 (78.65)	83.05 (83.05)	83.65 (83.55)	83.45 (83.30)	83.65 (83.30)	83.50 (83.45)	83.60 (83.65)
$m_{(2,5)}^{(j)2}$	(100,11)	84.10 (84.00)	78.15 (78.25)	81.55 (81.35)	82.15 (82.15)	82.40 (82.00)	82.70 (82.65)	82.75 (82.50)	82.75 (82.85)
$m_{(2,6)}^{(j)2}$	(100,11)	82.55 (82.55)	77.20 (78.45)	79.50 (80.20)	80.15 (81.80)	80.95 (82.20)	81.10 (82.10)	80.95 (81.85)	80.75 (81.35)
$m_{(2,7)}^{(j)2}$	(100,11)	80.50 (78.20)	75.00 (74.00)	77.35 (75.85)	77.80 (76.60)	78.40 (77.00)	78.75 (77.05)	78.95 (77.25)	79.00 (76.90)
$S_{(j)}$	(100,11)	0.00 (7.30)	3.80 (6.10)	83.45 (82.75)	76.15 (78.10)	69.50 (72.90)	64.15 (67.75)	57.55 (60.95)	52.50 (52.70)
$\tilde{J}_{(j)}$	(100,11)	1.40 (4.00)	4.55 (5.40)	83.90 (74.90)	76.45 (66.70)	69.50 (62.30)	64.10 (55.85)	57.50 (47.80)	52.55 (39.75)
$m_2^{(j)}$	(150,11)	92.50 (89.95)	91.50 (90.00)	92.40 (90.65)	92.45 (90.70)	92.50 (90.80)	92.55 (90.70)	92.55 (90.95)	92.55 (90.95)
$m_{(2,3)}^{(j)2}$	(150,11)	95.45 (96.00)	94.05 (94.65)	95.35 (95.50)	95.30 (95.85)	95.30 (95.90)	95.35 (95.95)	95.40 (95.95)	95.35 (95.90)
$m_{(2,4)}^{(j)2}$	(150,11)	96.55 (96.50)	95.30 (95.30)	96.45 (96.15)	96.70 (96.40)	96.60 (96.60)	96.60 (96.25)	96.55 (96.35)	96.55 (96.35)
$m_{(2,5)}^{(j)2}$	(150,11)	96.90 (97.20)	95.40 (96.15)	96.30 (96.55)	96.45 (96.85)	96.60 (96.85)	96.50 (96.95)	96.55 (97.00)	96.55 (96.95)
$m_{(2,6)}^{(j)2}$	(150,11)	96.40 (96.70)	95.15 (95.40)	95.85 (96.00)	95.85 (96.20)	95.95 (96.30)	95.90 (96.20)	95.90 (96.25)	95.95 (96.40)
$m_{(2,7)}^{(j)2}$	(150,11)	96.30 (96.30)	94.70 (95.10)	95.55 (95.90)	95.70 (96.10)	95.85 (96.35)	95.85 (96.25)	95.90 (96.25)	95.95 (96.20)
$S_{(j)}$	(150,11)	1.95 (25.20)	4.15 (4.85)	95.55 (95.65)	92.35 (92.95)	90.05 (91.10)	87.00 (87.30)	83.15 (85.10)	79.05 (79.50)
$\tilde{J}_{(j)}$	(150,11)	20.20 (16.65)	4.45 (4.90)	95.65 (91.70)	92.45 (89.30)	89.90 (83.80)	86.60 (81.35)	82.40 (75.45)	78.70 (70.80)
$m_2^{(j)}$	(200,11)	97.05 (96.45)	96.50 (95.75)	97.15 (96.20)	97.10 (96.30)	97.10 (96.35)	97.10 (96.35)	97.00 (96.35)	97.00 (96.35)
$m_{(2,3)}^{(j)2}$	(200,11)	98.65 (98.65)	98.15 (98.20)	98.40 (98.65)	98.50 (98.70)	98.55 (98.60)	98.55 (98.65)	98.55 (98.65)	98.55 (98.65)
$m_{(2,4)}^{(j)2}$	(200,11)	98.95 (98.90)	98.50 (98.40)	98.85 (98.75)	98.85 (98.85)	98.85 (98.85)	98.90 (98.85)	98.90 (98.85)	98.85 (98.85)
$m_{(2,5)}^{(j)2}$	(200,11)	99.15 (99.15)	98.65 (98.60)	98.85 (98.85)	98.85 (98.85)	98.85 (98.90)	98.95 (99.00)	98.95 (99.00)	98.95 (98.95)
$m_{(2,6)}^{(j)2}$	(200,11)	99.45 (99.50)	99.00 (99.15)	99.15 (99.35)	99.20 (99.35)	99.25 (99.35)	99.30 (99.35)	99.25 (99.40)	99.25 (99.40)
$m_{(2,7)}^{(j)2}$	(200,11)	99.30 (99.50)	98.75 (98.85)	99.15 (99.40)	99.30 (99.45)	99.25 (99.40)	99.30 (99.40)	99.30 (99.40)	99.30 (99.40)
$S_{(j)}$	(200,11)	28.20 (48.95)	5.25 (7.15)	98.10 (97.95)	97.75 (98.00)	97.10 (96.95)	95.65 (95.45)	94.00 (94.00)	92.20 (92.55)
$\tilde{J}_{(j)}$	(200,11)	46.10 (34.10)	5.20 (5.85)	98.25 (96.90)	97.75 (95.75)	97.10 (93.80)	95.45 (92.65)	93.50 (89.85)	91.65 (86.70)

Notes: This table reports the power (size-adjusted power) of $m_s^{(j)}$, $m_{(2,p)}^{(j)2}$, $S_{(j)}$ and $\tilde{J}_{(j)}$ under the MA(1) alternative; for $j = 0,1,2$; $p = 3,4,5,6,7$. For $j = 0$, we choose $\kappa = 2,3,4,5,6,7$. The data is generated as $y_{it} = \alpha_i + 0.5y_{i,t-1} + 0.5x_{it} + u_{it}$, where $x_{it} = 0.5x_{i,t-1} + 0.5u_{i,t-1} + v_{it}$, $v_{it} \sim i.i.d.N(0,\sigma_v^2)$, $i = 1,\ldots,N$; $t = -48,\ldots,T$, with $y_{i,-49} = x_{i,-49} = 0$. The first 50 observations are discarded. The signal-to-noise ratio is fixed to be 3 through σ_v^2. Under this alternative, $u_{it} = \sigma_\varepsilon\left(\varepsilon_{it} + \psi_1\varepsilon_{i,t-1}\right)$, where $\sigma_\varepsilon^2 = 1/(1+\psi_1^2)$ with $\psi_1 = 0.2$; and $\varepsilon_{it} \sim i.i.d.N(0,1)$.

Table 3: Power (Size-Adjusted Power) of Tests: MA(1)

		Type of Instrument Matrix/Number of Instruments						
		$j=1$	$j=0$					
Tests	(N,T)	2(T-2)	4	6	8	10	12	14
$m_2^{(j)}$	(100,20)	97.85 (97.85)	98.20 (98.15)	98.25 (98.20)	98.30 (98.10)	98.25 (98.20)	98.35 (98.25)	98.35 (98.25)
$m_{(2,3)}^{(j)2}$	(100,20)	99.25 (99.15)	99.45 (99.45)	99.50 (99.45)	99.50 (99.45)	99.45 (99.45)	99.45 (99.45)	99.45 (99.45)
$m_{(2,4)}^{(j)2}$	(100,20)	99.55 (99.60)	99.70 (99.70)	99.70 (99.70)	99.70 (99.70)	99.70 (99.75)	99.75 (99.75)	99.75 (99.75)
$m_{(2,5)}^{(j)2}$	(100,20)	99.60 (99.60)	99.65 (99.65)	99.65 (99.65)	99.70 (99.70)	99.70 (99.75)	99.70 (99.80)	99.70 (99.80)
$m_{(2,6)}^{(j)2}$	(100,20)	99.40 (99.55)	99.55 (99.60)	99.60 (99.65)	99.65 (99.65)	99.65 (99.70)	99.70 (99.75)	99.75 (99.75)
$m_{(2,7)}^{(j)2}$	(100,20)	99.50 (99.60)	99.55 (99.60)	99.55 (99.60)	99.65 (99.65)	99.65 (99.65)	99.65 (99.65)	99.65 (99.65)
$S_{(j)}$	(100,20)	1.60 (4.65)	99.95 (99.95)	99.75 (99.85)	99.45 (99.65)	99.25 (99.45)	98.75 (99.00)	98.40 (98.80)
$\tilde{J}_{(j)}$	(100,20)	4.05 (5.45)	99.95 (99.75)	99.75 (99.50)	99.50 (99.00)	99.25 (98.75)	98.75 (98.30)	98.40 (96.95)
$m_2^{(j)}$	(150,20)	99.85 (99.80)	99.90 (99.80)	99.95 (99.80)	99.95 (99.80)	99.95 (99.80)	99.95 (99.80)	99.95 (99.80)
$m_{(2,3)}^{(j)2}$	(150,20)	99.95 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,4)}^{(j)2}$	(150,20)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,5)}^{(j)2}$	(150,20)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,6)}^{(j)2}$	(150,20)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,7)}^{(j)2}$	(150,20)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$S_{(j)}$	(150,20)	2.25 (4.75)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	99.95 (99.95)	99.95 (99.95)
$\tilde{J}_{(j)}$	(150,20)	3.75 (4.75)	100.00 (100.00)	100.00 (100.00)	100.00 (99.95)	100.00 (99.95)	99.95 (99.95)	99.95 (99.95)
$m_2^{(j)}$	(200,20)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,3)}^{(j)2}$	(200,20)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,4)}^{(j)2}$	(200,20)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,5)}^{(j)2}$	(200,20)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,6)}^{(j)2}$	(200,20)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,7)}^{(j)2}$	(200,20)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$S_{(j)}$	(200,20)	2.70 (5.60)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$\tilde{J}_{(j)}$	(200,20)	4.15 (5.70)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_2^{(j)}$	(100,30)	99.85 (99.85)	99.85 (99.85)	99.85 (99.85)	99.85 (99.85)	99.85 (99.85)	99.85 (99.85)	99.85 (99.85)
$m_{(2,3)}^{(j)2}$	(100,30)	99.90 (99.90)	99.95 (99.95)	99.95 (99.95)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,4)}^{2}$	(100,30)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,5)}^{(j)2}$	(100,30)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,6)}^{(j)2}$	(100,30)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,7)}^{(j)2}$	(100,30)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$S_{(j)}$	(100,30)	0.30 (5.05)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$\tilde{J}_{(j)}$	(100,30)	4.75 (6.50)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_2^{(j)}$	(150,30)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,3)}^{(j)2}$	(150,30)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,4)}^{(j)2}$	(150,30)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,5)}^{(j)2}$	(150,30)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,6)}^{(j)2}$	(150,30)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,7)}^{(j)2}$	(150,30)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$S_{(j)}$	(150,30)	1.55 (5.05)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$\tilde{J}_{(j)}$	(150,30)	4.30 (4.50)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_2^{(j)}$	(200,30)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,3)}^{(j)2}$	(200,30)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,4)}^{(j)2}$	(200,30)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,5)}^{(j)2}$	(200,30)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,6)}^{(j)2}$	(200,30)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,7)}^{(j)2}$	(200,30)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$S_{(j)}$	(200,30)	2.45 (5.25)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$\tilde{J}_{(j)}$	(200,30)	4.25 (5.25)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)

Notes: This table reports the power (size-adjusted power) of $m_s^{(j)}$, $m_{(2,p)}^{(j)2}$, $S_{(j)}$ and $\tilde{J}_{(j)}$ under the MA(1) alternative; for $j = 0, 1, 2$; $p = 3, 4, 5, 6, 7$. For $j = 0$, we choose $\kappa = 2, 3, 4, 5, 6, 7$. The data is generated as $y_{it} = \alpha_i + 0.5y_{i,t-1} + 0.5x_{it} + u_{it}$, where $x_{it} = 0.5x_{i,t-1} + 0.5u_{i,t-1} + v_{it}$, $v_{it} \sim i.i.d.N(0, \sigma_v^2)$, $i = 1, \ldots, N$; $t = -48, \ldots, T$, with $y_{i,-49} = x_{i,-49} = 0$. The first 50 observations are discarded. The signal-to-noise ratio is fixed to be 3 through σ_v^2. Under this alternative, $u_{it} = \sigma_\varepsilon \left(\varepsilon_{it} + \psi_1 \varepsilon_{i,t-1}\right)$, where $\sigma_\varepsilon^2 = 1/(1+\psi_1^2)$ with $\psi_1 = 0.2$; and $\varepsilon_{it} \sim i.i.d.N(0,1)$.

Table 4: Power (Size-Adjusted Power) of Tests: MA(2)

		Type of Instrument Matrix/Number of Instruments							
		$j=2$	$j=1$	$j=0$					
Tests	(N,T)	T(T-2)	2(T-2)	4	6	8	10	12	14
$m_2^{(j)}$	(100,7)	7.90 (11.10)	8.55 (10.50)	8.15 (10.25)	7.85 (10.70)	7.85 (11.15)	7.85 (10.90)		
$m_{(2,3)}^{(j)2}$	(100,7)	12.35 (10.95)	9.35 (8.25)	8.95 (8.40)	10.25 (9.55)	10.30 (9.55)	10.45 (10.25)		
$m_{(2,4)}^{(j)2}$	(100,7)	14.10 (11.95)	9.10 (9.70)	8.50 (9.95)	9.65 (11.15)	9.70 (11.35)	9.95 (11.10)		
$S_{(j)}$	(100,7)	6.20 (12.65)	4.05 (4.85)	8.50 (8.95)	18.00 (19.35)	14.00 (16.00)	12.55 (13.95)		
$\tilde{J}_{(j)}$	(100,7)	9.75 (6.60)	4.05 (4.55)	9.10 (6.50)	18.20 (13.15)	14.05 (10.55)	12.50 (9.05)		
$m_2^{(j)}$	(150,7)	9.90 (12.80)	9.10 (11.00)	9.50 (11.90)	9.85 (12.35)	9.80 (12.05)	9.75 (11.80)		
$m_{(2,3)}^{(j)2}$	(150,7)	19.25 (20.60)	15.35 (17.50)	15.00 (17.70)	16.85 (18.95)	16.80 (19.55)	16.90 (19.45)		
$m_{(2,4)}^{(j)2}$	(150,7)	23.40 (24.15)	17.80 (19.60)	17.40 (20.05)	19.10 (21.15)	19.30 (21.75)	19.50 (21.85)		
$S_{(j)}$	(150,7)	16.45 (21.85)	4.35 (9.45)	11.65 (12.20)	27.45 (28.50)	22.90 (20.80)	20.15 (20.55)		
$\tilde{J}_{(j)}$	(150,7)	17.75 (14.45)	4.25 (4.80)	11.95 (9.05)	27.55 (17.80)	22.70 (14.90)	19.25 (14.15)		
$m_2^{(j)}$	(200,7)	11.45 (14.55)	11.75 (14.00)	11.55 (13.65)	11.50 (14.05)	11.55 (14.40)	11.55 (14.45)		
$m_{(2,3)}^{(j)2}$	(200,7)	24.50 (24.35)	20.75 (20.35)	20.70 (20.75)	22.00 (22.30)	22.10 (22.10)	22.20 (22.55)		
$m_{(2,4)}^{(j)2}$	(200,7)	32.95 (30.90)	26.05 (27.05)	25.85 (26.40)	27.85 (28.25)	28.10 (28.75)	28.05 (28.90)		
$S_{(j)}$	(200,7)	27.65 (28.60)	4.45 (9.50)	13.65 (13.80)	37.55 (37.65)	31.90 (32.40)	27.40 (30.80)		
$\tilde{J}_{(j)}$	(200,7)	27.35 (20.80)	4.20 (4.25)	14.15 (10.60)	37.55 (26.60)	31.40 (24.25)	26.55 (21.10)		
$m_2^{(j)}$	(100,11)	13.40 (16.20)	13.45 (16.65)	13.65 (15.70)	13.70 (15.55)	13.75 (15.65)	13.75 (15.75)	13.70 (15.65)	13.70 (15.65)
$m_{(2,3)}^{(j)2}$	(100,11)	32.65 (35.10)	29.55 (31.20)	29.90 (32.00)	32.60 (34.50)	32.35 (33.85)	32.30 (33.80)	32.25 (33.55)	32.15 (33.60)
$m_{(2,4)}^{(j)2}$	(100,11)	47.45 (46.75)	42.60 (40.85)	43.50 (43.50)	45.80 (45.75)	46.35 (46.05)	45.95 (45.45)	46.05 (45.85)	46.00 (46.00)
$m_{(2,5)}^{(j)2}$	(100,11)	50.65 (50.65)	44.05 (44.50)	44.70 (44.10)	47.65 (47.35)	48.45 (47.75)	48.65 (48.50)	48.50 (48.25)	48.60 (48.70)
$m_{(2,6)}^{(j)2}$	(100,11)	50.35 (50.45)	43.90 (45.45)	44.30 (45.30)	47.15 (48.85)	47.50 (49.95)	47.80 (49.80)	47.80 (49.35)	48.15 (49.20)
$m_{(2,7)}^{(j)2}$	(100,11)	49.50 (46.50)	42.10 (41.15)	42.55 (40.55)	44.80 (43.40)	45.35 (43.75)	45.30 (43.90)	45.65 (43.90)	45.50 (43.65)
$S_{(j)}$	(100,11)	0.00 (7.35)	3.25 (6.00)	14.90 (14.00)	57.00 (59.40)	54.30 (58.25)	46.40 (51.35)	41.60 (44.35)	38.45 (38.60)
$\tilde{J}_{(j)}$	(100,11)	2.90 (6.25)	3.90 (5.20)	15.55 (10.35)	57.35 (46.30)	54.35 (43.85)	46.30 (37.85)	41.40 (31.75)	38.50 (26.40)
$m_2^{(j)}$	(150,11)	16.60 (20.55)	15.80 (19.40)	16.15 (20.20)	16.40 (19.90)	16.30 (19.80)	16.25 (19.90)	15.95 (19.75)	15.95 (19.75)
$m_{(2,3)}^{(j)2}$	(150,11)	46.05 (47.95)	42.45 (44.40)	42.55 (44.00)	44.65 (47.30)	44.50 (47.15)	44.60 (47.35)	44.80 (47.35)	44.90 (47.10)
$m_{(2,4)}^{(j)2}$	(150,11)	67.00 (66.45)	62.35 (62.35)	62.15 (61.50)	65.45 (64.75)	65.85 (65.30)	65.60 (64.65)	65.70 (64.85)	65.65 (65.25)
$m_{(2,5)}^{(j)2}$	(150,11)	72.65 (74.15)	66.00 (68.30)	66.75 (68.45)	69.50 (71.50)	70.00 (72.25)	70.30 (72.35)	70.25 (72.40)	70.10 (72.60)
$m_{(2,6)}^{(j)2}$	(150,11)	73.20 (75.70)	67.25 (69.80)	67.80 (70.05)	70.15 (73.05)	70.45 (72.90)	70.85 (72.50)	70.95 (73.00)	70.90 (73.20)
$m_{(2,7)}^{(j)2}$	(150,11)	73.70 (74.10)	67.75 (69.50)	67.70 (69.35)	69.70 (71.60)	70.10 (72.80)	70.70 (72.45)	70.45 (72.80)	70.60 (72.20)
$S_{(j)}$	(150,11)	1.15 (17.45)	3.75 (4.25)	19.55 (20.15)	77.00 (78.15)	75.30 (77.45)	70.45 (71.20)	65.35 (67.45)	61.90 (62.15)
$\tilde{J}_{(j)}$	(150,11)	13.35 (11.00)	3.95 (4.55)	19.95 (14.30)	77.05 (70.15)	75.10 (65.95)	69.80 (61.95)	64.60 (55.70)	60.65 (50.40)
$m_2^{(j)}$	(200,11)	25.20 (21.55)	23.85 (21.20)	24.90 (20.95)	24.75 (21.35)	25.00 (21.45)	25.15 (21.35)	25.05 (21.45)	25.05 (21.45)
$m_{(2,3)}^{(j)2}$	(200,11)	60.85 (59.85)	57.60 (57.05)	58.40 (57.05)	60.30 (59.00)	59.95 (58.80)	59.80 (58.90)	60.00 (58.95)	60.10 (58.90)
$m_{(2,4)}^{(j)2}$	(200,11)	80.85 (81.25)	77.15 (77.55)	77.45 (78.05)	79.30 (79.70)	79.60 (79.80)	79.05 (79.75)	79.15 (79.85)	79.30 (79.85)
$m_{(2,5)}^{(j)2}$	(200,11)	85.45 (85.40)	81.20 (81.25)	82.00 (82.00)	84.00 (83.65)	84.20 (84.00)	84.15 (84.10)	84.25 (84.15)	84.35 (84.10)
$m_{(2,6)}^{(j)2}$	(200,11)	88.65 (87.75)	84.40 (83.25)	85.75 (84.00)	87.15 (85.35)	87.10 (85.70)	87.30 (85.90)	87.40 (85.95)	87.40 (85.90)
$m_{(2,7)}^{(j)2}$	(200,11)	89.05 (87.50)	84.90 (83.10)	85.95 (83.90)	87.70 (85.20)	87.60 (85.45)	87.60 (85.45)	87.35 (85.50)	87.35 (85.45)
$S_{(j)}$	(200,11)	32.20 (15.90)	5.75 (4.25)	24.45 (26.05)	90.10 (89.45)	88.20 (88.55)	86.05 (86.25)	81.55 (81.60)	79.70 (79.00)
$\tilde{J}_{(j)}$	(200,11)	29.85 (19.55)	4.25 (4.90)	26.75 (17.30)	89.45 (82.80)	88.40 (80.90)	85.75 (78.70)	80.80 (73.35)	77.80 (69.30)

Notes: This table reports the power (size-adjusted power) of $m_s^{(j)}$, $m_{(2,p)}^{(j)2}$, $S_{(j)}$ and $\tilde{J}_{(j)}$ under the MA(2) alternative; for $j = 0,1,2$; $p = 3,4,5,6,7$. For $j = 0$, we choose $\kappa = 2,3,4,5,6,7$. The data is generated as $y_{it} = \alpha_i + 0.5y_{i,t-1} + 0.5x_{it} + u_{it}$, where $x_{it} = 0.5x_{i,t-1} + 0.5u_{i,t-1} + v_{it}$, $v_{it} \sim i.i.d.N(0,\sigma_v^2)$, $i = 1,\ldots,N$; $t = -48,\ldots,T$, with $y_{i,-49} = x_{i,-49} = 0$. The first 50 observations are discarded. The signal-to-noise ratio is fixed to be 3 through σ_v^2. Under this alternative, $u_{it} = \sigma_\varepsilon\left(\varepsilon_{it} + \psi_1\varepsilon_{i,t-1} + \psi_2\varepsilon_{i,t-2}\right)$, where $\sigma_\varepsilon^2 = 1/(1+\psi_1^2+\psi_2^2)$ with $\psi_1 = 20/103; \psi_2 = 13/90$; and $\varepsilon_{it} \sim i.i.d.N(0,1)$.

Table 5: Power (Size-Adjusted Power) of Tests: MA(2)

		Type of Instrument Matrix/Number of Instruments						
		$j=1$	$j=0$					
Tests	(N,T)	2(T-2)	4	6	8	10	12	14
$m_2^{(j)}$	(100,20)	23.55 (26.80)	23.85 (25.85)	23.85 (26.10)	24.15 (26.30)	24.15 (26.50)	24.30 (26.50)	24.30 (26.50)
$m_{(2,3)}^{(j)2}$	(100,20)	66.95 (66.40)	66.85 (66.45)	68.75 (67.40)	68.40 (67.25)	68.55 (67.55)	68.50 (67.45)	68.65 (67.50)
$m_{(2,4)}^{(j)2}$	(100,20)	87.35 (88.00)	87.15 (88.00)	88.45 (89.20)	88.80 (89.40)	88.75 (89.60)	89.05 (89.55)	89.05 (89.60)
$m_{(2,5)}^{(j)2}$	(100,20)	90.50 (91.15)	90.30 (90.70)	91.65 (92.05)	92.10 (92.80)	92.30 (93.05)	92.50 (92.75)	92.55 (93.00)
$m_{(2,6)}^{(j)2}$	(100,20)	92.15 (92.75)	91.95 (92.80)	92.80 (93.25)	93.00 (93.40)	93.10 (93.80)	93.30 (94.55)	93.60 (94.55)
$m_{(2,7)}^{(j)2}$	(100,20)	92.30 (93.70)	92.10 (93.10)	93.00 (93.80)	93.20 (93.95)	93.25 (94.10)	93.35 (94.05)	93.50 (94.20)
$S_{(j)}$	(100,20)	1.35 (5.05)	30.95 (33.10)	96.85 (97.35)	97.40 (97.80)	95.45 (96.15)	94.25 (94.85)	92.90 (94.00)
$\tilde{J}_{(j)}$	(100,20)	4.25 (5.75)	31.80 (23.55)	96.95 (95.10)	97.40 (95.25)	95.45 (93.75)	94.25 (92.15)	92.90 (89.75)
$m_2^{(j)}$	(150,20)	32.75 (36.65)	33.55 (36.40)	33.85 (37.10)	33.75 (38.00)	33.70 (38.05)	33.90 (38.25)	33.90 (38.25)
$m_{(2,3)}^{(j)2}$	(150,20)	85.35 (85.50)	85.80 (86.00)	86.60 (86.90)	86.70 (86.90)	86.70 (86.90)	86.75 (86.90)	86.75 (86.85)
$m_{(2,4)}^{(j)2}$	(150,20)	97.25 (97.55)	97.40 (97.55)	97.80 (97.90)	97.80 (97.95)	97.85 (97.95)	97.90 (97.95)	97.90 (97.95)
$m_{(2,5)}^{(j)2}$	(150,20)	99.05 (99.20)	99.05 (99.05)	99.35 (99.40)	99.40 (99.40)	99.40 (99.45)	99.45 (99.45)	99.45 (99.55)
$m_{(2,6)}^{(j)2}$	(150,20)	99.35 (99.30)	99.35 (99.35)	99.45 (99.35)	99.55 (99.40)	99.60 (99.45)	99.70 (99.50)	99.70 (99.60)
$m_{(2,7)}^{(j)2}$	(150,20)	99.50 (99.50)	99.50 (99.30)	99.55 (99.45)	99.55 (99.55)	99.60 (99.60)	99.65 (99.60)	99.70 (99.65)
$S_{(j)}$	(150,20)	2.45 (4.70)	40.85 (39.40)	99.85 (99.85)	99.90 (99.90)	99.70 (99.85)	99.65 (99.65)	99.50 (99.65)
$\tilde{J}_{(j)}$	(150,20)	4.05 (5.40)	41.90 (28.10)	99.85 (99.75)	99.90 (99.80)	99.70 (99.45)	99.65 (99.50)	99.50 (99.35)
$m_2^{(j)}$	(200,20)	40.05 (38.15)	40.80 (39.05)	41.05 (38.90)	41.10 (38.75)	41.15 (38.80)	41.15 (38.75)	41.15 (38.75)
$m_{(2,3)}^{(j)2}$	(200,20)	93.10 (91.60)	93.55 (91.55)	94.20 (92.35)	94.20 (91.95)	94.25 (92.20)	94.25 (92.30)	94.30 (92.35)
$m_{(2,4)}^{(j)2}$	(200,20)	99.90 (99.85)	99.90 (99.90)	99.90 (99.90)	99.90 (99.90)	99.95 (99.90)	99.95 (99.90)	99.95 (99.90)
$m_{(2,5)}^{(j)2}$	(200,20)	99.95 (99.95)	99.95 (99.95)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,6)}^{(j)2}$	(200,20)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,7)}^{(j)2}$	(200,20)	100.00 (99.85)	100.00 (99.85)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$S_{(j)}$	(200,20)	4.40 (6.95)	55.20 (57.90)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$\tilde{J}_{(j)}$	(200,20)	5.30 (6.75)	55.80 (46.85)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_2^{(j)}$	(100,30)	37.05 (39.00)	37.25 (38.25)	37.25 (38.10)	37.10 (38.10)	37.20 (38.00)	37.25 (38.20)	37.25 (38.20)
$m_{(2,3)}^{(j)2}$	(100,30)	87.55 (88.15)	87.50 (87.55)	88.50 (88.75)	88.50 (88.55)	88.45 (88.55)	88.45 (88.55)	88.55 (88.80)
$m_{(2,4)}^{(j)2}$	(100,30)	98.50 (98.70)	98.60 (98.75)	98.85 (98.90)	98.85 (99.10)	98.85 (99.10)	98.90 (99.05)	98.85 (99.15)
$m_{(2,5)}^{(j)2}$	(100,30)	99.45 (99.65)	99.45 (99.60)	99.60 (99.75)	99.60 (99.75)	99.65 (99.75)	99.60 (99.75)	99.70 (99.75)
$m_{(2,6)}^{(j)2}$	(100,30)	99.60 (99.80)	99.70 (99.80)	99.80 (99.80)	99.80 (99.80)	99.80 (99.80)	99.80 (99.80)	99.85 (99.85)
$m_{(2,7)}^{(j)2}$	(100,30)	99.60 (99.65)	99.60 (99.60)	99.65 (99.65)	99.70 (99.65)	99.75 (99.70)	99.80 (99.70)	99.80 (99.70)
$S_{(j)}$	(100,30)	0.25 (5.10)	45.05 (43.40)	99.95 (99.95)	100.00 (100.00)	99.95 (99.95)	99.90 (99.95)	99.90 (99.90)
$\tilde{J}_{(j)}$	(100,30)	4.00 (5.60)	45.85 (34.80)	99.95 (99.85)	100.00 (99.95)	99.95 (99.95)	99.90 (99.80)	99.90 (99.70)
$m_2^{(j)}$	(150,30)	51.80 (55.15)	52.20 (55.55)	52.40 (56.35)	52.45 (55.65)	52.30 (55.25)	52.25 (55.70)	52.25 (55.70)
$m_{(2,3)}^{(j)2}$	(150,30)	97.30 (97.30)	97.30 (97.30)	97.60 (97.60)	97.60 (97.60)	97.65 (97.60)	97.65 (97.65)	97.65 (97.65)
$m_{(2,4)}^{(j)2}$	(150,30)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,5)}^{(j)2}$	(150,30)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,6)}^{(j)2}$	(150,30)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,7)}^{(j)2}$	(150,30)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$S_{(j)}$	(150,30)	1.60 (5.25)	60.05 (62.75)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$\tilde{J}_{(j)}$	(150,30)	4.25 (4.65)	60.75 (53.80)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_2^{(j)}$	(200,30)	63.45 (64.45)	64.50 (65.10)	65.05 (65.90)	64.75 (66.10)	64.75 (66.15)	64.85 (65.85)	64.85 (65.85)
$m_{(2,3)}^{(j)2}$	(200,30)	99.85 (99.85)	99.85 (99.85)	99.85 (99.85)	99.85 (99.85)	99.85 (99.85)	99.85 (99.85)	99.85 (99.85)
$m_{(2,4)}^{(j)2}$	(200,30)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,5)}^{(j)2}$	(200,30)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,6)}^{(j)2}$	(200,30)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,7)}^{(j)2}$	(200,30)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$S_{(j)}$	(200,30)	2.05 (5.00)	73.75 (73.35)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$\tilde{J}_{(j)}$	(200,30)	4.25 (5.70)	74.30 (64.60)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)

Notes: This table reports the power (size-adjusted power) of $m_s^{(j)}$, $m_{(2,p)}^{(j)2}$, $S_{(j)}$ and $\tilde{J}_{(j)}$ under the MA(2) alternative; for $j = 0, 1, 2$; $p = 3, 4, 5, 6, 7$. For $j = 0$, we choose $\kappa = 2, 3, 4, 5, 6, 7$. The data is generated as $y_{it} = \alpha_i + 0.5y_{i,t-1} + 0.5x_{it} + u_{it}$, where $x_{it} = 0.5x_{i,t-1} + 0.5u_{i,t-1} + v_{it}$, $v_{it} \sim i.i.d.N(0, \sigma_v^2)$, $i = 1, \ldots, N$; $t = -48, \ldots, T$, with $y_{i,-49} = x_{i,-49} = 0$. The first 50 observations are discarded. The signal-to-noise ratio is fixed to be 3 through σ_v^2. Under this alternative, $u_{it} = \sigma_\varepsilon \left(\varepsilon_{it} + \psi_1\varepsilon_{i,t-1} + \psi_2\varepsilon_{i,t-2}\right)$, where $\sigma_\varepsilon^2 = 1/(1 + \psi_1^2 + \psi_2^2)$ with $\psi_1 = 20/103; \psi_2 = 13/90$; and $\varepsilon_{it} \sim i.i.d.N(0, 1)$.

Table 6: Power (Size-Adjusted Power) of Tests: AR(1)

		Type of Instrument Matrix/Number of Instruments							
		$j=2$	$j=1$	$j=0$					
Tests	(N,T)	T(T-2)	2(T-2)	4	6	8	10	12	14
$m_2^{(j)}$	(100,7)	22.65 (16.45)	22.00 (16.25)	23.15 (16.80)	22.70 (16.90)	22.75 (16.90)	22.55 (17.05)		
$m_{(2,3)}^{(j)2}$	(100,7)	25.85 (24.30)	19.45 (18.10)	22.70 (22.05)	23.20 (22.25)	23.45 (22.70)	23.50 (23.15)		
$m_{(2,4)}^{(j)2}$	(100,7)	25.75 (23.20)	18.15 (19.05)	20.60 (22.45)	21.15 (22.85)	21.20 (23.10)	21.15 (22.50)		
$S_{(j)}$	(100,7)	7.20 (14.10)	4.20 (5.10)	26.70 (27.15)	23.70 (25.50)	19.70 (21.15)	16.35 (18.70)		
$\tilde{J}_{(j)}$	(100,7)	9.80 (7.30)	4.20 (4.55)	27.70 (19.90)	24.05 (18.00)	19.75 (15.40)	16.35 (12.15)		
$m_2^{(j)}$	(150,7)	31.00 (29.95)	29.75 (29.95)	31.65 (31.75)	31.60 (31.60)	31.45 (31.45)	31.35 (31.15)		
$m_{(2,3)}^{(j)2}$	(150,7)	36.65 (38.45)	31.00 (32.70)	34.05 (36.60)	34.55 (37.05)	34.45 (37.60)	34.70 (37.40)		
$m_{(2,4)}^{(j)2}$	(150,7)	37.00 (37.85)	29.45 (31.65)	32.70 (35.95)	33.30 (36.40)	33.50 (36.55)	33.80 (36.65)		
$S_{(j)}$	(150,7)	19.80 (26.25)	4.40 (9.60)	41.25 (42.80)	36.70 (37.90)	30.00 (27.35)	25.90 (26.70)		
$\tilde{J}_{(j)}$	(150,7)	21.55 (17.00)	4.20 (4.80)	42.20 (31.70)	36.80 (26.00)	29.80 (20.95)	25.05 (19.05)		
$m_2^{(j)}$	(200,7)	42.15 (34.80)	40.90 (34.70)	42.80 (35.95)	42.75 (35.30)	42.80 (35.50)	42.85 (35.55)		
$m_{(2,3)}^{(j)2}$	(200,7)	52.00 (51.90)	45.35 (44.35)	50.25 (50.45)	50.85 (51.30)	50.80 (50.80)	50.75 (51.35)		
$m_{(2,4)}^{(j)2}$	(200,7)	55.00 (51.75)	46.35 (47.70)	50.15 (50.90)	50.85 (51.95)	50.80 (52.25)	50.70 (51.90)		
$S_{(j)}$	(200,7)	35.10 (36.65)	4.25 (9.30)	55.70 (55.85)	50.90 (51.10)	45.10 (46.10)	39.00 (42.15)		
$\tilde{J}_{(j)}$	(200,7)	34.35 (26.85)	4.00 (4.80)	56.30 (44.55)	50.90 (39.55)	44.75 (36.85)	37.60 (31.10)		
$m_2^{(j)}$	(100,11)	46.05 (40.50)	44.65 (41.30)	46.95 (42.80)	46.55 (42.55)	46.65 (42.35)	46.65 (42.75)	46.70 (41.30)	46.70 (41.30)
$m_{(2,3)}^{(j)2}$	(100,11)	56.20 (58.95)	51.35 (52.75)	55.75 (57.30)	56.50 (57.90)	56.25 (57.95)	56.30 (58.10)	56.35 (58.05)	56.70 (57.85)
$m_{(2,4)}^{(j)2}$	(100,11)	61.10 (60.35)	55.05 (53.65)	58.70 (58.70)	59.80 (59.80)	60.05 (59.85)	60.15 (59.70)	59.95 (59.85)	60.15 (60.20)
$m_{(2,5)}^{(j)2}$	(100,11)	63.25 (63.05)	55.35 (55.80)	58.60 (58.30)	59.90 (59.70)	60.55 (60.40)	60.70 (60.65)	60.90 (60.85)	60.90 (60.95)
$m_{(2,6)}^{(j)2}$	(100,11)	61.20 (61.25)	54.85 (55.75)	57.50 (58.45)	58.40 (59.90)	59.00 (60.65)	59.05 (60.60)	59.05 (60.30)	59.05 (59.95)
$m_{(2,7)}^{(j)2}$	(100,11)	59.55 (56.55)	52.85 (51.70)	55.70 (54.20)	56.80 (55.25)	56.90 (55.20)	57.15 (55.45)	57.15 (55.45)	57.40 (55.50)
$S_{(j)}$	(100,11)	0.00 (5.85)	3.05 (5.15)	65.00 (63.95)	61.20 (62.90)	55.85 (59.75)	50.00 (54.75)	44.15 (46.85)	39.90 (40.05)
$\tilde{J}_{(j)}$	(100,11)	1.80 (3.60)	3.90 (5.45)	65.60 (53.20)	61.55 (51.30)	55.95 (47.45)	50.00 (42.25)	44.10 (34.55)	39.95 (29.45)
$m_2^{(j)}$	(150,11)	61.75 (56.80)	60.15 (56.70)	62.35 (57.95)	62.10 (57.95)	61.80 (57.95)	62.10 (58.00)	62.25 (58.10)	62.25 (58.10)
$m_{(2,3)}^{(j)2}$	(150,11)	76.85 (77.90)	71.95 (73.75)	75.35 (76.65)	76.55 (77.85)	76.45 (78.05)	76.45 (78.10)	76.40 (78.05)	76.55 (77.90)
$m_{(2,4)}^{(j)2}$	(150,11)	81.50 (80.90)	76.80 (76.90)	79.75 (78.75)	80.75 (80.20)	81.10 (80.75)	80.95 (80.35)	80.95 (80.55)	81.05 (80.70)
$m_{(2,5)}^{(j)2}$	(150,11)	84.05 (84.95)	78.75 (80.75)	81.30 (82.60)	82.35 (83.65)	82.60 (84.05)	82.40 (84.15)	82.40 (83.95)	82.50 (84.10)
$m_{(2,6)}^{(j)2}$	(150,11)	84.05 (85.10)	79.30 (81.05)	81.65 (82.95)	82.40 (84.15)	82.30 (83.65)	82.40 (83.25)	82.40 (83.45)	82.50 (83.55)
$m_{(2,7)}^{(j)2}$	(150,11)	83.00 (83.30)	78.95 (79.95)	80.75 (81.35)	81.10 (82.15)	81.45 (82.65)	81.50 (82.35)	81.50 (82.60)	81.65 (82.40)
$S_{(j)}$	(150,11)	0.80 (18.45)	3.65 (4.55)	82.70 (83.50)	82.90 (83.80)	78.95 (80.90)	74.55 (75.15)	70.00 (71.65)	65.35 (65.50)
$\tilde{J}_{(j)}$	(150,11)	14.20 (11.35)	3.75 (4.15)	83.30 (74.35)	83.00 (76.45)	78.55 (70.20)	74.15 (67.05)	69.10 (60.60)	64.30 (54.80)
$m_2^{(j)}$	(200,11)	74.30 (71.75)	72.95 (70.95)	74.40 (72.00)	74.30 (72.20)	74.40 (72.20)	74.45 (72.30)	74.35 (72.20)	74.35 (72.20)
$m_{(2,3)}^{(j)2}$	(200,11)	88.25 (88.80)	85.15 (85.30)	87.25 (87.95)	87.70 (88.40)	87.65 (88.30)	87.80 (88.35)	87.85 (88.30)	87.85 (88.35)
$m_{(2,4)}^{(j)2}$	(200,11)	92.55 (92.35)	89.85 (89.60)	91.50 (91.05)	91.95 (91.65)	92.15 (91.85)	92.15 (91.70)	92.15 (91.85)	92.15 (91.90)
$m_{(2,5)}^{(j)2}$	(200,11)	93.20 (93.20)	90.70 (90.60)	92.05 (92.05)	92.20 (92.35)	92.45 (92.60)	92.45 (92.60)	92.35 (92.45)	92.45 (92.50)
$m_{(2,6)}^{(j)2}$	(200,11)	94.45 (94.90)	91.60 (92.15)	92.65 (93.70)	93.00 (94.10)	93.20 (94.10)	93.20 (94.20)	93.25 (94.35)	93.25 (94.40)
$m_{(2,7)}^{(j)2}$	(200,11)	93.60 (94.55)	91.15 (92.35)	91.90 (93.15)	92.20 (93.65)	92.30 (93.70)	92.40 (93.60)	92.35 (93.75)	92.40 (93.55)
$S_{(j)}$	(200,11)	16.05 (33.25)	4.15 (5.50)	92.25 (91.50)	93.65 (93.90)	91.65 (91.40)	88.45 (88.30)	85.25 (85.20)	82.65 (83.30)
$\tilde{J}_{(j)}$	(200,11)	31.20 (20.20)	4.15 (5.20)	92.50 (86.45)	93.65 (88.95)	91.40 (85.20)	88.05 (81.70)	84.50 (78.20)	81.60 (74.55)

Notes: This table reports the power (size-adjusted power) of $m_s^{(j)}$, $m_{(2,p)}^{(j)2}$, $S_{(j)}$ and $\tilde{J}_{(j)}$ under the AR(1) alternative; for $j = 0, 1, 2;$ $p = 3, 4, 5, 6, 7$. For $j = 0$, we choose $\kappa = 2, 3, 4, 5, 6, 7$. The data is generated as $y_{it} = \alpha_i + 0.5y_{i,t-1} + 0.5x_{it} + u_{it}$, where $x_{it} = 0.5x_{i,t-1} + 0.5u_{i,t-1} + v_{it}$, $v_{it} \sim i.i.d.N(0, \sigma_v^2)$, $i = 1, \ldots, N$; $t = -48, \ldots, T$, with $y_{i,-49} = x_{i,-49} = 0$. The first 50 observations are discarded. The signal-to-noise ratio is fixed to be 3 through σ_v^2. Under this alternative, $u_{it} = \rho_1 u_{i,t-1} + \sigma_\varepsilon \varepsilon_{it}$, where $\sigma_\varepsilon^2 = 1/(1-\rho_1^2)$ with $\rho_1 = 0.2$; and $\varepsilon_{it} \sim i.i.d.N(0, 1)$.

Table 7: Power (Size-Adjusted Power) of Tests: AR(1)

		Type of Instrument Matrix/Number of Instruments						
		$j=1$	$j=0$					
Tests	(N,T)	2(T-2)	4	6	8	10	12	14
$m_2^{(j)}$	(100,20)	76.10 (75.95)	78.15 (77.50)	78.05 (77.30)	78.10 (77.50)	78.20 (77.65)	78.20 (77.95)	78.20 (77.95)
$m_{(2,3)}^{(j)2}$	(100,20)	89.55 (89.35)	91.10 (90.80)	91.60 (90.90)	91.55 (90.95)	91.65 (91.15)	91.65 (91.00)	91.60 (91.05)
$m_{(2,4)}^{(j)2}$	(100,20)	93.70 (94.30)	94.65 (95.00)	95.10 (95.25)	95.20 (95.35)	95.20 (95.50)	95.10 (95.55)	95.10 (95.55)
$m_{(2,5)}^{(j)2}$	(100,20)	94.40 (94.85)	95.25 (95.45)	95.70 (95.85)	95.85 (96.05)	96.05 (96.20)	96.20 (96.35)	96.20 (96.40)
$m_{(2,6)}^{(j)2}$	(100,20)	95.65 (95.75)	95.80 (96.00)	95.80 (96.15)	96.00 (96.25)	96.15 (96.35)	96.20 (96.60)	96.40 (96.65)
$m_{(2,7)}^{(j)2}$	(100,20)	96.00 (96.55)	96.40 (96.55)	96.50 (96.70)	96.60 (96.70)	96.75 (96.75)	96.80 (96.80)	96.75 (96.85)
$S_{(j)}$	(100,20)	1.50 (5.20)	97.05 (97.40)	98.05 (98.70)	97.35 (97.85)	96.15 (97.05)	93.75 (94.55)	92.25 (93.15)
$\tilde{J}_{(j)}$	(100,20)	4.45 (6.05)	97.15 (95.45)	98.15 (96.40)	97.35 (95.50)	96.15 (93.65)	93.75 (90.85)	92.25 (88.00)
$m_2^{(j)}$	(150,20)	90.00 (88.60)	90.95 (89.20)	90.65 (88.80)	90.70 (88.80)	90.70 (89.05)	90.90 (88.80)	90.90 (88.80)
$m_{(2,3)}^{(j)2}$	(150,20)	97.55 (97.60)	97.75 (97.75)	97.95 (97.95)	97.95 (98.05)	98.00 (98.05)	98.05 (98.10)	98.10 (98.10)
$m_{(2,4)}^{(j)2}$	(150,20)	99.20 (99.35)	99.45 (99.45)	99.55 (99.60)	99.55 (99.60)	99.55 (99.60)	99.60 (99.60)	99.60 (99.60)
$m_{(2,5)}^{(j)2}$	(150,20)	99.55 (99.65)	99.70 (99.70)	99.70 (99.70)	99.75 (99.75)	99.75 (99.75)	99.75 (99.75)	99.75 (99.75)
$m_{(2,6)}^{(j)2}$	(150,20)	99.55 (99.45)	99.70 (99.65)	99.75 (99.75)	99.75 (99.75)	99.75 (99.75)	99.75 (99.75)	99.75 (99.75)
$m_{(2,7)}^{(j)2}$	(150,20)	99.45 (99.45)	99.50 (99.45)	99.55 (99.55)	99.60 (99.55)	99.70 (99.55)	99.65 (99.60)	99.80 (99.60)
$S_{(j)}$	(150,20)	2.90 (5.45)	99.35 (99.30)	99.85 (99.85)	99.95 (99.95)	99.85 (99.85)	99.70 (99.80)	99.40 (99.45)
$\tilde{J}_{(j)}$	(150,20)	4.25 (5.10)	99.35 (98.90)	99.85 (99.70)	99.95 (99.65)	99.85 (99.65)	99.55 (99.45)	99.35 (98.90)
$m_2^{(j)}$	(200,20)	96.30 (96.10)	96.65 (96.55)	96.80 (96.50)	96.85 (96.60)	96.80 (96.75)	96.80 (96.70)	96.80 (96.70)
$m_{(2,3)}^{(j)2}$	(200,20)	99.65 (99.60)	99.70 (99.60)	99.70 (99.60)	99.70 (99.65)	99.70 (99.65)	99.80 (99.65)	99.80 (99.65)
$m_{(2,4)}^{(j)2}$	(200,20)	99.90 (99.90)	100.00 (99.90)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,5)}^{(j)2}$	(200,20)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,6)}^{(j)2}$	(200,20)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,7)}^{(j)2}$	(200,20)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$S_{(j)}$	(200,20)	3.50 (6.00)	99.95 (99.95)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	99.90 (99.90)
$\tilde{J}_{(j)}$	(200,20)	4.60 (6.85)	99.95 (99.95)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (99.90)	99.90 (99.90)
$m_2^{(j)}$	(100,30)	91.85 (92.10)	92.65 (92.75)	92.60 (92.75)	92.70 (92.80)	92.65 (92.75)	92.70 (92.70)	92.70 (92.70)
$m_{(2,3)}^{(j)2}$	(100,30)	98.80 (98.90)	99.00 (99.00)	99.05 (99.05)	99.00 (99.00)	98.95 (98.95)	98.90 (98.95)	98.90 (98.90)
$m_{(2,4)}^{(j)2}$	(100,30)	99.65 (99.75)	99.75 (99.80)	99.80 (99.80)	99.80 (99.80)	99.80 (99.80)	99.80 (99.80)	99.80 (99.80)
$m_{(2,5)}^{(j)2}$	(100,30)	99.85 (99.90)	99.90 (99.90)	99.90 (99.90)	99.90 (99.90)	99.90 (99.90)	99.90 (99.90)	99.90 (99.90)
$m_{(2,6)}^{(j)2}$	(100,30)	99.95 (99.95)	99.95 (99.95)	99.95 (99.95)	99.95 (99.95)	99.95 (99.95)	99.95 (99.95)	99.95 (99.95)
$m_{(2,7)}^{(j)2}$	(100,30)	99.95 (99.95)	99.95 (99.95)	99.95 (99.95)	99.95 (99.95)	99.95 (99.95)	100.00 (99.95)	100.00 (99.95)
$S_{(j)}$	(100,30)	0.45 (5.00)	99.85 (99.85)	99.90 (99.90)	100.00 (100.00)	99.90 (99.95)	99.90 (99.90)	99.90 (99.90)
$\tilde{J}_{(j)}$	(100,30)	3.70 (5.15)	99.85 (99.80)	99.90 (99.80)	100.00 (100.00)	99.90 (99.85)	99.90 (99.75)	99.90 (99.60)
$m_2^{(j)}$	(150,30)	97.95 (97.90)	98.30 (98.00)	98.25 (98.15)	98.25 (98.25)	98.25 (98.25)	98.25 (98.35)	98.25 (98.35)
$m_{(2,3)}^{(j)2}$	(150,30)	99.90 (99.95)	99.95 (99.95)	99.95 (99.95)	99.95 (99.95)	99.95 (99.95)	99.95 (99.95)	99.95 (99.95)
$m_{(2,4)}^{(j)2}$	(150,30)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,5)}^{(j)2}$	(150,30)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,6)}^{(j)2}$	(150,30)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,7)}^{(j)2}$	(150,30)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$S_{(j)}$	(150,30)	2.00 (5.25)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$\tilde{J}_{(j)}$	(150,30)	4.95 (5.40)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_2^{(j)}$	(200,30)	99.80 (99.75)	99.80 (99.80)	99.85 (99.80)	99.90 (99.80)	99.90 (99.80)	99.90 (99.80)	99.90 (99.80)
$m_{(2,3)}^{(j)2}$	(200,30)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,4)}^{(j)2}$	(200,30)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,5)}^{(j)2}$	(200,30)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,6)}^{(j)2}$	(200,30)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,7)}^{(j)2}$	(200,30)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$S_{(j)}$	(200,30)	3.05 (6.80)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$\tilde{J}_{(j)}$	(200,30)	5.55 (6.35)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)

Notes: This table reports the power (size-adjusted power) of $m_s^{(j)}$, $m_{(2,p)}^{(j)2}$, $S_{(j)}$ and $\tilde{J}_{(j)}$ under the AR(1) alternative; for $j = 0, 1, 2$; $p = 3, 4, 5, 6, 7$. For $j = 0$, we choose $\kappa = 2, 3, 4, 5, 6, 7$. The data is generated as $y_{it} = \alpha_i + 0.5 y_{i,t-1} + 0.5 x_{it} + u_{it}$, where $x_{it} = 0.5 x_{i,t-1} + 0.5 u_{i,t-1} + v_{it}$, $v_{it} \sim i.i.d.N(0, \sigma_v^2)$, $i = 1, \ldots, N$; $t = -48, \ldots, T$, with $y_{i,-49} = x_{i,-49} = 0$. The first 50 observations are discarded. The signal-to-noise ratio is fixed to be 3 through σ_v^2. Under this alternative, $u_{it} = \rho_1 u_{i,t-1} + \sigma_\varepsilon \varepsilon_{it}$, where $\sigma_\varepsilon^2 = 1/(1-\rho_1^2)$ with $\rho_1 = 0.2$; and $\varepsilon_{it} \sim i.i.d.N(0,1)$.

Table 8: Power (Size-Adjusted Power) of Tests: AR(2)

		Type of Instrument Matrix/Number of Instruments							
		$j=2$	$j=1$	$j=0$					
Tests	(N,T)	T(T-2)	2(T-2)	4	6	8	10	12	14
$m_2^{(j)}$	(100,7)	5.00 (6.20)	4.85 (5.70)	4.70 (5.45)	4.80 (5.60)	4.80 (5.70)	4.85 (5.70)		
$m_{(2,3)}^{(j)2}$	(100,7)	7.10 (6.10)	5.65 (4.95)	5.55 (5.15)	6.10 (5.70)	6.00 (5.95)	6.05 (6.00)		
$m_{(2,4)}^{(j)2}$	(100,7)	7.65 (6.50)	5.20 (5.55)	5.15 (5.95)	5.50 (6.20)	5.70 (6.40)	5.55 (6.05)		
$S_{(j)}$	(100,7)	4.25 (8.70)	3.90 (4.40)	10.15 (10.40)	13.05 (14.35)	12.80 (13.80)	10.45 (12.00)		
$\tilde{J}_{(j)}$	(100,7)	6.55 (5.30)	3.90 (4.40)	10.75 (6.75)	13.35 (10.45)	12.85 (9.55)	10.45 (8.65)		
$m_2^{(j)}$	(150,7)	4.55 (5.90)	4.35 (5.35)	4.75 (5.75)	4.65 (5.95)	4.60 (5.75)	4.65 (5.70)		
$m_{(2,3)}^{(j)2}$	(150,7)	9.30 (9.95)	7.55 (8.60)	7.65 (9.15)	8.40 (9.80)	8.45 (9.90)	8.45 (9.70)		
$m_{(2,4)}^{(j)2}$	(150,7)	10.55 (11.05)	6.95 (7.70)	7.05 (8.65)	7.70 (9.25)	8.05 (9.40)	8.00 (9.70)		
$S_{(j)}$	(150,7)	9.55 (12.65)	4.95 (9.35)	12.15 (12.70)	18.45 (19.40)	15.90 (13.95)	14.30 (14.60)		
$\tilde{J}_{(j)}$	(150,7)	10.50 (9.05)	4.75 (4.45)	12.45 (9.55)	18.60 (12.00)	15.55 (10.75)	13.75 (9.85)		
$m_2^{(j)}$	(200,7)	5.35 (5.30)	5.35 (5.35)	5.35 (5.00)	5.35 (5.10)	5.25 (5.20)	5.30 (5.35)		
$m_{(2,3)}^{(j)2}$	(200,7)	9.70 (9.70)	8.00 (7.50)	8.15 (8.20)	8.55 (8.60)	8.40 (8.40)	8.50 (8.65)		
$m_{(2,4)}^{(j)2}$	(200,7)	13.35 (12.25)	9.50 (10.25)	10.00 (10.25)	10.40 (10.95)	10.55 (11.15)	10.80 (11.40)		
$S_{(j)}$	(200,7)	15.90 (16.80)	4.10 (8.60)	15.25 (15.35)	24.10 (24.35)	22.15 (22.90)	19.15 (21.80)		
$\tilde{J}_{(j)}$	(200,7)	15.35 (10.65)	3.85 (4.90)	15.35 (13.35)	24.10 (16.95)	21.60 (17.65)	18.15 (14.95)		
$m_2^{(j)}$	(100,11)	4.95 (5.55)	4.85 (5.95)	4.95 (5.60)	5.10 (5.40)	5.15 (5.50)	5.10 (5.55)	5.15 (5.30)	5.15 (5.30)
$m_{(2,3)}^{(j)2}$	(100,11)	10.45 (12.35)	9.20 (9.85)	9.65 (11.05)	10.55 (12.00)	10.10 (11.30)	10.30 (11.50)	10.25 (11.40)	10.15 (11.10)
$m_{(2,4)}^{(j)2}$	(100,11)	18.25 (17.70)	14.80 (14.05)	15.80 (15.80)	17.10 (17.05)	17.20 (17.05)	17.30 (16.95)	17.55 (17.30)	17.55 (17.55)
$m_{(2,5)}^{(j)2}$	(100,11)	21.70 (21.55)	17.45 (17.70)	18.10 (17.95)	19.70 (19.45)	20.45 (19.95)	20.50 (20.45)	20.45 (20.35)	20.60 (20.70)
$m_{(2,6)}^{(j)2}$	(100,11)	23.65 (23.65)	18.75 (19.70)	19.30 (19.80)	20.60 (22.30)	21.45 (23.20)	21.60 (23.20)	21.70 (23.00)	21.90 (22.50)
$m_{(2,7)}^{(j)2}$	(100,11)	23.05 (20.30)	18.25 (17.45)	18.20 (17.15)	19.65 (18.30)	19.85 (18.95)	20.30 (18.90)	20.20 (19.10)	20.10 (19.10)
$S_{(j)}$	(100,11)	0.00 (6.40)	3.35 (5.30)	17.75 (16.95)	35.10 (37.65)	34.05 (38.25)	31.40 (34.70)	26.75 (29.35)	25.05 (25.25)
$\tilde{J}_{(j)}$	(100,11)	2.25 (4.85)	4.20 (5.85)	18.20 (12.15)	35.80 (25.05)	34.20 (25.65)	31.40 (23.90)	26.75 (19.65)	25.10 (15.70)
$m_2^{(j)}$	(150,11)	4.90 (5.95)	4.90 (5.95)	5.00 (5.85)	4.95 (5.85)	4.95 (5.85)	4.90 (5.85)	4.85 (5.80)	4.85 (5.80)
$m_{(2,3)}^{(j)2}$	(150,11)	16.05 (17.20)	14.80 (15.35)	15.30 (16.20)	16.25 (17.25)	15.85 (17.05)	16.00 (17.30)	16.00 (17.35)	15.95 (17.30)
$m_{(2,4)}^{(j)2}$	(150,11)	28.80 (28.30)	24.55 (24.55)	26.10 (25.30)	27.25 (26.60)	27.70 (27.30)	27.45 (26.70)	27.55 (26.75)	27.45 (26.75)
$m_{(2,5)}^{(j)2}$	(150,11)	35.55 (38.20)	29.25 (31.85)	30.80 (32.30)	32.85 (35.55)	33.55 (36.00)	33.45 (36.30)	33.60 (35.75)	33.65 (36.25)
$m_{(2,6)}^{(j)2}$	(150,11)	38.10 (40.55)	30.95 (33.05)	32.40 (34.60)	33.70 (37.70)	34.35 (37.35)	34.50 (36.85)	34.65 (36.90)	34.75 (37.15)
$m_{(2,7)}^{(j)2}$	(150,11)	39.20 (39.70)	31.90 (33.60)	32.60 (34.10)	34.65 (36.60)	34.80 (37.90)	35.20 (37.60)	35.65 (38.00)	35.75 (37.10)
$S_{(j)}$	(150,11)	0.70 (12.45)	4.05 (4.85)	23.65 (24.80)	53.15 (54.90)	54.80 (58.40)	49.95 (50.40)	44.75 (48.00)	42.05 (42.25)
$\tilde{J}_{(j)}$	(150,11)	9.40 (8.45)	4.25 (5.5)	24.65 (16.50)	53.25 (44.90)	54.35 (42.25)	49.25 (41.50)	43.60 (35.55)	41.35 (31.40)
$m_2^{(j)}$	(200,11)	6.90 (5.90)	6.60 (5.70)	6.65 (5.75)	6.60 (5.95)	6.60 (5.85)	6.55 (5.85)	6.45 (5.85)	6.45 (5.85)
$m_{(2,3)}^{(j)2}$	(200,11)	21.10 (20.10)	18.95 (18.55)	20.20 (19.35)	21.15 (20.20)	20.65 (20.15)	20.85 (20.15)	21.10 (20.15)	21.15 (20.20)
$m_{(2,4)}^{(j)2}$	(200,11)	37.55 (37.90)	31.40 (31.95)	33.15 (33.95)	35.05 (35.85)	35.60 (36.35)	34.75 (36.10)	34.95 (35.95)	35.05 (36.00)
$m_{(2,5)}^{(j)2}$	(200,11)	47.00 (46.90)	38.80 (39.15)	40.75 (40.75)	44.15 (43.50)	44.70 (44.10)	44.75 (44.25)	44.40 (44.15)	44.20 (44.10)
$m_{(2,6)}^{(j)2}$	(200,11)	53.20 (51.25)	46.35 (44.15)	48.80 (45.25)	51.55 (47.80)	51.30 (48.70)	51.75 (48.85)	52.40 (48.90)	52.15 (48.70)
$m_{(2,7)}^{(j)2}$	(200,11)	56.20 (53.35)	49.50 (45.40)	51.65 (46.40)	54.30 (48.00)	54.50 (48.75)	54.20 (49.25)	54.25 (49.55)	53.90 (49.50)
$S_{(j)}$	(200,11)	19.10 (8.85)	5.00 (3.80)	29.05 (31.25)	68.65 (67.45)	70.65 (71.50)	66.40 (66.95)	63.00 (63.05)	59.75 (58.70)
$\tilde{J}_{(j)}$	(200,11)	17.85 (11.40)	4.25 (5.05)	32.10 (19.60)	67.45 (57.70)	71.05 (58.30)	66.15 (56.00)	61.40 (49.85)	57.25 (46.85)

Notes: This table reports the power (size-adjusted power) of $m_s^{(j)}$, $m_{(2,p)}^{(j)2}$, $S_{(j)}$ and $\tilde{J}_{(j)}$ under the AR(2) alternative; for $j = 0, 1, 2$; $p = 3, 4, 5, 6, 7$. For $j = 0$, we choose $\kappa = 2, 3, 4, 5, 6, 7$. The data is generated as $y_{it} = \alpha_i + 0.5y_{i,t-1} + 0.5x_{it} + u_{it}$, where $x_{it} = 0.5x_{i,t-1} + 0.5u_{i,t-1} + v_{it}$, $v_{it} \sim i.i.d.N(0, \sigma_v^2)$, $i = 1, \ldots, N$; $t = -48, \ldots, T$, with $y_{i,-49} = x_{i,-49} = 0$. The first 50 observations are discarded. The signal-to-noise ratio is fixed to be 3 through σ_v^2. Under this alternative, $u_{it} = \rho_1 u_{i,t-1} + \rho_2 u_{i,t-2} + \sigma_\varepsilon \varepsilon_{it}$, where $\sigma_\varepsilon^2 = (1+\rho_2)\left[(1-\rho_2)^2 - \rho_1^2\right]/(1-\rho_2)$ with $\rho_1 = 0.2; \rho_2 = 0.1$; and $\varepsilon_{it} \sim i.i.d.N(0,1)$.

Table 9: Power (Size-Adjusted Power) of Tests: AR(2)

		Type of Instrument Matrix/Number of Instruments						
		$j=1$	$j=0$					
Tests	(N,T)	2(T-2)	4	6	8	10	12	14
$m_2^{(j)}$	(100,20)	6.05 (7.25)	6.15 (7.00)	5.95 (6.80)	5.90 (6.75)	5.95 (6.95)	6.00 (6.80)	6.00 (6.80)
$m_{(2,3)}^{(j)2}$	(100,20)	21.95 (21.15)	22.35 (21.95)	23.45 (22.30)	23.05 (22.20)	23.20 (22.35)	23.35 (22.25)	23.40 (22.45)
$m_{(2,4)}^{(j)2}$	(100,20)	40.50 (41.70)	40.80 (42.20)	43.60 (44.80)	44.55 (45.80)	43.90 (46.00)	44.20 (46.25)	44.30 (46.55)
$m_{(2,5)}^{(j)2}$	(100,20)	52.65 (53.90)	53.20 (53.65)	55.10 (56.10)	56.25 (57.30)	56.65 (57.85)	56.90 (57.75)	56.95 (58.45)
$m_{(2,6)}^{(j)2}$	(100,20)	59.40 (61.25)	60.05 (61.55)	62.30 (63.85)	63.40 (64.45)	64.05 (65.55)	64.35 (66.35)	64.45 (66.45)
$m_{(2,7)}^{(j)2}$	(100,20)	62.20 (67.00)	62.85 (65.75)	64.70 (66.85)	65.85 (68.35)	66.40 (68.95)	66.90 (68.60)	67.35 (69.00)
$S_{(j)}$	(100,20)	1.85 (5.00)	34.35 (36.50)	81.90 (84.85)	85.80 (88.35)	84.45 (86.45)	80.65 (82.70)	76.55 (80.10)
$\tilde{J}_{(j)}$	(100,20)	3.75 (5.45)	35.40 (26.95)	82.25 (75.00)	85.80 (78.90)	84.40 (77.95)	80.65 (75.05)	76.55 (69.40)
$m_2^{(j)}$	(150,20)	6.10 (6.65)	6.05 (6.20)	6.00 (6.20)	5.85 (6.45)	5.90 (6.60)	5.90 (6.60)	5.90 (6.60)
$m_{(2,3)}^{(j)2}$	(150,20)	31.70 (31.75)	32.25 (32.50)	33.95 (34.45)	33.90 (34.10)	33.85 (34.30)	33.95 (34.35)	33.95 (34.15)
$m_{(2,4)}^{(j)2}$	(150,20)	59.30 (60.20)	60.95 (62.00)	63.25 (65.50)	64.45 (66.00)	64.30 (65.65)	64.45 (65.40)	64.65 (65.35)
$m_{(2,5)}^{(j)2}$	(150,20)	74.30 (76.05)	75.20 (76.05)	77.10 (78.20)	78.15 (79.00)	78.45 (79.45)	78.65 (79.50)	78.90 (79.60)
$m_{(2,6)}^{(j)2}$	(150,20)	82.00 (81.30)	82.75 (81.00)	83.85 (82.35)	84.35 (83.25)	84.95 (83.40)	85.10 (83.40)	85.20 (83.60)
$m_{(2,7)}^{(j)2}$	(150,20)	86.35 (85.60)	86.10 (84.90)	87.00 (86.05)	87.75 (87.10)	88.05 (87.40)	88.25 (87.85)	88.40 (87.85)
$S_{(j)}$	(150,20)	2.70 (5.05)	50.35 (49.55)	94.95 (95.15)	97.80 (97.95)	97.05 (97.75)	96.70 (97.30)	95.55 (95.80)
$\tilde{J}_{(j)}$	(150,20)	4.15 (4.75)	51.15 (38.10)	94.95 (92.00)	97.70 (95.15)	97.05 (96.05)	96.35 (94.80)	95.35 (93.30)
$m_2^{(j)}$	(200,20)	5.65 (5.25)	6.25 (5.40)	6.40 (5.70)	6.35 (5.90)	6.50 (5.90)	6.55 (5.90)	6.55 (5.90)
$m_{(2,3)}^{(j)2}$	(200,20)	40.65 (36.25)	41.90 (37.20)	43.60 (38.30)	42.85 (38.30)	42.95 (38.65)	43.30 (38.65)	43.30 (38.70)
$m_{(2,4)}^{(j)2}$	(200,20)	74.05 (70.95)	75.35 (73.45)	76.80 (75.55)	77.15 (75.95)	77.25 (75.65)	77.30 (75.80)	77.40 (76.00)
$m_{(2,5)}^{(j)2}$	(200,20)	87.15 (86.25)	87.45 (85.80)	89.15 (87.35)	89.65 (87.90)	89.95 (88.70)	90.10 (88.80)	90.10 (88.95)
$m_{(2,6)}^{(j)2}$	(200,20)	92.95 (92.95)	93.70 (92.60)	94.35 (93.70)	94.95 (94.05)	95.10 (94.60)	95.15 (94.80)	95.15 (95.00)
$m_{(2,7)}^{(j)2}$	(200,20)	95.55 (94.65)	95.90 (94.25)	95.95 (95.00)	96.30 (95.25)	96.55 (95.60)	96.95 (95.60)	96.90 (96.20)
$S_{(j)}$	(200,20)	2.60 (5.50)	60.60 (63.60)	97.85 (98.25)	99.30 (99.40)	99.65 (99.65)	99.60 (99.60)	99.35 (99.30)
$\tilde{J}_{(j)}$	(200,20)	3.55 (4.95)	61.05 (54.15)	97.85 (97.35)	99.30 (98.95)	99.65 (99.15)	99.50 (99.10)	99.35 (98.65)
$m_2^{(j)}$	(100,30)	5.70 (6.15)	5.45 (5.90)	5.55 (5.85)	5.55 (6.05)	5.65 (6.00)	5.70 (6.05)	5.70 (6.05)
$m_{(2,3)}^{(j)2}$	(100,30)	37.95 (39.30)	38.60 (38.70)	40.35 (40.90)	40.30 (40.40)	40.45 (40.50)	40.40 (40.60)	40.30 (40.80)
$m_{(2,4)}^{(j)2}$	(100,30)	65.80 (68.70)	67.25 (68.75)	69.20 (70.60)	70.10 (71.60)	70.30 (71.60)	70.40 (71.75)	70.50 (72.10)
$m_{(2,5)}^{(j)2}$	(100,30)	81.35 (83.75)	81.60 (83.85)	83.15 (85.60)	84.10 (86.40)	84.65 (86.65)	84.95 (87.05)	84.95 (87.05)
$m_{(2,6)}^{(j)2}$	(100,30)	88.10 (89.95)	88.25 (89.40)	89.70 (90.55)	90.30 (90.80)	90.80 (90.90)	90.95 (91.00)	91.05 (91.15)
$m_{(2,7)}^{(j)2}$	(100,30)	91.95 (92.45)	92.15 (90.40)	93.00 (91.80)	93.45 (92.50)	93.80 (92.70)	93.95 (93.20)	94.40 (93.25)
$S_{(j)}$	(100,30)	0.20 (4.35)	55.90 (54.50)	97.80 (98.20)	99.15 (99.35)	99.15 (99.30)	98.80 (99.00)	98.45 (98.60)
$\tilde{J}_{(j)}$	(100,30)	4.10 (5.55)	57.05 (45.80)	97.85 (96.55)	99.15 (98.45)	99.15 (98.45)	98.75 (98.10)	98.55 (96.90)
$m_2^{(j)}$	(150,30)	8.20 (9.30)	8.10 (9.05)	7.95 (9.25)	7.80 (8.95)	7.80 (8.95)	7.80 (8.95)	7.80 (8.95)
$m_{(2,3)}^{(j)2}$	(150,30)	52.85 (53.05)	53.85 (53.85)	55.65 (55.70)	55.50 (55.85)	55.80 (55.40)	55.70 (55.40)	55.90 (55.55)
$m_{(2,4)}^{(j)2}$	(150,30)	84.60 (85.70)	85.20 (85.95)	86.50 (87.45)	87.05 (88.15)	87.10 (87.90)	87.40 (88.20)	87.45 (88.45)
$m_{(2,5)}^{(j)2}$	(150,30)	94.90 (95.70)	95.45 (95.70)	95.90 (96.35)	96.05 (96.80)	96.20 (96.95)	96.60 (97.10)	96.70 (97.25)
$m_{(2,6)}^{(j)2}$	(150,30)	97.75 (98.35)	97.95 (98.40)	98.30 (98.70)	98.65 (98.70)	98.75 (98.95)	98.85 (99.05)	99.00 (99.05)
$m_{(2,7)}^{(j)2}$	(150,30)	99.05 (99.35)	99.05 (99.10)	99.15 (99.45)	99.30 (99.55)	99.50 (99.60)	99.50 (99.60)	99.50 (99.60)
$S_{(j)}$	(150,30)	1.65 (4.75)	74.10 (76.50)	99.75 (99.90)	100.00 (100.00)	99.95 (100.00)	100.00 (100.00)	99.90 (100.00)
$\tilde{J}_{(j)}$	(150,30)	4.25 (4.35)	74.90 (68.65)	99.75 (99.55)	100.00 (99.90)	99.95 (99.95)	100.00 (99.95)	99.90 (99.90)
$m_2^{(j)}$	(200,30)	8.10 (8.40)	8.30 (8.50)	8.05 (8.50)	7.90 (8.45)	8.10 (8.50)	8.25 (8.50)	8.25 (8.50)
$m_{(2,3)}^{(j)2}$	(200,30)	65.20 (65.70)	66.30 (66.50)	67.20 (67.80)	67.45 (67.70)	67.45 (67.90)	67.75 (67.90)	67.85 (68.20)
$m_{(2,4)}^{(j)2}$	(200,30)	94.75 (94.75)	95.40 (95.40)	96.15 (96.20)	96.20 (96.30)	96.35 (96.40)	96.40 (96.50)	96.40 (96.55)
$m_{(2,5)}^{(j)2}$	(200,30)	99.10 (99.10)	99.10 (99.15)	99.30 (99.30)	99.30 (99.30)	99.35 (99.55)	99.45 (99.50)	99.45 (99.55)
$m_{(2,6)}^{(j)2}$	(200,30)	99.80 (99.80)	99.85 (99.85)	99.90 (99.95)	99.95 (99.95)	99.95 (99.95)	99.95 (99.95)	99.95 (99.95)
$m_{(2,7)}^{(j)2}$	(200,30)	99.90 (99.90)	99.90 (99.90)	99.90 (99.95)	99.90 (99.95)	99.95 (99.95)	99.95 (100.00)	99.95 (100.00)
$S_{(j)}$	(200,30)	3.20 (6.35)	86.40 (85.95)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$\tilde{J}_{(j)}$	(200,30)	5.20 (6.40)	86.80 (78.80)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)

Notes: This table reports the power (size-adjusted power) of $m_s^{(j)}$, $m_{(2,p)}^{(j)2}$, $S_{(j)}$ and $\tilde{J}_{(j)}$ under the AR(2) alternative; for $j=0,1,2;$ $p=3,4,5,6,7$. For $j=0$, we choose $\kappa=2,3,4,5,6,7$. The data is generated as $y_{it}=\alpha_i+0.5y_{i,t-1}+0.5x_{it}+u_{it}$, where $x_{it}=0.5x_{i,t-1}+0.5u_{i,t-1}+v_{it}$, $v_{it}\sim i.i.d.N(0,\sigma_v^2)$, $i=1,\dots,N$; $t=-48,\dots,T$, with $y_{i,-49}=x_{i,-49}=0$. The first 50 observations are discarded. The signal-to-noise ratio is fixed to be 3 through σ_v^2. Under this alternative, $u_{it}=\rho_1 u_{i,t-1}+\rho_2 u_{i,t-2}+\sigma_\varepsilon \varepsilon_{it}$, where $\sigma_\varepsilon^2=(1+\rho_2)\left[(1-\rho_2)^2-\rho_1^2\right]/(1-\rho_2)$ with $\rho_1=0.2;\rho_2=0.1;$ and $\varepsilon_{it}\sim i.i.d.N(0,1)$.

Table 10: Power (Size-Adjusted Power) of Tests: Heterogeneous Slopes

		Type of Instrument Matrix/Number of Instruments							
		$j=2$	$j=1$	$j=0$					
Tests	(N,T)	T(T-2)	2(T-2)	4	6	8	10	12	14
$m_2^{(j)}$	(100,7)	17.05 (11.30)	16.35 (10.55)	17.85 (12.30)	17.70 (12.40)	17.75 (12.15)	17.40 (12.15)		
$m_{(2,3)}^{(j)2}$	(100,7)	17.70 (15.70)	10.05 (8.90)	13.80 (13.05)	14.40 (14.00)	14.70 (14.15)	14.65 (14.25)		
$m_{(2,4)}^{(j)2}$	(100,7)	16.55 (15.00)	9.00 (9.50)	11.55 (12.90)	12.05 (13.70)	12.35 (13.75)	12.45 (13.60)		
$S_{(j)}$	(100,7)	4.45 (8.95)	2.95 (3.95)	35.20 (35.90)	28.10 (30.35)	23.70 (25.75)	18.60 (21.20)		
$\tilde{J}_{(j)}$	(100,7)	6.65 (4.45)	2.90 (3.05)	36.55 (24.45)	28.65 (21.75)	23.75 (17.15)	18.50 (11.90)		
$m_2^{(j)}$	(150,7)	25.25 (24.65)	22.65 (22.90)	25.75 (26.00)	25.70 (25.70)	25.40 (25.50)	25.25 (25.35)		
$m_{(2,3)}^{(j)2}$	(150,7)	27.70 (28.80)	16.95 (18.90)	23.20 (25.50)	24.20 (26.85)	24.50 (27.75)	24.70 (27.40)		
$m_{(2,4)}^{(j)2}$	(150,7)	27.55 (28.25)	16.10 (17.55)	21.60 (23.60)	21.90 (24.80)	22.05 (24.95)	22.40 (25.00)		
$S_{(j)}$	(150,7)	13.80 (19.80)	3.05 (4.95)	52.40 (54.40)	48.30 (49.70)	41.00 (36.45)	35.00 (35.90)		
$\tilde{J}_{(j)}$	(150,7)	15.05 (11.50)	2.80 (2.85)	53.65 (41.65)	48.50 (35.65)	40.50 (28.35)	34.10 (26.05)		
$m_2^{(j)}$	(200,7)	29.80 (22.75)	26.20 (21.15)	29.45 (23.90)	29.70 (23.45)	29.45 (23.75)	29.25 (23.70)		
$m_{(2,3)}^{(j)2}$	(200,7)	35.30 (35.20)	23.10 (22.50)	30.80 (30.85)	32.15 (32.30)	32.05 (32.15)	32.05 (32.60)		
$m_{(2,4)}^{(j)2}$	(200,7)	35.90 (32.90)	20.70 (21.75)	28.30 (29.50)	30.15 (31.25)	30.85 (32.10)	30.90 (32.15)		
$S_{(j)}$	(200,7)	25.00 (26.25)	3.35 (4.90)	63.05 (63.15)	61.55 (61.65)	53.85 (54.35)	45.45 (49.70)		
$\tilde{J}_{(j)}$	(200,7)	24.20 (16.25)	3.05 (3.20)	63.65 (51.20)	61.55 (49.15)	52.95 (44.40)	44.40 (37.30)		
$m_2^{(j)}$	(100,11)	31.00 (26.75)	29.15 (26.05)	32.70 (28.65)	32.70 (28.75)	32.55 (28.50)	32.35 (28.95)	32.25 (28.00)	32.25 (28.00)
$m_{(2,3)}^{(j)2}$	(100,11)	40.55 (43.75)	32.40 (33.55)	38.05 (39.90)	40.30 (43.30)	40.05 (41.95)	40.05 (42.35)	39.85 (42.40)	39.65 (41.45)
$m_{(2,4)}^{(j)2}$	(100,11)	43.85 (42.85)	34.60 (32.70)	39.50 (39.45)	42.05 (42.00)	43.20 (43.05)	43.30 (42.80)	42.90 (42.65)	42.70 (42.80)
$m_{(2,5)}^{(j)2}$	(100,11)	43.45 (43.35)	33.80 (34.40)	38.60 (38.40)	41.15 (41.00)	42.60 (41.90)	42.80 (42.75)	43.05 (42.60)	42.60 (42.70)
$m_{(2,6)}^{(j)2}$	(100,11)	43.90 (43.90)	34.50 (36.25)	38.00 (38.65)	40.00 (41.60)	40.60 (43.20)	40.75 (43.10)	41.45 (42.95)	41.45 (42.30)
$m_{(2,7)}^{(j)2}$	(100,11)	39.90 (36.05)	30.70 (29.70)	33.80 (31.40)	35.55 (33.85)	36.40 (33.90)	36.80 (35.00)	36.80 (34.80)	36.75 (34.50)
$S_{(j)}$	(100,11)	0.00 (4.55)	1.30 (2.75)	67.55 (66.65)	67.10 (69.70)	61.95 (66.10)	55.55 (60.10)	49.30 (52.80)	44.35 (44.85)
$\tilde{J}_{(j)}$	(100,11)	2.15 (3.75)	1.90 (3.20)	68.75 (54.80)	67.35 (55.65)	61.95 (51.65)	55.35 (45.60)	49.25 (38.00)	44.60 (30.65)
$m_2^{(j)}$	(150,11)	45.70 (40.05)	41.80 (37.20)	46.30 (41.55)	46.25 (41.65)	46.30 (41.95)	46.70 (42.20)	46.75 (42.40)	46.75 (42.40)
$m_{(2,3)}^{(j)2}$	(150,11)	59.25 (60.90)	49.60 (51.75)	57.20 (58.50)	58.90 (60.80)	59.15 (61.10)	59.05 (61.60)	59.25 (61.80)	59.15 (61.25)
$m_{(2,4)}^{(j)2}$	(150,11)	66.10 (65.45)	54.25 (54.30)	60.50 (59.45)	63.25 (62.75)	64.45 (63.95)	64.65 (63.55)	64.75 (63.55)	64.65 (63.80)
$m_{(2,5)}^{(j)2}$	(150,11)	66.50 (69.30)	55.25 (57.90)	61.10 (63.50)	64.15 (67.00)	65.45 (68.55)	65.85 (69.10)	65.80 (69.15)	65.45 (69.10)
$m_{(2,6)}^{(j)2}$	(150,11)	66.75 (69.20)	55.90 (58.70)	60.35 (63.20)	63.10 (67.20)	64.35 (67.20)	64.50 (67.10)	64.75 (67.25)	64.90 (67.45)
$m_{(2,7)}^{(j)2}$	(150,11)	64.50 (65.20)	54.10 (56.05)	57.80 (59.60)	60.15 (62.60)	61.40 (64.15)	61.65 (63.70)	61.75 (64.85)	62.15 (64.15)
$S_{(j)}$	(150,11)	0.20 (10.05)	2.20 (2.65)	85.20 (85.60)	87.80 (88.45)	86.05 (87.55)	83.30 (83.80)	78.10 (81.00)	74.60 (74.65)
$\tilde{J}_{(j)}$	(150,11)	7.25 (5.45)	2.50 (3.05)	85.60 (77.60)	87.80 (82.50)	85.85 (79.70)	82.85 (76.70)	77.30 (68.75)	73.35 (62.75)
$m_2^{(j)}$	(200,11)	56.55 (53.90)	51.90 (49.90)	56.95 (53.40)	56.55 (53.90)	56.50 (53.90)	56.95 (54.20)	56.90 (54.10)	56.90 (54.10)
$m_{(2,3)}^{(j)2}$	(200,11)	73.65 (74.90)	64.30 (65.20)	71.60 (72.75)	73.55 (74.75)	73.05 (74.00)	73.35 (74.15)	73.35 (74.50)	73.05 (74.35)
$m_{(2,4)}^{(j)2}$	(200,11)	81.25 (80.90)	71.15 (70.50)	76.95 (76.00)	79.20 (78.40)	80.15 (79.70)	80.30 (79.10)	80.25 (79.25)	80.40 (79.45)
$m_{(2,5)}^{(j)2}$	(200,11)	83.15 (83.20)	72.15 (72.15)	78.40 (78.40)	80.55 (80.95)	81.90 (82.25)	82.35 (82.45)	82.10 (82.40)	82.10 (82.10)
$m_{(2,6)}^{(j)2}$	(200,11)	83.60 (84.95)	74.20 (75.55)	78.95 (81.45)	80.80 (83.35)	81.55 (83.60)	82.25 (84.10)	82.30 (84.30)	82.50 (84.30)
$m_{(2,7)}^{(j)2}$	(200,11)	83.60 (85.15)	73.45 (76.65)	78.90 (81.75)	80.65 (83.25)	80.95 (83.75)	81.30 (83.70)	81.35 (83.65)	81.45 (83.95)
$S_{(j)}$	(200,11)	7.70 (20.50)	2.85 (3.90)	93.05 (92.10)	95.95 (96.50)	95.55 (95.30)	94.40 (94.20)	91.80 (91.80)	89.60 (90.10)
$\tilde{J}_{(j)}$	(200,11)	18.50 (10.70)	2.80 (3.75)	93.40 (87.15)	95.95 (92.45)	95.30 (90.55)	94.00 (89.90)	91.25 (86.00)	88.70 (82.05)

Notes: This table reports the power (size-adjusted power) of $m_s^{(j)}$, $m_{(2,p)}^{(j)2}$, $S_{(j)}$ and $\tilde{J}_{(j)}$ under the heterogeneous slopes alternative; for $j = 0, 1, 2$; $p = 3, 4, 5, 6, 7$. For $j = 0$, we choose $\kappa = 2, 3, 4, 5, 6, 7$. The data is generated as $y_{it} = \alpha_i + 0.5 y_{i,t-1} + \beta_i x_{it} + u_{it}$, where $u_{it} \sim i.i.d.N(0, 1)$; $x_{it} = 0.5 x_{i,t-1} + 0.5 u_{i,t-1} + v_{it}$, $v_{it} \sim i.i.d.N(0, \sigma_v^2)$, $i = 1, \ldots, N$; $t = -48, \ldots, T$, with $y_{i,-49} = x_{i,-49} = 0$; $\beta_i \sim i.i.d.N(0.5, 1)$. The first 50 observations are discarded. The signal-to-noise ratio is fixed to be 3 through σ_v^2.

Table 11: Power (Size-Adjusted Power) of Tests: Heterogeneous Slopes

		Type of Instrument Matrix/Number of Instruments						
		$j=1$	$j=0$					
Tests	(N,T)	2(T-2)	4	6	8	10	12	14
$m_2^{(j)}$	(100,20)	53.40 (53.10)	59.95 (59.50)	59.95 (59.10)	60.70 (59.75)	61.10 (60.25)	61.10 (60.25)	61.10 (60.25)
$m_{(2,3)}^{(j)2}$	(100,20)	69.75 (68.80)	74.40 (73.65)	76.80 (75.10)	76.65 (75.35)	76.85 (75.75)	77.35 (76.25)	77.15 (76.15)
$m_{(2,4)}^{(j)2}$	(100,20)	79.10 (80.45)	80.00 (81.30)	82.75 (83.65)	83.35 (84.60)	83.75 (85.50)	84.10 (85.75)	84.20 (85.90)
$m_{(2,5)}^{(j)2}$	(100,20)	79.70 (81.10)	79.95 (80.35)	82.30 (82.85)	83.50 (84.25)	84.10 (85.15)	84.15 (84.75)	84.35 (85.30)
$m_{(2,6)}^{(j)2}$	(100,20)	80.70 (82.80)	80.70 (82.15)	82.45 (83.90)	83.45 (84.55)	84.35 (85.75)	84.95 (86.60)	84.85 (86.50)
$m_{(2,7)}^{(j)2}$	(100,20)	80.10 (84.75)	79.60 (82.35)	81.35 (83.45)	82.55 (84.80)	83.15 (84.80)	83.55 (84.45)	83.55 (84.70)
$S_{(j)}$	(100,20)	0.50 (2.25)	94.65 (95.15)	96.40 (96.95)	95.85 (96.90)	93.90 (95.30)	91.40 (92.15)	90.30 (92.25)
$\tilde{J}_{(j)}$	(100,20)	2.50 (3.90)	94.90 (91.05)	96.50 (94.05)	95.85 (93.15)	93.80 (90.85)	91.40 (88.25)	90.50 (85.25)
$m_2^{(j)}$	(150,20)	71.00 (68.50)	76.70 (74.35)	77.40 (74.50)	78.00 (74.85)	78.50 (75.65)	78.80 (75.60)	78.80 (75.60)
$m_{(2,3)}^{(j)2}$	(150,20)	88.95 (89.25)	92.05 (92.15)	93.70 (93.80)	93.90 (94.05)	94.00 (94.30)	94.30 (94.40)	94.30 (94.45)
$m_{(2,4)}^{(j)2}$	(150,20)	94.25 (94.55)	95.50 (95.95)	96.30 (96.65)	96.50 (97.00)	96.90 (97.15)	97.15 (97.25)	97.20 (97.30)
$m_{(2,5)}^{(j)2}$	(150,20)	95.90 (96.55)	96.30 (96.50)	97.20 (97.50)	97.55 (97.70)	97.60 (97.90)	97.65 (97.80)	97.65 (98.00)
$m_{(2,6)}^{(j)2}$	(150,20)	96.85 (96.60)	96.85 (96.45)	97.35 (96.85)	97.70 (97.00)	97.95 (97.35)	98.15 (97.75)	98.20 (97.80)
$m_{(2,7)}^{(j)2}$	(150,20)	96.60 (96.15)	96.35 (96.00)	97.15 (96.65)	97.45 (97.05)	97.60 (97.30)	97.75 (97.25)	97.80 (97.40)
$S_{(j)}$	(150,20)	0.90 (2.30)	99.40 (99.25)	99.90 (99.90)	99.80 (99.80)	99.65 (99.75)	99.30 (99.55)	99.20 (99.25)
$\tilde{J}_{(j)}$	(150,20)	1.95 (2.80)	99.40 (98.10)	99.90 (99.70)	99.80 (99.45)	99.65 (99.30)	99.30 (98.75)	99.05 (98.15)
$m_2^{(j)}$	(200,20)	85.05 (83.90)	90.30 (89.40)	90.00 (89.90)	90.65 (90.30)	90.75 (90.55)	90.95 (90.45)	90.95 (90.45)
$m_{(2,3)}^{(j)2}$	(200,20)	96.90 (95.85)	97.55 (97.10)	98.25 (97.45)	98.30 (97.70)	98.45 (97.95)	98.45 (97.95)	98.45 (98.10)
$m_{(2,4)}^{(j)2}$	(200,20)	98.95 (98.70)	99.40 (99.15)	99.60 (99.60)	99.70 (99.70)	99.70 (99.70)	99.70 (99.60)	99.70 (99.65)
$m_{(2,5)}^{(j)2}$	(200,20)	99.45 (99.45)	99.40 (99.35)	99.80 (99.50)	99.90 (99.90)	99.95 (99.90)	99.95 (99.90)	99.95 (99.90)
$m_{(2,6)}^{(j)2}$	(200,20)	99.60 (99.60)	99.70 (99.45)	99.75 (99.75)	99.85 (99.75)	99.85 (99.60)	99.85 (99.60)	99.60 (99.60)
$m_{(2,7)}^{(j)2}$	(200,20)	99.75 (99.35)	99.80 (99.45)	99.80 (99.70)	99.95 (99.80)	99.95 (99.90)	99.95 (99.90)	99.90 (99.90)
$S_{(j)}$	(200,20)	1.05 (2.50)	99.80 (99.80)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$\tilde{J}_{(j)}$	(200,20)	1.85 (4.20)	99.80 (99.60)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (99.90)
$m_2^{(j)}$	(100,30)	72.50 (73.05)	77.95 (78.15)	78.55 (78.80)	79.80 (80.25)	80.45 (80.85)	80.50 (80.50)	80.50 (80.50)
$m_{(2,3)}^{(j)2}$	(100,30)	90.25 (90.95)	92.10 (92.15)	93.80 (94.00)	93.80 (93.80)	94.05 (94.20)	94.00 (94.40)	94.25 (94.50)
$m_{(2,4)}^{(j)2}$	(100,30)	95.25 (96.00)	96.00 (96.40)	96.90 (97.20)	97.30 (97.75)	97.35 (97.65)	97.35 (97.65)	97.40 (97.80)
$m_{(2,5)}^{(j)2}$	(100,30)	96.05 (97.05)	96.05 (96.95)	97.10 (97.95)	97.65 (98.25)	97.65 (98.45)	97.70 (98.70)	97.70 (98.60)
$m_{(2,6)}^{(j)2}$	(100,30)	96.20 (97.25)	95.60 (96.00)	96.70 (97.10)	97.15 (97.30)	97.30 (97.45)	97.40 (97.50)	97.50 (97.75)
$m_{(2,7)}^{(j)2}$	(100,30)	96.45 (96.85)	95.65 (94.45)	96.30 (95.85)	96.75 (95.95)	97.25 (96.20)	97.20 (96.60)	97.15 (96.50)
$S_{(j)}$	(100,30)	0.00 (1.85)	99.10 (99.00)	99.85 (99.90)	99.65 (99.80)	99.10 (99.25)	98.45 (98.90)	98.00 (98.15)
$\tilde{J}_{(j)}$	(100,30)	3.35 (5.20)	99.20 (98.25)	99.85 (99.60)	99.65 (98.90)	99.10 (98.40)	98.45 (97.40)	98.05 (95.95)
$m_2^{(j)}$	(150,30)	87.25 (86.80)	91.80 (91.55)	92.45 (91.95)	93.15 (93.05)	93.30 (93.45)	93.45 (93.60)	93.45 (93.60)
$m_{(2,3)}^{(j)2}$	(150,30)	98.05 (98.20)	99.05 (99.05)	99.30 (99.30)	99.35 (99.35)	99.25 (99.25)	99.40 (99.35)	99.40 (99.40)
$m_{(2,4)}^{(j)2}$	(150,30)	99.60 (99.65)	99.65 (99.65)	99.65 (99.70)	99.70 (99.75)	99.75 (99.75)	99.75 (99.75)	99.75 (99.75)
$m_{(2,5)}^{(j)2}$	(150,30)	99.80 (99.90)	99.75 (99.85)	99.90 (99.95)	99.95 (100.00)	99.95 (100.00)	99.95 (100.00)	99.95 (100.00)
$m_{(2,6)}^{(j)2}$	(150,30)	99.80 (99.85)	99.75 (99.75)	99.80 (99.80)	99.80 (99.90)	99.90 (99.95)	99.80 (99.95)	99.80 (99.95)
$m_{(2,7)}^{(j)2}$	(150,30)	99.80 (99.80)	99.70 (99.75)	99.75 (99.75)	99.80 (99.80)	99.80 (99.85)	99.80 (99.85)	99.80 (99.85)
$S_{(j)}$	(150,30)	0.40 (1.80)	99.90 (99.90)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	99.95 (99.95)
$\tilde{J}_{(j)}$	(150,30)	2.85 (3.90)	99.90 (99.75)	100.00 (99.95)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	99.95 (99.95)
$m_2^{(j)}$	(200,30)	95.05 (94.30)	97.65 (97.25)	97.90 (97.45)	98.10 (97.80)	98.20 (97.95)	98.25 (98.00)	98.25 (98.00)
$m_{(2,3)}^{(j)2}$	(200,30)	99.50 (99.60)	99.80 (99.80)	99.95 (99.95)	99.95 (99.95)	99.95 (99.95)	99.95 (99.95)	99.95 (99.95)
$m_{(2,4)}^{(j)2}$	(200,30)	99.90 (99.90)	99.90 (99.95)	99.95 (99.95)	99.95 (99.95)	99.95 (99.95)	99.95 (99.95)	99.95 (99.95)
$m_{(2,5)}^{(j)2}$	(200,30)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,6)}^{(j)2}$	(200,30)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$m_{(2,7)}^{(j)2}$	(200,30)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$S_{(j)}$	(200,30)	0.55 (1.50)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)
$\tilde{J}_{(j)}$	(200,30)	2.50 (3.85)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)	100.00 (100.00)

Notes: This table reports the power (size-adjusted power) of $m_s^{(j)}$, $m_{(2,p)}^{(j)2}$, $S_{(j)}$ and $\tilde{J}_{(j)}$ under the heterogeneous slopes alternative; for $j = 0, 1, 2$; $p = 3, 4, 5, 6, 7$. For $j = 0$, we choose $\kappa = 2, 3, 4, 5, 6, 7$. The data is generated as $y_{it} = \alpha_i + 0.5y_{i,t-1} + \beta_i x_{it} + u_{it}$, where $u_{it} \sim i.i.d.N(0, 1)$; $x_{it} = 0.5x_{i,t-1} + 0.5u_{i,t-1} + v_{it}$, $v_{it} \sim i.i.d.N(0, \sigma_v^2)$, $i = 1, \ldots, N$; $t = -48, \ldots, T$, with $y_{i,-49} = x_{i,-49} = 0$; $\beta_i \sim i.i.d.N(0.5, 1)$. The first 50 observations are discarded. The signal-to-noise ratio is fixed to be 3 through σ_v^2.

Table 12: Power (Size-Adjusted Power) of Tests: Cross-Sectional Dependence (Factor)

		Type of Instrument Matrix/Number of Instruments							
		$j=2$	$j=1$	$j=0$					
Tests	(N,T)	T(T-2)	2(T-2)	4	6	8	10	12	14
$m_2^{(j)}$	(100,7)	35.65 (34.70)	35.80 (34.80)	35.65 (34.30)	35.80 (34.75)	35.65 (35.10)	35.65 (34.80)		
$m_{(2,3)}^{(j)2}$	(100,7)	49.85 (48.15)	49.05 (47.50)	48.90 (48.25)	49.10 (48.75)	49.10 (48.60)	49.15 (49.05)		
$m_{(2,4)}^{(j)2}$	(100,7)	55.85 (54.00)	52.20 (52.60)	51.55 (53.00)	52.00 (54.00)	52.20 (54.15)	52.05 (53.25)		
$S_{(j)}$	(100,7)	38.15 (49.00)	22.45 (23.95)	25.45 (25.75)	31.70 (33.55)	35.60 (36.95)	35.85 (37.85)		
$\tilde{J}_{(j)}$	(100,7)	44.80 (37.20)	22.45 (17.15)	26.15 (20.00)	32.10 (27.65)	35.60 (31.70)	35.85 (30.55)		
$m_2^{(j)}$	(150,7)	41.00 (41.90)	40.70 (41.85)	40.80 (42.15)	40.85 (42.20)	40.70 (42.30)	40.65 (42.20)		
$m_{(2,3)}^{(j)2}$	(150,7)	56.40 (57.30)	54.90 (56.00)	54.75 (56.35)	54.90 (56.55)	54.85 (56.95)	55.15 (56.90)		
$m_{(2,4)}^{(j)2}$	(150,7)	63.30 (63.50)	60.15 (60.85)	60.15 (61.35)	60.40 (61.75)	60.75 (62.15)	60.80 (62.50)		
$S_{(j)}$	(150,7)	64.70 (68.75)	31.65 (37.65)	32.05 (32.80)	40.30 (41.25)	45.00 (43.20)	45.75 (45.90)		
$\tilde{J}_{(j)}$	(150,7)	65.85 (62.15)	31.15 (25.90)	32.75 (26.80)	40.40 (34.30)	44.75 (38.35)	45.35 (41.55)		
$m_2^{(j)}$	(200,7)	45.95 (45.15)	46.25 (44.60)	45.95 (44.50)	45.80 (44.55)	45.85 (44.85)	45.85 (44.95)		
$m_{(2,3)}^{(j)2}$	(200,7)	62.50 (62.35)	61.70 (61.40)	61.15 (61.20)	61.20 (61.45)	61.20 (61.20)	61.15 (61.45)		
$m_{(2,4)}^{(j)2}$	(200,7)	69.45 (68.60)	67.80 (68.25)	67.15 (67.65)	67.35 (67.80)	67.85 (68.10)	67.80 (68.35)		
$S_{(j)}$	(200,7)	77.30 (78.00)	38.95 (39.35)	35.60 (35.65)	47.05 (47.20)	53.90 (54.50)	55.30 (58.05)		
$\tilde{J}_{(j)}$	(200,7)	76.85 (73.00)	38.35 (33.10)	35.90 (30.00)	47.05 (41.75)	53.50 (49.70)	54.70 (51.35)		
$m_2^{(j)}$	(100,11)	36.60 (35.85)	36.10 (37.05)	36.25 (36.05)	36.30 (36.10)	36.20 (36.15)	36.30 (36.15)	36.20 (35.85)	36.20 (35.85)
$m_{(2,3)}^{(j)2}$	(100,11)	51.95 (53.65)	51.80 (52.45)	51.80 (53.00)	51.95 (53.20)	52.15 (53.10)	52.10 (53.35)	51.85 (53.15)	51.85 (53.00)
$m_{(2,4)}^{(j)2}$	(100,11)	60.50 (59.90)	60.55 (59.40)	60.10 (60.10)	60.35 (60.35)	60.85 (60.85)	60.60 (60.25)	60.75 (60.65)	60.55 (60.60)
$m_{(2,5)}^{(j)2}$	(100,11)	67.05 (66.80)	66.25 (66.65)	65.80 (65.80)	66.30 (66.25)	66.25 (66.00)	66.45 (66.40)	66.35 (66.00)	66.70 (66.80)
$m_{(2,6)}^{(j)2}$	(100,11)	71.65 (71.65)	71.40 (72.05)	71.40 (71.80)	71.60 (72.15)	71.70 (72.20)	71.60 (72.20)	71.70 (72.05)	71.60 (71.80)
$m_{(2,7)}^{(j)2}$	(100,11)	74.20 (72.05)	73.55 (73.00)	73.70 (72.70)	73.70 (73.05)	73.55 (72.75)	73.45 (72.70)	73.80 (72.75)	73.80 (72.85)
$S_{(j)}$	(100,11)	0.00 (7.60)	26.65 (34.25)	27.30 (26.95)	37.30 (39.10)	42.95 (46.70)	46.85 (50.20)	47.75 (49.65)	50.85 (51.15)
$\tilde{J}_{(j)}$	(100,11)	0.60 (4.45)	27.20 (23.25)	27.95 (22.45)	37.55 (32.10)	42.95 (38.20)	46.75 (41.00)	47.65 (41.90)	50.90 (42.05)
$m_2^{(j)}$	(150,11)	44.90 (44.65)	45.00 (45.45)	45.20 (45.05)	45.20 (44.85)	45.00 (44.70)	44.95 (44.65)	44.95 (44.80)	44.95 (44.80)
$m_{(2,3)}^{(j)2}$	(150,11)	61.45 (62.50)	61.15 (62.35)	61.25 (62.05)	61.20 (62.35)	61.20 (62.45)	61.35 (62.55)	61.25 (62.75)	61.50 (62.60)
$m_{(2,4)}^{(j)2}$	(150,11)	70.70 (70.55)	70.55 (70.55)	70.55 (70.45)	70.45 (70.05)	70.45 (70.30)	70.60 (70.10)	70.50 (70.15)	70.65 (70.20)
$m_{(2,5)}^{(j)2}$	(150,11)	77.25 (77.90)	77.25 (77.90)	77.05 (77.85)	77.00 (77.85)	77.05 (77.95)	77.15 (78.00)	77.25 (77.95)	77.20 (78.05)
$m_{(2,6)}^{(j)2}$	(150,11)	81.75 (82.45)	81.60 (82.35)	81.55 (82.60)	81.45 (82.55)	81.45 (82.35)	81.50 (82.15)	81.50 (82.25)	81.55 (82.40)
$m_{(2,7)}^{(j)2}$	(150,11)	84.20 (84.35)	84.15 (84.65)	83.60 (84.05)	83.65 (84.40)	83.90 (84.50)	84.00 (84.45)	84.10 (84.50)	84.00 (84.40)
$S_{(j)}$	(150,11)	10.80 (53.35)	44.10 (46.40)	36.40 (37.00)	49.45 (50.15)	57.10 (58.75)	61.55 (61.95)	63.80 (65.40)	66.65 (66.75)
$\tilde{J}_{(j)}$	(150,11)	47.25 (43.80)	43.50 (37.90)	36.90 (30.80)	49.60 (44.25)	56.85 (49.95)	61.05 (56.50)	63.25 (57.50)	65.95 (59.20)
$m_2^{(j)}$	(200,11)	52.15 (52.25)	52.20 (52.25)	51.90 (52.35)	52.00 (52.70)	52.10 (52.55)	52.25 (52.50)	52.30 (52.55)	52.30 (52.55)
$m_{(2,3)}^{(j)2}$	(200,11)	68.85 (69.50)	68.45 (68.65)	68.15 (68.80)	68.25 (69.05)	68.25 (68.70)	68.35 (68.85)	68.35 (69.00)	68.35 (69.00)
$m_{(2,4)}^{(j)2}$	(200,11)	77.25 (77.20)	77.25 (76.85)	77.05 (76.45)	76.80 (76.45)	76.90 (76.60)	76.85 (76.35)	76.90 (76.40)	76.90 (76.55)
$m_{(2,5)}^{(j)2}$	(200,11)	83.70 (83.70)	83.85 (83.80)	83.55 (83.55)	83.60 (83.65)	83.45 (83.65)	83.45 (83.65)	83.50 (83.65)	83.55 (83.60)
$m_{(2,6)}^{(j)2}$	(200,11)	88.00 (88.50)	87.70 (88.10)	87.30 (88.20)	87.50 (88.35)	87.70 (88.40)	87.55 (88.35)	87.60 (88.55)	87.60 (88.45)
$m_{(2,7)}^{(j)2}$	(200,11)	90.55 (91.25)	90.55 (91.25)	90.30 (91.15)	90.35 (91.55)	90.30 (91.35)	90.35 (91.25)	90.40 (91.15)	90.40 (91.15)
$S_{(j)}$	(200,11)	72.70 (83.75)	57.85 (61.65)	43.90 (42.65)	60.00 (60.85)	68.85 (68.50)	72.00 (71.75)	76.15 (76.15)	78.80 (79.20)
$\tilde{J}_{(j)}$	(200,11)	83.15 (76.75)	56.85 (53.45)	44.65 (38.20)	60.00 (54.40)	68.60 (61.00)	71.60 (66.55)	75.45 (69.95)	77.80 (73.75)

Notes: This table reports the power (size-adjusted power) of $m_s^{(j)}$, $m_{(2,p)}^{(j)2}$, $S_{(j)}$ and $\tilde{J}_{(j)}$ under the cross-sectional dependence (factor) alternative; for $j = 0,1,2$; $p = 3,4,5,6,7$. For $j = 0$, we choose $\kappa = 2,3,4,5,6,7$. The data is generated as $y_{it} = \alpha_i + 0.5y_{i,t-1} + 0.5x_{it} + u_{it}$, where $x_{it} = 0.5x_{i,t-1} + 0.5u_{i,t-1} + v_{it}$, $v_{it} \sim i.i.d.N(0,\sigma_v^2)$, $i = 1,\ldots,N$; $t = -48,\ldots,T$, with $y_{i,-49} = x_{i,-49} = 0$. The first 50 observations are discarded. The signal-to-noise ratio is fixed to be 3 through σ_v^2. Under this alternative, $u_{it} = 0.75\left(\lambda_i f_t + \sigma_\varepsilon \varepsilon_{it}\right)$, where $\lambda_i \sim i.i.d.U[-1,1]$, $f_t \sim i.i.d.N(0,1)$, $\sigma_\varepsilon = 1$; and $\varepsilon_{it} \sim i.i.d.N(0,1)$.

Table 13: Power (Size-Adjusted Power) of Tests: Cross-Sectional Dependence (Factor)

		Type of Instrument Matrix/Number of Instruments						
		$j=1$	$j=0$					
Tests	(N,T)	2(T-2)	4	6	8	10	12	14
$m_2^{(j)}$	(100,20)	44.55 (45.85)	44.50 (45.45)	44.60 (45.75)	44.55 (45.75)	44.60 (45.90)	44.70 (45.80)	44.70 (45.80)
$m_{(2,3)}^{(j)2}$	(100,20)	59.15 (58.55)	59.40 (59.05)	59.35 (58.45)	59.50 (58.65)	59.45 (59.05)	59.55 (58.65)	59.55 (58.75)
$m_{(2,4)}^{(j)2}$	(100,20)	67.45 (68.20)	67.65 (68.60)	67.55 (68.55)	67.65 (68.35)	67.85 (68.85)	67.60 (68.80)	67.55 (68.65)
$m_{(2,5)}^{(j)2}$	(100,20)	73.25 (73.80)	74.00 (74.45)	74.00 (74.45)	73.65 (74.55)	73.60 (74.50)	73.70 (74.15)	73.60 (74.15)
$m_{(2,6)}^{(j)2}$	(100,20)	78.10 (79.15)	78.70 (79.10)	78.80 (79.35)	78.95 (79.30)	78.65 (79.10)	78.80 (79.60)	78.80 (79.65)
$m_{(2,7)}^{(j)2}$	(100,20)	81.70 (83.65)	82.40 (83.50)	82.40 (83.45)	82.40 (83.50)	82.25 (83.50)	82.25 (82.95)	82.30 (82.90)
$S_{(j)}$	(100,20)	23.15 (38.10)	33.95 (35.80)	46.15 (48.30)	52.45 (55.15)	56.50 (59.30)	60.70 (61.80)	64.70 (67.80)
$\tilde{J}_{(j)}$	(100,20)	29.55 (30.80)	34.85 (29.90)	46.80 (40.35)	52.45 (46.80)	56.45 (51.55)	60.70 (56.70)	64.75 (57.80)
$m_2^{(j)}$	(150,20)	51.45 (52.25)	51.45 (52.25)	51.45 (52.10)	51.55 (52.20)	51.45 (52.15)	51.45 (52.15)	51.45 (52.15)
$m_{(2,3)}^{(j)2}$	(150,20)	70.10 (69.70)	70.10 (70.00)	70.05 (69.15)	69.90 (69.20)	70.05 (69.55)	69.95 (69.35)	70.20 (69.65)
$m_{(2,4)}^{(j)2}$	(150,20)	79.35 (80.05)	79.20 (79.80)	79.50 (79.75)	79.20 (79.80)	79.35 (80.10)	79.50 (80.00)	79.40 (80.10)
$m_{(2,5)}^{(j)2}$	(150,20)	85.40 (86.05)	85.30 (85.65)	85.15 (85.60)	85.30 (85.55)	85.20 (85.45)	85.15 (85.35)	85.20 (85.60)
$m_{(2,6)}^{(j)2}$	(150,20)	88.55 (89.05)	88.85 (89.25)	88.90 (89.35)	88.75 (89.10)	88.85 (89.15)	89.00 (89.45)	88.95 (89.50)
$m_{(2,7)}^{(j)2}$	(150,20)	90.75 (91.85)	90.90 (91.60)	91.05 (91.50)	91.05 (91.65)	91.00 (91.65)	91.15 (91.60)	91.15 (91.65)
$S_{(j)}$	(150,20)	57.00 (67.95)	43.90 (45.10)	59.30 (61.25)	68.70 (70.85)	74.55 (76.95)	78.30 (79.50)	81.35 (82.85)
$\tilde{J}_{(j)}$	(150,20)	58.20 (55.65)	44.45 (36.25)	59.50 (54.00)	68.60 (61.25)	74.15 (71.05)	77.90 (73.95)	80.85 (76.45)
$m_2^{(j)}$	(200,20)	55.95 (55.05)	55.75 (55.05)	55.80 (55.00)	55.95 (55.05)	56.00 (55.25)	55.95 (55.10)	55.95 (55.10)
$m_{(2,3)}^{(j)2}$	(200,20)	74.00 (72.35)	74.55 (72.25)	74.45 (72.05)	74.80 (72.50)	74.70 (72.70)	74.65 (72.70)	74.60 (72.65)
$m_{(2,4)}^{(j)2}$	(200,20)	83.95 (82.65)	83.60 (82.65)	83.70 (83.00)	83.75 (82.90)	83.85 (83.00)	83.70 (82.75)	83.80 (82.85)
$m_{(2,5)}^{(j)2}$	(200,20)	88.25 (87.75)	88.30 (87.30)	88.45 (87.60)	88.30 (87.60)	88.30 (87.60)	88.35 (87.60)	88.25 (87.80)
$m_{(2,6)}^{(j)2}$	(200,20)	92.20 (92.30)	92.10 (91.65)	92.15 (91.95)	92.05 (91.95)	92.10 (91.85)	92.20 (91.85)	92.35 (91.95)
$m_{(2,7)}^{(j)2}$	(200,20)	93.85 (93.30)	93.70 (92.80)	93.65 (92.90)	93.75 (93.00)	93.80 (93.10)	93.80 (93.05)	93.80 (93.25)
$S_{(j)}$	(200,20)	76.90 (81.85)	47.25 (48.75)	64.80 (66.90)	75.10 (76.05)	81.20 (81.05)	85.30 (85.55)	88.15 (87.20)
$\tilde{J}_{(j)}$	(200,20)	76.80 (76.60)	47.50 (43.70)	64.80 (61.85)	74.95 (70.95)	80.65 (76.50)	84.85 (82.05)	87.75 (84.15)
$m_2^{(j)}$	(100,30)	42.15 (43.25)	42.60 (42.90)	42.85 (43.35)	42.65 (43.25)	42.65 (43.00)	42.60 (43.05)	42.60 (43.05)
$m_{(2,3)}^{(j)2}$	(100,30)	58.80 (59.60)	59.65 (59.65)	59.65 (60.00)	59.40 (59.65)	59.40 (59.75)	59.35 (59.85)	59.45 (59.75)
$m_{(2,4)}^{(j)2}$	(100,30)	68.55 (70.10)	69.40 (70.85)	69.10 (70.40)	69.15 (70.35)	69.05 (70.75)	69.20 (70.65)	69.15 (70.70)
$m_{(2,5)}^{(j)2}$	(100,30)	75.35 (76.75)	76.30 (77.70)	76.65 (77.70)	76.45 (77.70)	76.50 (77.75)	76.15 (77.70)	76.25 (77.70)
$m_{(2,6)}^{(j)2}$	(100,30)	79.70 (81.45)	80.40 (81.05)	80.60 (81.45)	80.55 (81.05)	80.65 (81.10)	80.70 (80.95)	80.70 (80.95)
$m_{(2,7)}^{(j)2}$	(100,30)	83.50 (84.45)	84.20 (82.75)	84.45 (83.10)	84.15 (82.50)	83.95 (82.75)	84.10 (83.05)	84.20 (82.70)
$S_{(j)}$	(100,30)	6.55 (35.70)	35.25 (34.60)	48.20 (50.20)	57.80 (61.00)	63.15 (65.10)	66.15 (68.45)	70.70 (71.50)
$\tilde{J}_{(j)}$	(100,30)	23.05 (26.75)	36.45 (30.30)	48.60 (43.95)	57.80 (52.95)	63.10 (57.75)	66.10 (61.50)	70.70 (63.65)
$m_2^{(j)}$	(150,30)	53.30 (54.00)	53.80 (54.20)	53.95 (54.45)	53.85 (54.45)	53.95 (54.55)	53.95 (54.55)	53.95 (54.55)
$m_{(2,3)}^{(j)2}$	(150,30)	71.25 (71.45)	71.35 (71.35)	71.20 (71.25)	71.25 (71.35)	71.25 (71.05)	71.25 (71.10)	71.25 (70.95)
$m_{(2,4)}^{(j)2}$	(150,30)	80.20 (81.15)	80.65 (81.10)	80.30 (81.35)	80.20 (81.35)	80.35 (81.00)	80.40 (81.05)	80.45 (81.35)
$m_{(2,5)}^{(j)2}$	(150,30)	86.00 (86.80)	86.60 (86.95)	86.60 (87.10)	86.60 (87.05)	86.65 (87.15)	86.50 (87.30)	86.40 (87.25)
$m_{(2,6)}^{(j)2}$	(150,30)	90.10 (90.70)	90.10 (90.90)	90.25 (90.80)	90.25 (90.80)	90.15 (90.80)	90.35 (90.65)	90.20 (90.65)
$m_{(2,7)}^{(j)2}$	(150,30)	92.00 (93.35)	92.30 (92.35)	92.25 (93.00)	92.35 (92.90)	92.45 (92.90)	92.40 (92.95)	92.35 (93.10)
$S_{(j)}$	(150,30)	55.85 (70.95)	45.15 (46.80)	61.70 (62.20)	72.55 (75.25)	77.85 (79.70)	82.30 (84.40)	85.70 (86.85)
$\tilde{J}_{(j)}$	(150,30)	62.35 (61.25)	45.45 (42.45)	61.80 (54.65)	72.45 (69.10)	77.40 (75.65)	81.80 (79.25)	85.55 (81.70)
$m_2^{(j)}$	(200,30)	58.75 (58.15)	58.75 (57.90)	58.90 (57.95)	58.65 (57.95)	58.75 (57.95)	58.90 (57.85)	58.90 (57.85)
$m_{(2,3)}^{(j)2}$	(200,30)	77.00 (77.35)	77.00 (77.10)	77.00 (77.25)	77.10 (77.20)	77.25 (77.35)	77.20 (77.20)	77.15 (77.25)
$m_{(2,4)}^{(j)2}$	(200,30)	87.20 (87.35)	87.05 (87.05)	86.95 (87.15)	86.90 (87.05)	86.95 (87.20)	86.90 (87.25)	87.00 (87.30)
$m_{(2,5)}^{(j)2}$	(200,30)	92.20 (92.95)	92.70 (92.80)	92.75 (92.90)	92.75 (93.10)	92.70 (93.05)	92.75 (92.85)	92.65 (92.75)
$m_{(2,6)}^{(j)2}$	(200,30)	94.65 (94.65)	94.45 (94.45)	94.40 (94.50)	94.55 (94.55)	94.55 (94.50)	94.60 (94.55)	94.65 (94.65)
$m_{(2,7)}^{(j)2}$	(200,30)	96.40 (96.60)	96.45 (96.85)	96.40 (96.75)	96.45 (96.65)	96.50 (96.60)	96.40 (96.60)	96.40 (96.65)
$S_{(j)}$	(200,30)	81.25 (87.75)	51.55 (51.15)	70.40 (70.30)	80.10 (80.10)	87.90 (87.75)	90.45 (90.60)	93.20 (93.20)
$\tilde{J}_{(j)}$	(200,30)	83.25 (81.75)	51.95 (45.20)	70.40 (65.10)	79.70 (77.25)	87.75 (83.65)	90.10 (88.00)	93.05 (91.55)

Notes: This table reports the power (size-adjusted power) of $m_s^{(j)}$, $m_{(2,p)}^{(j)2}$, $S_{(j)}$ and $\tilde{J}_{(j)}$ under the cross-sectional dependence (factor) alternative; for $j = 0, 1, 2$; $p = 3, 4, 5, 6, 7$. For $j = 0$, we choose $\kappa = 2, 3, 4, 5, 6, 7$. The data is generated as $y_{it} = \alpha_i + 0.5y_{i,t-1} + 0.5x_{it} + u_{it}$, where $x_{it} = 0.5x_{i,t-1} + 0.5u_{i,t-1} + v_{it}$, $v_{it} \sim i.i.d.N(0, \sigma_v^2)$, $i = 1, \ldots, N$; $t = -48, \ldots, T$, with $y_{i,-49} = x_{i,-49} = 0$. The first 50 observations are discarded. The signal-to-noise ratio is fixed to be 3 through σ_v^2. Under this alternative, $u_{it} = 0.75\left(\lambda_i f_t + \sigma_\varepsilon \varepsilon_{it}\right)$, where $\lambda_i \sim i.i.d.U[-1, 1]$, $f_t \sim i.i.d.N(0, 1)$, $\sigma_\varepsilon = 1$; and $\varepsilon_{it} \sim i.i.d.N(0, 1)$.

Appendix

This appendix includes proofs of the main results in the text. The appendix includes t-wo parts: Part A includes some useful lemmas which are frequently used in the proofs of Theorems; Part B gives the proofs of all the theorems included in the paper.

A Some Useful Lemmas

Lemma A.1 *Under the Assumptions 1-3, we have*

(a) $(NT_s)^{-1/2}\sum_{i=1}^{N}\sum_{t=3}^{T-s}\Delta u_{i,t+s}Z_{it}'\left(\hat{\theta}_{FDGMM2}^{(j)}-\theta\right)=o_p\left(1\right),\ s=2,\ldots,p,\ j=0,1,2;$

(b) $(NT_s)^{-1/2}\sum_{i=1}^{N}\sum_{t=3}^{T-s}Z_{it}'\left(\hat{\theta}_{FDGMM2}^{(j)}-\theta\right)\left(\hat{\theta}_{FDGMM2}^{(j)}-\theta\right)'Z_{i,t+s}=o_p\left(1\right),\ s=2,\ldots,p,$ $j=0,1,2.$

Proof of Lemma A.1.

Consider part (a).

$$
\begin{aligned}
&\left(\frac{1}{NT_s}\right)^{1/2}\sum_{i=1}^{N}\sum_{t=3}^{T-s}\Delta u_{i,t+s}Z_{it}'\left(\hat{\theta}_{FDGMM2}^{(j)}-\theta\right)\\
=&\left(\frac{1}{NT_s}\right)^{1/2}\sum_{i=1}^{N}\sum_{t=3}^{T-s}\left(u_{i,t+s}-u_{i,t+s-1}\right)\left(Z_{it}-Z_{i,t-1}\right)'\left(\hat{\theta}_{FDGMM2}^{(j)}-\theta\right)\\
=&\left(\frac{1}{NT_s}\right)^{1/2}\sum_{i=1}^{N}\sum_{t=3}^{T-s}\left(u_{i,t+s}Z_{it}'-u_{i,t+s}Z_{i,t-1}'-u_{i,t+s-1}Z_{it}'+u_{i,t+s-1}Z_{i,t-1}'\right)\left(\hat{\theta}_{FDGMM2}^{(j)}-\theta\right).
\end{aligned}
$$

Consider the term $(NT_s)^{-1/2}\sum_{i=1}^{N}\sum_{t=3}^{T-s}u_{i,t+s}Z_{it}'$, it can be rewritten as

$$
\left(\frac{1}{NT_s}\right)^{1/2}\sum_{i=1}^{N}\sum_{t=3}^{T-s}u_{i,t+s}Z_{it}'=\left(\frac{1}{NT_s}\right)^{1/2}\sum_{i=1}^{N}\sum_{t=3}^{T-s}\left(u_{i,t+s}y_{i,t-1},u_{i,t+s}x_{it}\right)'.
$$

To calculate the term $(NT_s)^{-1/2}\sum_{i=1}^{N}\sum_{t=3}^{T-s}u_{i,t+s}y_{i,t-1}$.

$$\begin{aligned}
&\text{var}\left[\left(\frac{1}{NT_s}\right)^{1/2}\sum_{i=1}^{N}\sum_{t=3}^{T-s}u_{i,t+s}y_{i,t-1}\right]\\
=&\frac{1}{NT_s}\sum_{i=1}^{N}var\left[\sum_{t=3}^{T-s}u_{i,t+s}y_{i,t-1}\right]\\
=&\frac{1}{NT_s}\sum_{i=1}^{N}E\left(\sum_{t=3}^{T-s}\sum_{\tau=3}^{T-s}u_{i,t+s}y_{i,t-1}u_{i,\tau+s}y_{i,\tau-1}\right)\\
=&\frac{1}{NT_s}\sum_{i=1}^{N}E\left[\sum_{t=3}^{T-s}\left(u_{i,t+s}y_{i,t-1}\right)^2\right]+\frac{2}{NT_s}\sum_{i=1}^{N}E\left(\sum_{t=3}^{T-s}\sum_{t<\tau}u_{i,t+s}y_{i,t-1}u_{i,\tau+s}y_{i,\tau-1}\right)\\
=&\frac{1}{NT_s}\sum_{i=1}^{N}E\left[\sum_{t=3}^{T-s}\left(u_{i,t+s}y_{i,t-1}\right)^2\right].
\end{aligned}$$

By Assumptions 2, we obtain $\mathrm{E}\left[\sum_{t=3}^{T-s}\left(u_{i,t+s}y_{i,t-1}\right)^2\right]=O_p\left(T_s\right)$, therefore

$$\text{var}\left[\left(\frac{1}{NT_s}\right)^{1/2}\sum_{i=1}^{N}\sum_{t=3}^{T-s}u_{i,t+s}y_{i,t-1}\right]=O_p\left(1\right).$$

Similarly, using Assumptions 1-2, we have $\mathrm{E}\left(\|x_{it}\|^2\right)=O_p(1)$, which further leads to

$$\left(\frac{1}{NT_s}\right)^{1/2}\left\|\sum_{i=1}^{N}\sum_{t=3}^{T-s}u_{i,t+s}x_{it}\right\|=O_p\left(1\right).$$

The above results together result in $(NT_s)^{-1/2}\left\|\sum_{i=1}^{N}\sum_{t=3}^{T-s}u_{i,t+s}Z_{it}'\right\|=O_p(1)$. By using the same procedures, and with Assumptions 1-2, we obtain

$$\left(\frac{1}{NT_s}\right)^{1/2}\left\|\sum_{i=1}^{N}\sum_{t=3}^{T-s}u_{i,t+s}Z_{i,t-1}'\right\|=O_p\left(1\right);$$

$$\left(\frac{1}{NT_s}\right)^{1/2}\left\|\sum_{i=1}^{N}\sum_{t=3}^{T-s}u_{i,t+s-1}Z_{it}'\right\|=O_p\left(1\right);$$

and

$$\left(\frac{1}{NT_s}\right)^{1/2}\left\|\sum_{i=1}^{N}\sum_{t=3}^{T-s}u_{i,t+s-1}Z'_{i,t-1}\right\|=O_p(1).$$

With the fact that $\left\|\hat{\theta}^{(j)}_{FDGMM2}-\theta\right\|=O_p\left[(NT)^{-1/2}\right]$, we get part (a).

For part (b). From Assumptions 1-2, we have $\|Z'_{it}Z_{i,t+s}\|=O_p(1)$, hence $\left\|\sum_{i=1}^{N}\sum_{t=3}^{T-s}Z'_{it}Z_{i,t+s}\right\|=O_p(NT_s)$ and $\left\|\hat{\theta}^{(j)}_{FDGMM2}-\theta\right\|=O_p\left[(NT)^{-1/2}\right]$, the term on the left-hand side of part (b) is $O_p\left[(NT_s)^{-1/2}\right]$. Part (b) holds.

Lemma A.2 *Under the Assumptions 1-3, we have*

(a) $\sup_T T_s^{-2}E\Big(\sum_{t=3}^{T-s}\Delta u_{it}\Delta u_{i,t+s}\Big)^4=O(1),\ s=2,\ldots,p;$

(b) $\sup_T T_s^{-2}E\Big(\sum_{t=s+1}^{T-s}y_{i,t-s}\Delta u_{i,t}\Big)^4=O(1),\ s=2,\ldots,max\{p,\kappa\};$

(c) $\sup_T T_s^{-2}E\Big(\sum_{t=s+1}^{T-s}x_{i,t-s+1,j}\Delta u_{i,t}\Big)^4=O(1),$ *for* $j=1,\ldots,k-1,\ s=2,\ldots,max\{p,\kappa\}.$

Proof of Lemma A.2.

Consider part (a), for $s=2,\ldots,p$;

$$\begin{aligned}&\frac{1}{T_s^2}\sup_T \mathrm{E}\left(\sum_{t=3}^{T-s}\Delta u_{i,t}\Delta u_{i,t+s}\right)^4\\
=&\frac{1}{T_s^2}\sup_T \mathrm{E}\left[\sum_{t=3}^{T-s}(u_{it}u_{i,t+s}+u_{i,t-1}u_{i,t+s-1}-u_{it}u_{i,t+s-1}-u_{i,t-1}u_{i,t+s})\right]^4\\
\leq&\frac{4^3}{T_s^2}\sup_T \mathrm{E}\left(\sum_{t=3}^{T-s}u_{it}u_{i,t+s}\right)^4+\frac{4^3}{T_s^2}\sup_T \mathrm{E}\left(\sum_{t=3}^{T-s}u_{i,t-1}u_{i,t+s-1}\right)^4\\
&+\frac{4^3}{T_s^2}\sup_T \mathrm{E}\left(\sum_{t=3}^{T-s}u_{it}u_{i,t+s-1}\right)^4+\frac{4^3}{T_s^2}\sup_T \mathrm{E}\left(\sum_{t=3}^{T-s}u_{i,t-1}u_{i,t+s}\right)^4.\end{aligned}$$

To prove part (a), we consider the above four terms. For the first term,

$$
\begin{aligned}
&\sup_T \frac{4^3}{T_s^2}\mathrm{E}\left(\sum_{t=3}^{T-s} u_{it}u_{i,t+s}\right)^4 \\
=&\sup_T \frac{4^3}{T_s^2}\mathrm{E}\left[\sum_{t=3}^{T-s}\left(u_{it}u_{i,t+s}\right)^4\right]+\sup_T \frac{4^3}{T_s^2}\mathrm{E}\left[\sum_{(t\neq\tau)=3}^{T-s}\left(u_{it}u_{i,t+s}\right)^2\left(u_{i\tau}u_{i,\tau+s}\right)^2\right]=O\left(1\right).
\end{aligned}
$$

Similarly, we can show the other three terms are bounded. Part (a) holds.

Consider part (b). For any $s=2,\ldots,max\{p,\kappa\}$;

$$
\begin{aligned}
&\sup_T \frac{1}{T_s^2}\mathrm{E}\left(\sum_{t=3}^{T-s} y_{i,t-s}\Delta u_{i,t}\right)^4 \\
=&\sup_T \frac{1}{T_s^2}\mathrm{E}\left(\sum_{t=3}^{T-s} y_{i,t-s}u_{it}-\sum_{t=3}^{T-s} y_{i,t-s}u_{i,t-1}\right)^4 \\
\leq&\sup_T \frac{8}{T_s^2}\mathrm{E}\left(\sum_{t=3}^{T-s} y_{i,t-s}u_{it}\right)^4+\sup_T \mathrm{E}\left(\sum_{t=3}^{T-s} y_{i,t-s}u_{i,t-1}\right)^4.
\end{aligned}
$$

To calculate the above two terms, we apply the Rosenthal's inequality for point process martingales (Wood (1999)). Define $S_\tau=\sum_{t=s+1}^{\tau-s} y_{i,t-s}u_{it}$ and $\mathcal{F}_\tau$ as the σ-field generated by $\{u_1,\ldots u_\tau\}$. $X_{i,\tau}=S_\tau-S_{\tau-1}$ is a martingale difference sequence with respect to $\mathcal{F}_{\tau-1}$. Applying the Rosenthal's inequality given by Wood (1999), we have

$$
\begin{aligned}
&\sup_{\tau\in T}\mathrm{E}|S_\tau|^4 \\
\leq&C_4\sup_{\tau\in T}\mathrm{E}\left\{\left(\sum_{t=s+1}^{\tau-s} E\left[\left(y_{i,t-s}u_{it}\right)^2|\mathcal{F}_{t-1}\right]\right)^2+\sum_{t=s+1}^{\tau-s}\left(y_{i,t-s}u_{it}\right)^4\right\} \\
=&C_4\left[\sup_{\tau\in T}\mathrm{E}\left(\sum_{t=s+1}^{\tau-s}\sigma^2 y_{i,t-s}^2\right)^2+\sup_{\tau\in T}\sum_{t=s+1}^{\tau-s} E\left(y_{i,t-s}u_{it}\right)^4\right] \\
=&C_4\left[\sup_{\tau\in T}\mathrm{E}\left(\sum_{t=s+1}^{\tau-s}\sigma^4 y_{i,t-s}^4\right)+\sup_{\tau\in T}\mathrm{E}\left(\sum_{t=s+1,t\neq\eta}^{\tau-s}\sigma^4 y_{i,t-s}^2 y_{i,\eta-s}^2\right)+\sup_{\tau\in T}\sum_{t=s+1}^{\tau-s}\mathrm{E}\left(y_{i,t-s}u_{it}\right)^4\right];
\end{aligned}
$$

where C_4 is a constant finite positive number. By Assumption 2, we can easily show that $\mathrm{E}\left(y_{i,t-s}^4\right)$, $\mathrm{E}\left(y_{i,t-s}^2 y_{i,\eta-s}^2\right)$ and $\mathrm{E}(y_{i,t-s}u_{it})^4$ are bounded for all t. Hence, $\sup_T \mathrm{E}|S_T|^4 \leq O(T_s^2)$, which implies $\sup_T 8T_s^{-2}\mathrm{E}\left(\sum_{t=3}^{T-s} y_{i,t-s}u_{it}\right)^4 = O\left(1\right)$. Similarly, we obtain $\sup_T 8T_s^{-2}$ $\mathrm{E}\left(\sum_{t=3}^{T-s} y_{i,t-s}u_{i,t-1}\right)^4 = O(1)$. Hence, $\sup_T T_s^{-2}\mathrm{E}\left(\sum_{t=s+1}^{T-s} y_{it-s}\Delta u_{it}\right)^4 = O\left(1\right)$. Part (c) can be proved similar to part (b).

Consider part (d). For any $s = 2, \ldots, max\{p, \kappa\}$;

$$\begin{aligned}
&\sup_T \frac{1}{T_s^2}\mathrm{E}\left(\sum_{t=s+1}^{T-s} x_{i,t-s+1,j}\Delta u_{it}\right)^4 \\
=&\sup_T \frac{8}{T_s^2}\mathrm{E}\left[\sum_{t=s+1}^{T-s}\left(x_{i,t-s+1,j}u_{it} + x_{i,t-s+1,j}u_{i,t-1}\right)\right]^4 \\
\leq&\sup_T \frac{8}{T_s^2}\mathrm{E}\left(\sum_{t=s+1}^{T-s} x_{i,t-s+1,j}u_{it}\right)^4 + \sup_T \frac{8}{T_s^2}\mathrm{E}\left(\sum_{t=s+1}^{T-s} x_{i,t-s+1,j}u_{i,t-1}\right)^4 .
\end{aligned}$$

Applying Rosenthal's inequality, we have the above two terms are bounded, which further proves part (c). Part (e) can be proved similar to part (d).

Lemma A.3 *Under Assumptions 1-3, as $(N,T) \to \infty$, we have*

(a) $N^{-1}\sum_{i=1}^N \left((T_s)^{-1/2}\sum_{t=3}^{T-s}\Delta\hat{u}_{it}^{(j)}\Delta\hat{u}_{i,t+s}^{(j)}\right)^2 - E\left((T_s)^{-1/2}\sum_{t=3}^{T-s}\Delta\hat{u}_{it}^{(j)}\Delta\hat{u}_{i,t+s}^{(j)2}\right) = o_p\left(1\right)$, $s = 2, \ldots, p$, $j = 0, 1, 2$;

(b) $\left\|(NT_s)^{-1}\sum_{i=1}^N\left(W_{li}^{(0)\prime}\Delta\check{u}_i^{(0)}\Delta\check{u}_i^{(0)\prime}W_{li}^{(0)}\right) - E\left(w_{iTs}^{(0)}w_{iTs}^{(0)\prime}\right)\right\| = o_p\left(1\right)$, $s = 2, \ldots, max\{p, \kappa\}$;

(c) $\left\|(NT_s)^{-1}\sum_{i=1}^N\sum_{t=3}^{T-s}\Delta\hat{u}_{it}^{(j)}Z_{i,t+s}^{\prime} - E\left(\Delta u_{it}Z_{i,t+s}^{\prime}\right)\right\| = o_p\left(1\right)$, $s = 2, \ldots, max\{p, \kappa\}$, $j = 0, 1, 2$.

Proof of Lemma A.3.

Consider part (a), since

$$\begin{aligned}
&\left(\frac{1}{T_s}\right)^{1/2}\sum_{t=3}^{T-s}\Delta\hat{u}_{it}^{(j)}\Delta\hat{u}_{i,t+s}^{(j)} \\
&=\left(\frac{1}{T_s}\right)^{1/2}\sum_{t=3}^{T-s}\Delta u_{it}\Delta u_{i,t+s}-\left(\frac{1}{T_s}\right)^{1/2}\sum_{t=3}^{T-s}\Delta u_{it}Z_{i,t+s}'\left(\hat{\theta}_{FDGMM2}^{(j)}-\theta\right) \\
&\quad-\left(\frac{1}{T_s}\right)^{1/2}\sum_{t=3}^{T-s}\Delta u_{i,t+s}Z_{it}'\left(\hat{\theta}_{FDGMM2}^{(j)}-\theta\right) \\
&\quad+\left(\frac{1}{T_s}\right)^{1/2}\sum_{t=3}^{T-s}Z_{it}'\left(\hat{\theta}_{FDGMM2}^{(j)}-\theta\right)Z_{i,t+s}'\left(\hat{\theta}_{FDGMM2}^{(j)}-\theta\right).
\end{aligned}$$

Since $T_s^{-1}\sum_{t=3}^{T-s}\Delta u_{it}Z_{i,t+s}'=O_p(1)$ and $\left\|\hat{\theta}_{FDGMM2}^{(j)}-\theta\right\|=O_p\left[(NT)^{-1/2}\right]$, the second term of above equation is $O_p\left(N^{-1/2}\right)$. The third and the fourth terms are of smaller order than the second term. Those lead to $(T_s)^{1/2}\sum_{t=3}^{T-s}\Delta\hat{u}_{it}^{(j)}\Delta\hat{u}_{i,t+s}^{(j)}=(T_s)^{-1/2}\sum_{t=3}^{T-s}\Delta u_{it}\Delta u_{i,t+s}+O_p\left(N^{-1/2}\right)$, which further implies

$$\left((T_s)^{-1/2}\sum_{t=3}^{T-s}\Delta\hat{u}_{it}^{(j)}\Delta\hat{u}_{i,t+s}^{(j)}\right)^2=\left((T_s)^{-1/2}\sum_{t=3}^{T-s}\Delta u_{it}^{(j)}\Delta u_{i,t+s}^{(j)}\right)^2+O_p\left(N^{-1}\right).$$

Since $(T_s)^{-1/2}\sum_{t=3}^{T-s}\Delta u_{it}^{(j)}\Delta u_{i,t+s}^{(j)}$ are independent across i, part (a) holds.

Consider part (b). To prove it, we only need to show that $\left(T_s^{-1}W_{li}^{(0)\prime}\Delta\breve{u}_i^{(0)}\Delta\breve{u}_i^{(0)\prime}W_{li}^{(0)}\right)\xrightarrow{p}w_{iTs}^{(0)}w_{iTs}^{(0)\prime}$. Since

$$\left(\frac{1}{T_s}\right)^{1/2}W_{li}^{(0)\prime}\Delta\breve{u}_i^{(0)}=\left(\frac{1}{T_s}\right)^{1/2}\sum_{t=3}^{T-s}W_{lit}^{(0)}\Delta u_{it}-\left(\frac{1}{T_s}\right)^{1/2}\sum_{t=3}^{T-s}W_{lit}^{(0)}\Delta Z_{it}'\left(\hat{\theta}_{FDGMM1}^{(0)}-\theta\right).$$

Using the facts that $\left\|\sum_{t=3}^{T-s}W_{lit}^{(0)}\Delta Z_{it}'\right\|=O_p(T_s)$ and $\left\|\hat{\theta}_{FDGMM1}^{(0)}-\theta\right\|=O_p\left[(NT)^{-1/2}\right]$, we conclude that the second term of above is $O_p\left(N^{-1/2}\right)$. Hence $\left(T_s^{-1}W_{li}^{(0)\prime}\Delta\breve{u}_i^{(0)}\Delta\breve{u}_i^{(0)\prime}W_{li}^{(0)}\right)\xrightarrow{p}w_{iTs}^{(0)}w_{iTs}^{(0)\prime}$, then part (b) holds.

Consider part (c). Since

$$\Delta\hat{u}_{it}^{(j)} Z_{i,t+s}' = \Delta u_{it} Z_{i,t+s}' + Z_{it}' \left(\hat{\theta}_{FDGMM2}^{(j)} - \theta\right) Z_{i,t+s}'.$$

By using the fact that $\left\|\hat{\theta}_{FDGMM2}^{(j)} - \theta\right\| = O_p\left[(NT)^{-1/2}\right]$, we can easily show part (c) holds.

B Proof of Theorems

Proof of Theorem 1

Define $\hat{r}_s^{(0)} = N^{-1/2}\sum_{i=1}^{N}\hat{v}_{iTs}^{(0)}$. Recall that $\Delta\hat{u}_t^{(0)} = \Delta y_{i,t} - Z_{it}'\hat{\theta}_{FDGMM2}^{(0)} = \Delta u_{it} - Z_{it}'\left(\hat{\theta}_{FDGMM2}^{(0)} - \theta\right)$, hence

$$\begin{aligned}
\hat{r}_s^{(0)} =& \left(\frac{1}{N}\right)^{1/2}\sum_{i=1}^{N} v_{iTs} - \left(\frac{1}{NT_s}\right)^{1/2}\sum_{i=1}^{N}\sum_{t=3}^{T-s}\Delta u_{it}Z_{i,t+s}'\left(\hat{\theta}_{FDGMM2}^{(0)} - \theta\right) \\
&- \left(\frac{1}{NT_s}\right)^{1/2}\sum_{i=1}^{N}\sum_{t=3}^{T-s}\Delta u_{i,t+s}Z_{it}'\left(\hat{\theta}_{FDGMM2}^{(0)} - \theta\right) \\
&+ \left(\frac{1}{NT_s}\right)^{1/2}\sum_{i=1}^{N}\sum_{t=3}^{T-s}Z_{it}'\left(\hat{\theta}_{FDGMM2}^{(0)} - \theta\right)Z_{i,t+s}'\left(\hat{\theta}_{FDGMM2}^{(0)} - \theta\right).
\end{aligned}$$

where $v_{iTs} = T_s^{-1/2}\sum_{t=3}^{T-s}\Delta u_{it}\Delta u_{i,t+s}$. By Lemma A.1, the third term and the fourth term are $o_p(1)$. $\hat{r}_s$ can be rewritten as

$$\begin{aligned}
\hat{r}_s &= \left(\frac{1}{N}\right)^{1/2}\sum_{i=1}^{N} v_{iTs} - \Xi_1'\left(\frac{1}{N}\right)^{1/2}\sum_{i=1}^{N} w_{iTs}^{(0)} + o_p(1) \\
&= (1, -\Xi_1')\, N^{-1/2}\sum_{i=1}^{N}\begin{pmatrix} v_{iTs} \\ w_{iTs}^{(0)} \end{pmatrix} + o_p(1).
\end{aligned}$$

where $w_{iTs}^{(0)} = T_s^{-1/2}W_{li}^{(0)}\Delta u_i$ and $\Xi_1' = \varpi_{Ns}^{(0)\prime}\hat{\Psi}_N^{-1(0)}\left[(NT_s)^{-1}Z'W^{(0)}\hat{\Omega}_{(0)}^{-1}\right]$. Define $Q_{iTs}^{(0)} = \left(v_{iTs}, w_{iTs}^{(0)\prime}\right)'$, and $\Sigma_{2s}^{(0)} = \mathrm{E}\left(Q_{iTs}^{(0)}Q_{iTs}^{(0)\prime}\right)$. To derive the limiting distribution of the test statistics with large T, we apply Theorem 3 (the Joint Limit CLT for Scaled Variates) of Phillips and Moon (1999). To apply this theorem, we need to verify the following conditions:

(i) $\liminf_T \lambda_{\min}\left(\Sigma_{2s}^{(0)}\right) > 0$ and $\Sigma_{2s}^{(0)}$ are positive definite;

(ii) $\left\|Q_{iTs}^{(0)}\right\|^2$ are uniformly integral in T.

By Assumption 3, condition (i) holds. Next we consider $\left\|Q_{iTs}^{(0)}\right\|^2$.

$$\left\|Q_{iTs}^{(0)}\right\|^2 = \operatorname{tr}\left(Q_{iTs}^{(0)}Q_{iTs}^{(0)\prime}\right) = v_{iTs}v_{iTs}' + \operatorname{tr}\left(w_{iTs}^{(0)}w_{iTs}^{(0)\prime}\right).$$

Let $\delta > 0$,

$$\begin{aligned}
&\sup_T \mathrm{E}\left(\left[v_{iTs}v_{iTs}' + \operatorname{tr}\left(w_{iTs}^{(0)}w_{iTs}^{(0)\prime}\right)\right] I_{\left(v_{iTs}v_{iTs}' + \operatorname{tr}\left(w_{iTs}^{(0)}w_{iTs}^{(0)\prime}\right) > \delta\right)}\right) \\
\leq &\sup_T \sqrt{\mathrm{E}\left[v_{iTs}v_{iTs}' + \operatorname{tr}\left(w_{iTs}^{(0)}w_{iTs}^{(0)\prime}\right)\right]^2}\sqrt{P(\left|v_{iTs}v_{iTs}' + \operatorname{tr}\left(w_{iTs}^{(0)}w_{iTs}^{(0)\prime}\right)\right| > \delta)} \\
\leq &\frac{1}{\delta}\sup_T \sqrt{\mathrm{E}\left[v_{iTs}v_{iTs}' + \operatorname{tr}\left(w_{iTs}^{(0)}w_{iTs}^{(0)\prime}\right)\right]^2}\operatorname{var}\sqrt{v_{iTs}v_{iTs}' + \operatorname{tr}\left(w_{iTs}^{(0)}w_{iTs}^{(0)\prime}\right)} \\
\leq &\frac{1}{\delta}\sup_T \mathrm{E}\left[v_{iTs}^2 + \operatorname{tr}\left(w_{iTs}^{(0)}w_{iTs}^{(0)\prime}\right)\right]^2 \\
\leq &\frac{2}{\delta}\left[\sup_T \mathrm{E}\left(v_{iTs}^4\right) + \sup_T \mathrm{E}\left[\operatorname{tr}\left(w_{iTs}^{(0)}w_{iTs}^{(0)\prime}\right)\right]^2\right].
\end{aligned}$$

Consider the term $\mathrm{E}\left[\operatorname{tr}\left(w_{iTs}^{(0)}w_{iTs}^{(0)\prime}\right)\right]^2$.

$$\begin{aligned}
&\mathrm{E}\left[\operatorname{tr}\left(w_{iTs}^{(0)}w_{iTs}^{(0)\prime}\right)\right]^2 \\
=&\frac{1}{T_s^2}\mathrm{E}\left(\sum_{t=3}^{T-s} y_{i,t-2}\Delta u_{i,t} + \ldots + \sum_{t=\kappa+1}^{T-s} y_{i,t-\kappa}\Delta u_{i,t} + \sum_{t=3}^{T-s} x_{i,t-1}'\Delta u_{i,t} + \ldots + \sum_{t=\kappa+1}^{T-s} x_{i,t-\kappa+1}'\Delta u_{i,t}\right)^4 \\
\leq&\frac{K}{T_s^2}\left[\mathrm{E}\left(\sum_{t=3}^{T-s} y_{it-2}\Delta u_{i,t}\right)^4 + \ldots + \mathrm{E}\left(\sum_{t=\kappa+1}^{T-s} y_{i,t-\kappa}\Delta u_{i,t}\right)^4\right] \\
&+\frac{K}{T_s^2}\left[\mathrm{E}\left(\sum_{t=3}^{T-s} x_{i,t-1}'\Delta u_{i,t}\right)^4 + \ldots + \mathrm{E}\left(\sum_{t=\kappa}^{T-s} x_{i,t-\kappa+1}'\Delta u_{i,t}\right)^4\right].
\end{aligned}$$

From part (b) and part (c) of Lemma A.2, we have $\sup_T T_s^{-2}\mathrm{E}\Big(\sum_{t=s+1}^{T-s} y_{i,t-s}\Delta u_{i,t}\Big)^4, s = 2, \ldots, \kappa$, are bounded. Next, we consider the left terms term. For any $s = 2, \ldots, \kappa$;

$$
\begin{aligned}
&\frac{1}{T_s^2}\mathrm{E}\left(\sum_{t=s+1}^{T-s} \Delta x'_{i,t-s+1}\Delta u_{it}\right)^4 \\
=&\frac{1}{T_s^2}\mathrm{E}\left(\sum_{t=s+1}^{T-s} x_{i.t-s+1,1}\Delta u_{it} + \ldots + \sum_{t=s+1}^{T-s} x_{i,t-s+1,k-1}\Delta u_{i,t}\right)^4 \\
\leq&\frac{(k-1)^3}{T_s^2}\left[\mathrm{E}\left(\sum_{t=s+1}^{T-s} x_{i,t-s+1,1}\Delta u_{it}\right)^4 + \ldots + \mathrm{E}\left(\sum_{t=s+1}^{T-s} (x_{i,t-s+1,k-1}\Delta u_{it})\right)^4\right].
\end{aligned}
$$

Using part (c) of Lemma A.2, $\sup_T T_s^{-2}\mathrm{E}\Big(\sum_{t=s+1}^{T-s} \Delta x'_{i,t-s+1}\Delta u_{it}\Big)^4 = O(1)$. All the above results lead to $\sup_T \mathrm{E}\Big[\mathrm{tr}\left(w_{iTs}^{(0)}w_{iTs}^{(0)\prime}\right)\Big]^2 = O(1)$. From part (a) of Lemma A.2**,** we also have $\sup_T \mathrm{E}(v_{iTs}^2) = O(1)$, hence

$$
\sup_T \mathrm{E}\left(\left[v_{iTs}v'_{iTs} + \mathrm{tr}\left(w_{iTs}^{(0)}w_{iTs}^{(0)\prime}\right)\right] I_{\left(v_{iTs}v'_{iTs}+\mathrm{tr}\left(w_{iTs}^{(0)}w_{iTs}^{(0)\prime}\right)>\delta\right)}\right) \to 0, \text{as } \delta \to \infty;
$$

which makes condition (ii) hold. With the satisfaction of those two conditions, as $(N, T) \to \infty$,

$$
\Sigma_{2s}^{(0)-1/2}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\begin{pmatrix} v_{iTs} \\ w_{iTs}^{(0)} \end{pmatrix} \xrightarrow{d} N(0, I_{(\kappa-1)k+1}).
$$

By using part (a) and part (b) of Lemma A.3**,** we have $\hat{\Sigma}_{2s}^{(0)} \xrightarrow{p} \Sigma_{2s}^{(0)}$, with

$$
\hat{\Sigma}_{2s}^{(0)} = \frac{1}{N}\sum_{i=1}^{N}\begin{pmatrix} \hat{v}_{iTs}^{(0)2} & \hat{v}_{iTs}^{(0)\prime}\hat{w}_{iTs}^{(0)} \\ \hat{v}_{iTs}^{(0)}\hat{w}_{iTs}^{(0)\prime} & \hat{w}_{iTs}^{(0)}\hat{w}_{iTs}^{(0)\prime} \end{pmatrix}.
$$

Since $\hat{r}_s = (1, -\Xi'_1)\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\binom{v_{iTs}}{w_{iTs}^{(0)}} + o_p(1)$, by using part (c) of Lemma A.3, $\hat{\varpi}_{Ns}^{(0)\prime} \xrightarrow{p} \varpi_{Ns}^{(0)\prime}$, therefore $\hat{\Xi}'_1 \xrightarrow{p} \Xi'_1$, where $\hat{\Xi}'_1 = \hat{\varpi}_{Ns}^{(0)\prime}\hat{\Psi}_N^{-1(0)}\left[(NT_s)^{-1} Z'W^{(0)}\hat{\Omega}_{(0)}^{-1}\right]$. Hence together with the result above, we have

$$
m_s^{(0)} = \hat{\gamma}_{(0)}^{-1}\hat{r}_s \xrightarrow{d} N(0, 1),
$$

where

$$\hat{\gamma}_{(0)}^2 = \left(1, -\hat{\Xi}_1'\right) \hat{\Sigma}_2 \left(1, -\hat{\Xi}_1'\right)'.$$

Proof of Theorem 2

Similar to the proof of Theorem 1, we can rewrite $\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \hat{v}_{ip}^{(0)}$ as,

$$\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \hat{v}_{ip}^{(0)} = [I_{p-1}, -\Xi_2'] \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \begin{pmatrix} v_{ip} \\ w_{iT2} \end{pmatrix} + o_p(1),$$

where $\Xi_2' = \varpi \hat{\Psi}_N^{(0)-1} \left[(NT_s)^{-1} Z' W^{(0)} \hat{\Omega}_{(0)}^{-1} \right]$, where $\varpi = \left((T_2/T_2)^{1/2} \varpi_{N2}, \ldots, (T_2/T_p)^{1/2} \varpi_{Np} \right)$ with $\varpi_{Ns} = (NT_s)^{-1} \sum_{i=1}^{N} \sum_{t=3}^{T-s} \Delta u_{it} \Delta Z_{i,t+s}$, $s = 2, \ldots, p$. Define $Q_i = \left(v_{ip}', w_{iT2}'\right)'$. Since $\mathrm{tr}\left(v_{ip} v_{ip}'\right) = \sum_{s=2}^{p} v_{iTs}^2$, hence

$$\sup_T \mathrm{E} \left[\mathrm{tr}\left(v_{ip} v_{ip}'\right) \right]^2 \leq \sup_T \mathrm{E} \left(\sum_{s=2}^{p} v_{iTs} \right)^4 \leq (p-1)^4 \sup_T \sum_{s=2}^{p} \mathrm{E}\left(v_{iTs}\right)^4 = O(1).$$

Combining the results that $\sup_T \mathrm{E}\left[\mathrm{tr}\left(w_{iT2}^{(0)} w_{iT2}^{(0)\prime} \right) \right]^2$ is bounded, $\left\| Q_i^{(0)} \right\|^2$ are uniformly integrable in T. Using part (a) and part (b) of Lemma A.3, we have $\hat{\Sigma}_1^{(0)} \xrightarrow{p} \Sigma_1^{(0)}$,where

$$\hat{\Sigma}_1^{(0)} = \frac{1}{N} \sum_{i=1}^{N} \begin{pmatrix} \hat{v}_{ip}^{(0)} \hat{v}_{ip}^{(0)\prime} & \hat{v}_{ip}^{(0)\prime} \hat{w}_{iT2}^{(0)} \\ \hat{v}_{ip}^{(0)} \hat{w}_{iT2}^{(0)\prime} & \hat{w}_{iT2}^{(0)} \hat{w}_{iT2}^{(0)\prime} \end{pmatrix}.$$

Applying Theorem 3 (the Joint Limit CLT for Scaled Variates) of Phillips and Moon (1999), and as $(N, T) \to \infty$,

$$\hat{\Sigma}_1^{(0)-1/2} \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \begin{pmatrix} v_{ip} \\ w_{iT2}^{(0)} \end{pmatrix} \xrightarrow{d} N\left(0, I_{(\kappa-1)k+p-1}\right).$$

Since $\hat{\varpi}_{Ns}^{(0)} \xrightarrow{p} \varpi_{Ns}$, then $\hat{\varpi}^{(0)} \xrightarrow{p} \varpi$,where $\hat{\varpi}^{(0)} = \left(\hat{\varpi}_{N2}^{(0)}, \ldots, \hat{\varpi}_{Np}^{(0)}\right)$, which further implies $\hat{\Xi}_2' \xrightarrow{p} \Xi_2'$, with $\hat{\Xi}_2 = \hat{\varpi}^{(0)} \hat{\Psi}_N^{(0)-1} \left[(NT_s)^{-1} Z' W^{(0)} \hat{\Omega}_{(0)}^{-1}\right]$. Then we obtain

$$\hat{\gamma}_2^{-1}[I_{p-1}, -\hat{\Xi}_2'] \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \begin{pmatrix} v_i \\ w_{iT}^{(0)} \end{pmatrix} \xrightarrow{d} N(0, I_{p-1}).$$

where $\hat{\gamma}_2^2 = \left(1, -\hat{\Xi}_2'\right) \hat{\Sigma}_2 \left(1, -\hat{\Xi}_2'\right)'$. This directly shows, as $(N,T) \to \infty$,

$$m_{(2,p)}^{(0)2} \xrightarrow{d} \chi_{p-1}^2.$$

Proof of Corollary 1

For $j = 1, 2$, we have

$$\begin{aligned}
&\left(\frac{1}{T}\right)^{1/2} \hat{v}_{iTs}^{(j)} \\
&= \left(\frac{1}{T}\right)^{1/2} \sum_{t=3}^{T-2} \Delta \hat{u}_{it}^{(j)} \Delta \hat{u}_{i,t+2}^{(j)} \\
&= \left(\frac{1}{T}\right)^{1/2} \sum_{t=3}^{T-2} \Delta u_{it} \Delta u_{i,t+2} - \left(\frac{1}{T}\right)^{1/2} \sum_{t=3}^{T-2} \Delta u_{it} Z_{i,t+2}' (\hat{\theta}_{FDGMM2}^{(j)} - \theta) + o_p(1) \\
&= \left(\frac{1}{T}\right)^{1/2} \sum_{t=3}^{T-2} \Delta u_{it} \Delta u_{i,t+2} + O_p(N^{-1/2}).
\end{aligned}$$

Since $\left\| T^{-1} \sum_{t=3}^{T-2} \Delta u_{it} Z_{i,t+2}' \right\| = O_p(1)$ and $\left\| (\hat{\theta}_{FDGMM2}^{(2)} - \theta) \right\| = O_p(1/\sqrt{NT})$, we then have

$$\left(\frac{1}{T}\right)^{1/2} \hat{v}_{iTs}^{(j)} = \left(\frac{1}{T}\right)^{1/2} \sum_{t=3}^{T-2} \Delta u_{it} \Delta u_{i,t+2} + O_p(N^{-1/2}).$$

Next we consider $\frac{1}{T} \hat{\gamma}_{(j)}^2$.

$$
\begin{aligned}
\frac{1}{T}\hat{\gamma}_{(j)}^{2} =& \frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{\sqrt{T}}\hat{v}_{iTs}^{(j)}\right)^{2} + \frac{1}{T}\hat{\varpi}_{Ns}^{(j)\prime}\hat{\Psi}_{N}^{(j)-1}\hat{\varpi}_{Ns}^{(j)} \\
& - 2\hat{\varpi}_{Ns}^{(j)\prime}\hat{\Psi}_{N}^{(j)-1}\left(\frac{Z'W^{(j)}\hat{\Omega}_{(j)}^{-1}}{N}\right)\left(\frac{1}{NT}\sum_{i=1}^{N}W_{li}^{(j)\prime}\Delta\breve{u}_{i}^{(j)}\hat{v}_{iTs}^{(j)}\right).
\end{aligned}
$$

with $\hat{\varpi}_{Ns}^{(j)} = 1/N\sum_{i=1}^{N}\left(\sum_{t=3}^{T-s}\Delta\hat{u}_{it}^{(j)}Z_{i,t+s}\right)$ and

$$
\hat{\Psi}_{N}^{(j)} = \left[1/N\sum_{i=1}^{N}W_{li}^{(j)\prime}Z_{i}\right]'\hat{\Omega}_{(j)}^{-1}\left[1/N\sum_{i=1}^{N}W_{li}^{(j)\prime}Z_{i}\right]
$$

. Using Lemma A.1., we have $1/(NT)\sum_{i=1}^{N}\left(\sum_{t=3}^{T-s}\Delta\hat{u}_{it}^{(j)}Z_{i,t+s}\right) = O_{p}\left((NT)^{-1/2}\right)$. Since $\left\|1/T^{2}\hat{\Omega}_{(j)}\right\| = O_{p}(1)$ and $1/(NT)\sum_{i=1}^{N}W_{li}^{(j)\prime}Z_{i} = O_{p}(1)$, then $\left\|\hat{\Psi}_{N}^{(j)}\right\| = O_{p}(1)$. Therefore, $T^{-1}\hat{\varpi}_{Ns}^{(j)\prime}\hat{\Psi}_{N}^{(j)-1}\hat{\varpi}_{Ns}^{(j)} = O_{p}(N^{-1})$. For the last term in the above equation, we have

$$
\left\|\frac{Z'W^{(j)}\hat{\Omega}_{(j)}^{-1}}{N}\right\| = O_{p}(1) \text{ and } \left\|\frac{1}{NT}\sum_{i=1}^{N}W_{li}^{(j)\prime}\Delta\breve{u}_{i}^{(j)}\hat{v}_{iTs}^{(j)}\right\| = O_{p}(1).
$$

Hence, the last term is $O_{p}\left((NT)^{-1/2}\right)$. Hence $\frac{1}{T}\hat{\gamma}_{(j)}^{2} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{\sqrt{T}}\hat{v}_{iTs}^{(j)}\right)^{2} + O_{p}\left(\frac{1}{N}\right) + O_{p}\left(\frac{1}{\sqrt{NT}}\right)$. From

$$
\hat{\gamma}_{(j)}^{2} \xrightarrow{p} \frac{1}{N}\sum_{i=1}^{N}\left[\left(\frac{1}{T}\right)^{1/2}\sum_{t=3}^{T-2}\Delta u_{it}\Delta u_{i,t+2}\right]^{2} \xrightarrow{p} \text{var}\left[\left(\frac{1}{T}\right)^{1/2}\sum_{t=3}^{T-2}\Delta u_{it}\Delta u_{i,t+2}\right].
$$

By using the joint limit CLT for scaled variates of Phillips and Moon (1999), we have

$$
\text{var}^{1/2}\left[\left(\frac{1}{T}\right)^{1/2}\sum_{t=3}^{T-2}\Delta u_{it}\Delta u_{i,t+2}\right]\frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{T}\right)^{1/2}\sum_{t=3}^{T-2}\Delta u_{it}\Delta u_{i,t+2} \xrightarrow{d} N(0,1).
$$

With the results above, we have

$$m_2^{(j)} = \hat{\gamma}_{(j)}^{-1} \frac{1}{N} \sum_{i=1}^{N} \hat{v}_{iTs}^{(j)} \xrightarrow{d} N(0,1).$$

Further, similar to the proof of Theorem 2, we have

$$m_{(2,p)}^{(j)2} \xrightarrow{d} \chi_{p-1}^2.$$

Proof of Theorem 3

Let $\hat{\Omega}_{(0)}^{-1} = C_{NT} C_{NT}'$, where C_N is a $q_0 \times q_0$ non-singular matrix. Note that

$$\begin{aligned}
&(NT_0)^{-1/2} \sum_{i=1}^{N} W_{li}^{(0)\prime} \Delta \hat{u}_{2i} \\
=& (NT_0)^{-1/2} \sum_{i=1}^{N} W_{li}^{(0)\prime} \left[\Delta u_i - Z_i' \left(\hat{\theta}_{FDGMM2}^{(0)} - \theta \right) \right] \\
=& \left(I_{q_0} - W^{(0)\prime} Z \left[Z' W^{(0)} \hat{\Omega}_{(0)}^{-1} W^{(0)\prime} Z \right]^{-1} Z' W^{(0)} \hat{\Omega}_{(0)}^{-1} \right) (NT_0)^{-1/2} \sum_{i=1}^{N} W_{li}^{(0)\prime} \Delta u_i \\
=& \left(I_{q_0} - \frac{W^{(0)\prime} Z}{NT_0} \left[\frac{Z' W^{(0)}}{NT_0} C_N C_N' \frac{W^{(0)\prime} Z}{NT_0} \right]^{-1} \frac{Z' W^{(0)}}{NT_0} C_N C_N' \right) (NT_0)^{-1/2} \sum_{i=1}^{N} W_{li}^{(0)\prime} \Delta u_i.
\end{aligned}$$

then

$$\begin{aligned}
S_{(0)} =& \left[(NT_0)^{-1/2} \sum_{i=1}^{N} W_{li}^{(0)\prime} \Delta u_i \right]' C_{NT} \left[I_{q_0} - D_{NT} \left(D_{NT}' D_{NT} \right)^{-1} D_{NT}' \right] \\
& \times C_{NT}' \left[(NT_0)^{-1/2} \sum_{i=1}^{N} W_{li}^{(0)\prime} \Delta u_i \right],
\end{aligned}$$

where $D_{NT} = C_{NT}' \frac{W^{(0)\prime} Z}{NT_0}$. From the proof of theorem 1, we have $\sup_T \mathrm{tr} \left[\left(T_0^{-1} W_{li}^{(0)\prime} \Delta u_i \Delta u_i' W_{li}^{(0)} \right) \right] = O(1)$; and with the Assumption 3, we have the two conditions satisfying the joint limit CLT for scaled variates of Phillips and Moon (1999), together with part (b) of Lemma A.3; as

$(N, T) \to \infty$,

$$\hat{\Omega}_{(0)}^{-1/2} (NT_0)^{-1/2} \sum_{i=1}^{N} W_{li}^{(0)\prime} \Delta u_i \xrightarrow{d} N(0, I_{q_0}).$$

Since $\text{rank}\left[I_{q_0} - D_{NT}(D_{NT}' D_{NT})^{-1} D_{NT}'\right] = q_0 - k$, using the argument that $\varepsilon' M \varepsilon \xrightarrow{d} \chi^2_{q_0 - k}$ where $\varepsilon \sim N(0, I_{q_0})$ and $M = I_{p_1} - D_{NT}(D_{NT}' D_{NT})^{-1} D_{NT}'$. Hence, as $(N, T) \to \infty$, $S_{(0)} \xrightarrow{d} \chi^2_{q_0 - k}$.

Proof of Theorem 4

Similar to the proof of Theorem 3, $S_{(0)I}$ can be rewritten as

$$\begin{aligned} S_{(0)I} = & \left[(NT_0)^{-1/2} \sum_{i=1}^{N} W_{li1}^{(0)\prime} \Delta u_i\right]' C_{NT1} \left[I_{q_{02}} - D_{NT1}(D_{NT1}' D_{NT1})^{-1} D_{NT1}'\right] \\ & \times C_{NT1}' \left[(NT_0)^{-1/2} \sum_{i=1}^{N} W_{li1}^{(0)\prime} \Delta u_i\right], \end{aligned}$$

with $C_{NT1} C_{NT1}' = \left(\frac{1}{NT_0} \sum_{i=1}^{N} W_{li1}^{(0)\prime} \Delta \breve{u}_{i1}^{(0)} \Delta \breve{u}_{i1}^{(0)\prime} W_{li1}^{(0)}\right)^{-1}$ and $D_{NT1} = C_{NT1}' \frac{W_1^{(0)\prime} Z}{NT_0}$; where $\Delta \breve{u}_{i1}^{(0)}$ is the one-step GMM residual using valid instrument matrix $W_{li1}^{(0)}$, and $W_1^{(0)} = \left(W_{l11}^{(0)\prime}, \ldots, W_{lN1}^{(0)\prime}\right)'$. Similar to the proof of Theorem 3, we also have $S_{(0)I} \xrightarrow{d} \varepsilon' M_I \varepsilon$, where $M_I = I_{q_{02}} - D_{NT1}(D_{NT1}' D_{NT1})^{-1} D_{NT1}$ and $rank(M_I) = q_{02} - k$. Let D_{NT1}^* contain the top q_{02} rows of D_{NT}. Notice that $D_{NT1}^* - D_{NT1} \to 0$; therefore,

$$DS_{(0)} = S_{(0)} - S_{(0)I} \xrightarrow{d} \varepsilon' M \varepsilon - \varepsilon' \begin{pmatrix} M_I & 0 \\ 0 & 0 \end{pmatrix} \varepsilon,$$

Finally, since

$$\left[M - \begin{pmatrix} M_I & 0 \\ 0 & 0 \end{pmatrix}\right]$$

is symmetric and idempotent with rank $q_0 - q_{02}$, and also

$$\left[M - \begin{pmatrix} M_I & 0 \\ 0 & 0 \end{pmatrix} \right] \begin{pmatrix} M_I & 0 \\ 0 & 0 \end{pmatrix} = 0$$

we have $DS_{(0)} \xrightarrow{d} \chi^2_{(q_0 - q_{02})}$.

VITA

NAME OF AUTHOR: Bin Peng

PLACE OF BIRTH: Hengshan, China

DATE OF BIRTH: May 28th, 1984

GRADUATE AND UNDERGRADUATE SCHOOLS ATTEND:

Huazhong University of Sci. and Tech., Wuhan, China

Syracuse University, Syracuse, U.S.

DEGREES AWARDED:

Master of Arts in Economics, 2010, Huazhong University of Sci. and Tech.

Bachelor of Arts in Economics, 2007, Huazhong University of Sci. and Tech.

AWARDS AND HONORS:

Graduate Assistantship, Syracuse University, 2011-2016

Maxwell Dean Summer Fellowship, Maxwell School, 2011-2015

Travel Grant, Syracuse University, 2014

Travel Grant, National Science Foundation, 2014

PROFESSIONAL EXPERIENCE:

Teaching Assistant, Department of Economics, Syracuse University, 2011-2016