Determination of Different Types of Fixed Effects in Three-Dimensional Panels*

Xun Lu^a, Ke Miao^b and Liangjun Su^b

^{a}Department of Economics, The Chinese University of Hong Kong

^{b}School of Economics, Singapore Management University, Singapore

July 31, 2019

Abstract

In this paper, we propose a jackknife method to determine the type of fixed effects in three-dimensional panel data models. We show that with probability approaching 1, the method can select the correct type of fixed effects in the presence of only weak serial or cross-sectional dependence among the error terms. In the presence of strong serial correlation, we propose a modified jackknife method and justify its selection consistency. Monte Carlo simulations demonstrate the excellent finite sample performance of our method. Applications to two datasets in macroeconomics and international trade reveal the usefulness of our method.

Key words: Consistency; Cross-validation; Fixed effect; Individual effect; Jackknife; Three-dimensional panel.

JEL Classification: C23, C33, C51, C52, F17, F47.

1 Introduction

Standard two-dimensional (2-D) fixed effects panel data models (see, e.g., Baltagi (2013) and Hsiao (2014)) have the advantage of modeling heterogeneity by introducing time effect (λ_t) and individual effect (α_i). In recent years, three-dimensional (3-D) panel data models are employed to study various phenomena in many economic fields, such as international trade, transportation, labor, housing, and migration (see, e.g., Mátyás

---

(2017) for a recent review). In the trade literature, the 2-D panel model was firstly extended to the 3-D framework by Mátyás (1997) who considered Model 2 in (1.1) below. Thereafter, other 3-D panel data models were proposed in the trade literature. Egger and Pfaffermayr (2003) proposed a panel gravity model taking into account the bilateral interaction effect by including the bilateral specific effect γ_{ij}. Baltagi and Egger (2003), Cheng and Wall (2005), Baldwin and Taglioni (2006) and Baier and Bergstrand (2007) also proposed several variations of the 3-D fixed effects panel data models. Balazsi, Mátyás, and Wansbeek (2018) showed that the least squares dummy variable (LSDV) method can be applied to estimate the coefficient consistently and illustrated that these results can be generalized for higher dimensional panel data models.

In the 2-D fixed effects panel data models, there are only four types of specifications of fixed effects in the absence of interactive fixed effects. However, in the 3-D models, the number of possible specifications of fixed effects can be as large as sixty-four (2^6) theoretically. Therefore, it is a highly empirically relevant question to determine which model to use in practice. The goal of this paper is to provide a practical method to select the correct specification of fixed effects in the 3-D panel data models. Specifically, we consider seven commonly used candidate models as suggested by Balazsi, Mátyás, and Wansbeek (2017, 2018):

$$
\begin{aligned}
\text{Model 1}: & \quad y_{ijt} = x'_{ijt}\beta + u_{ijt}, \\
\text{Model 2}: & \quad y_{ijt} = x'_{ijt}\beta + \alpha_i + \gamma_j + \lambda_t + u_{ijt}, \\
\text{Model 3}: & \quad y_{ijt} = x'_{ijt}\beta + \gamma_{ij} + u_{ijt}, \\
\text{Model 4}: & \quad y_{ijt} = x'_{ijt}\beta + \gamma_{ij} + \lambda_t + u_{ijt}, \\
\text{Model 5}: & \quad y_{ijt} = x'_{ijt}\beta + \alpha^*_{jt} + u_{ijt}, \\
\text{Model 6}: & \quad y_{ijt} = x'_{ijt}\beta + \alpha_{it} + \alpha^*_{jt} + u_{ijt}, \\
\text{Model 7}: & \quad y_{ijt} = x'_{ijt}\beta + \gamma_{ij} + \alpha_{it} + \alpha^*_{jt} + u_{ijt},
\end{aligned}
\tag{1.1}
$$

for $i = 1, ..., N$, $j = 1, ..., M$, and $t = 1, ..., T$, where y_{ijt} is the dependent variable, e.g., the volumes of trades (exports) from country i to country j in year t, x_{ijt} is a $k \times 1$ vector of regressors that contains a constant term and may also include the lagged dependent variables, u_{ijt} is the idiosyncratic error term, and α_i, γ_j, λ_t, γ_{ij}, α^*_{jt}, and α_{it} are fixed effects that are treated as fixed parameters to be estimated.

In practice, there are two main motivations for model selection. First, an economic theory may suggest certain types of models and it would be interesting to know which model is likely to be true empirically. In our context, different specifications of fixed effects may be interpreted differently and it would be useful to understand the types of interactions of the unobserved heterogeneities. For example, consider the gravity model in international trade where y_{ijt} is the trade volume (exports) from country i to country j in year t. Country fixed effects have been argued to be important for the gravity models (see, e.g., Feenstra (2016, p.143)), as they represent unobservable multilateral resistance levels termed by Anderson and van Wincoop (2003). Therefore, if the multilateral resistance levels are time-varying, represented by α_{it} and α^*_{jt} here, the trade theory would support Model 6 and Model 7. Our method can select the correct model consistently and thus can be used to confirm or reject the theory. Taking another example, let y_{ijt} be the wage for worker-type i employed by firm j at time t. In an assortative matching model, Shimer and Smith (2000) argue that there might be complementarities between firms' productivity and workers' ability. Given that firms' productivity and workers' ability are typically unobservable to econometricians, their theory would suggest that the interaction term γ_{ij} is important and Model 3, 4 and 7 would be appropriate.

Second, model selection is important for the estimation and inference for the parameter of interest (typically β here). If we apply a misspecified model that is smaller than the true model, we may suffer from the notorious omitted variable bias (OVB) issue. If we adopt a larger model that nests the true model, we may have substantial efficiency loss as we have included many redundant dummy variables generated by the fixed effects. When N, M, and T are all large, the number of redundant dummy variables can be huge and thus tends to result in enormous efficiency loss. For this reason, it is not always desirable to adopt the largest model (Model 7) in empirical studies. To illustrate this point, we conduct a simple simulation exercise where the true data generating process (DGP) is

$$y_{ijt} = \beta_0 + \beta_1 y_{ij,t-1} + u_{ijt}, \tag{1.2}$$

$(\beta_0, \beta_1) = (1, 0.75)$ and u_{ijt}'s are IID $N(0, 1)$ random variables. Hence, here Model 1 is the true model. Table 1 compares the mean squared errors (MSEs) of the estimates of β_1 based on Models 1-7. Given this is a dynamic model, we consider both non-bias corrected estimators and bias-corrected estimators where the bias correction is based on the half-panel jackknife method as proposed in Dhaene and Jochmans (2015) or the analytic formula derived in online Appendix D. For both types of estimators, the estimators based on the true model (Model 1) achieve the smallest MSEs as expected. Adopting larger model results in substantial efficiency loss. For example, when $(N, M, T) = (10, 10, 10)$, the MSE of the non-biased corrected estimator based on Model 7 is 100 times as large as that based on Model 1. The bias-corrected estimator based on the analytic formula works but not as well as the jackknife one. For the jackknife bias-corrected estimator, the MSE based on Model 7 is about seven times as large as that based on Model 1. Interestingly, we find that the estimates based on Models 3, 4 and 7 perform similarly in finite samples. Following the lead of Balazsi et al. (2018), we can study the asymptotic Nickell bias of the least squares dummy variable (LSDV) estimator $\hat{\beta}_1^{(m)}$ of β_1 in (1.2) based on Model m. Table D2 in the online supplement reports the Nickell biases. It suggests that when N, M, and T pass to infinity jointly at the same rate as we have here in Table 1, the asymptotic biases of $\hat{\beta}_1^{(m)}$, $m = 3, 4, 7$, share the same dominant term $\frac{-(1+\beta_1)}{T}$, whereas the asymptotic biases of the other four estimators (i.e., $\hat{\beta}_1^{(1)}, \hat{\beta}_1^{(2)}, \hat{\beta}_1^{(5)}, \hat{\beta}_1^{(6)}$) are all $o(\frac{1}{T})$. This observation, in conjunction with the fact that all seven estimators share the same asymptotic variance when Model 1 is true and given by (1.2), explains why the performance of the estimators based on Models 3, 4 and 7 are similar in Table 1 despite the fact that Model 7 contains far more parameters than Models 3 and 4.[1]

Given the existence of many flexible ways of including fixed effects in the 3-D panel data models, the specification problem is more severe and complicated than the 2-D framework. To the best of our knowledge, so far there exists no systematic way of determining fixed effects specifications in the 3-D panel models in the literature. In the traditional 2-D models, Wu and Li (2014) proposed two Hausman-type tests for individual and time effects in a two-way error component model. Their method involves multiple hypothesis tests and suffers from severe size distortion in the 3-D case because the number of hypothesis tests increases exponentially as the number of models increases. Most recently Lu and Su (2019a) proposed a jackknife methodology to determine the inclusion of individual effects, time effects, or both through the leave-one-out

---

[1] If we further decompose MSE into bias and variance terms in the simulations, we can also find that for M3, M4 and M7, both the bias and variance terms are important no matter whether bias-correction is corrected. For M1, M2, M5 and M6, the bias terms are relatively small, and therefore, the variance terms play a dominant role regardless of whether one corrects the bias or not.

Table 1: Comparisons of MSEs of β_1 (true model: Model 1)

		Adopted Models						
		M1	M2	M3	M4	M5	M6	M7
	N=10, M=10, T=5	**8.76**	19.42	1720.48	1720.35	9.62	10.22	1726.75
Non-bias	N=20, M=20, T=5	**2.25**	4.73	1700.29	1700.29	2.41	2.56	1702.56
correction	N=10, M=10, T=10	**4.28**	10.38	442.65	443.36	4.67	5.16	446.09
	N=10, M=10, T=20	**2.12**	4.23	101.52	101.52	2.34	2.75	102.58
	N=20, M=20, T=20	**0.60**	1.14	98.77	98.76	0.64	0.67	98.87
Bias	N=10, M=10, T=5	**9.20**	21.21	118.65	119.11	10.14	10.63	138.52
correction	N=20, M=20, T=5	**2.34**	5.23	59.26	59.25	2.50	2.66	61.42
based on	N=10, M=10, T=10	**4.28**	7.53	22.12	22.72	4.69	5.16	28.95
jackknife	N=10, M=10, T=20	**2.11**	2.58	7.77	7.91	2.34	2.74	9.47
	N=20, M=20, T=20	**0.60**	0.72	2.65	2.66	0.64	0.67	2.88
Bias	N=10, M=10, T=5	**8.72**	28.80	97.64	98.11	9.62	10.22	107.43
correction	N=20, M=20, T=5	**2.24**	6.99	69.12	69.13	2.41	2.56	70.74
based on	N=10, M=10, T=10	**4.23**	6.50	23.91	24.20	4.67	5.16	27.02
the analytic	N=10, M=10, T=20	**2.12**	2.51	5.21	5.27	2.34	2.75	6.19
formula	N=20, M=20, T=20	**0.60**	0.66	2.43	2.44	0.64	0.67	2.57

Note: Numbers in the main entries are MSEs$\times 10^4$ of the estimates of β_1. The number of replications is 1000.

cross-validation (CV) in the 2-D framework. For a detailed review of the specification of fixed effects in the 2-D models, see Lu and Su (2019a).

Jackknife or CV has been applied to conduct model selection in many different contexts, though often without rigorous justification.[2] In the panel context, although Lu and Su (2019a) showed that the jackknife method could consistently select the correct model in 2-D panels, it was unclear whether jackknife would work for 3-D panels. There are substantial differences between the 2-D and 3-D cases. First, there are a large number of candidate models in 3-D panels that require different asymptotic analyses. The fixed effect specifications are much more complicated in the 3-D case than those in the 2-D case. For example, in Model 7 above, to control for the fixed effects, we need to include $(NM + NT + MT - N - M - T)$ dummy variables. We focus on the seven models in (1.1) that are commonly used in practice but conjecture that our method remains valid for a larger subset of candidate models. Because we allow each of these seven models to be either true or misspecified, there are 49 scenarios under our investigation. To prove the selection consistency, we need to carefully compare the correctly specified models and misspecified models under these 49 scenarios. Second, to expedite the asymptotic analysis, we allow N, M, and T to pass to infinity jointly and the asymptotic analysis along the three dimensions is quite challenging. We have to pay particular attention to the interactions of the three dimensions in our proofs, as we do not impose any conditions on the relative rates at which N, M, and T pass to infinity. Therefore, it is much more challenging to show the selection consistency in the 3-D case.

Despite the involved theoretical proofs, the new methodology is easy to implement and has excellent

---

[2] The jackknife method was originally proposed by Quenouille (1956) and Tukey (1958). It can be used for different purposes, such as bias-correction, inference and model selection. The theoretical work on jackknife for model specification includes Allen (1974), Stone (1974), Geisser (1974), Wahba and Wold (1975), Li (1987), Efron (1983, 1986), Picard and Cook (1984), Andrews (1991), Shao (1993), Hansen and Racine (2012), and Lu and Su (2015), among others.

performance in Monte Carlo simulations. In particular, it can easily handle unbalanced panels, which is a common phenomenon in multi-dimensional panel data. Asymptotically, we prove that this method can determine the correct model with probability approaching one as all the three dimensions go to infinity. As well, we argue that this method can be extended to higher-dimensional fixed effects panel data models. Although here we focus on seven popular candidate models in our asymptotic theory, we expect that our methodology can be applied to the other 3-D models or even 3-D nonlinear panels.

It is worth mentioning that here we focus on the selection consistency of our approach and leave the post-selection inference issue untouched. For the post-selection estimation and inference for the parameter of interest (β), it is desirable to consider uniform inference, which remains a challenging question in the model selection literature and certainly goes beyond the scope of this paper.

We provide two empirical applications to illustrate the usefulness of our new method. In the first application, we apply our method to the dataset used in Samaniego and Sun (2015), where they adopt Model 7 to investigate which technological characteristics lead industries to experience the most difficulty during the recession period. The dependent variable is the growth of industry j in country i at time t and the key independent variable is the interaction term between the recession indicator and industry technological characteristics. Our method finds that Model 6 is an appropriate model and country-industry fixed effects are redundant. In the second application, we apply our method to gravity equations in international trade. The dependent variable is the logarithm of the export of country i to country j in year t, and the independent variables include the logarithm of the sum of country i's GDP and country j's GDP in year t and the logarithm of the sum of country i's population and country j's population in year t. We show that the largest model (Model 7) is an appropriate model for gravity equations.

The rest of the paper is structured as follows. In Section 2, we discuss the 3-D panels with different types of fixed effects and introduce the notation to put all these models in a unified framework. We propose the jackknife method to determine the types of fixed effects in the 3-D panels and study its asymptotic properties in Section 3. We propose a modified jackknife method to incorporate strong serial dependence and study its consistency in Section 4. Section 5 reports Monte Carlo simulation results and compares our new methods with information criterion (IC)-based methods for both static and dynamic panel DGPs. In Section 6, we apply our method to two datasets to study (i) the interaction between technology and business cycles and (ii) the gravity models in international trade. Section 7 concludes. The proof of the main result (Theorem 3.1) is relegated to Appendix A. The proof of Theorem 4.1, the proofs of the technical lemmas and the derivation of the Nickell biases for the panel AR(1) models are relegated to the online Appendices B, C, and D, respectively.

Notation. For an $m \times n$ real matrix A, we denote its transpose as A', its Frobenius norm as $\|A\|$ and its spectral norm as $\|A\|_{\mathrm{sp}}$. Let $P_A \equiv A (A'A)^{-1} A'$ and $M_A \equiv I_m - P_A$, where I_m denotes an $m \times m$ identity matrix. When $A = \{a_{ij}\}$ is symmetric, we use $\lambda_{\max}(A)$ and $\lambda_{\min}(A)$ to denote its maximum and minimum eigenvalues, respectively. Let $J_m = \iota_m \iota'_m$ where ι_m denotes an $m \times 1$ vector of ones. Let $\otimes$ denote Kronecker product and $\xrightarrow{p}$ convergence in probability. We use $(N, M, T) \to \infty$ to denote that N, M, and T pass to infinity jointly.

2 Models and Notation

We consider a 3-D panel where the dependent and independent variables are given by y_{ijt} and x_{ijt}, respectively, for $i = 1, ..., N$, $j = 1, ..., M_i$, and $t = 1, ..., T_{ij}$. For notational simplicity, we will assume that $M_i = M$ for each i and $T_{ij} = T$ for each pair (i, j) and remark that our asymptotic theory continues to hold for the general case of unbalanced panels but with more complicated notation. As Balazsi, Mátyás, and Wansbeek (2017, 2018) remark, there are 2^6 ways to formulate the fixed effects in a 3-D panel, but only a small subset of these are considered and applied in empirical applications. Following these authors, we only consider the selection of the seven models as in (1.1) that are frequently employed.

Model 1 is a pooled regression model that ignores unobserved heterogeneity. Model 2 allows the specific effects to enter the model additively. Model 3 only allows a pairwise interaction between the i- and j-specific fixed effects but the model can be studied as if one studies the usual 2-D model with individual fixed effects by treating observation along the (i, j) dimensions as one single dimension. Similarly, we can study Model 4 as if we study the usual 2-D model with two-way error components by treating (i, j) dimensions as a single dimension. Model 5 allows the interaction between the j- and t-specific effects while Model 6 allows two pairwise interactions of specific effects. Model 7 encompasses all three pairwise effects and nests Models 1-6 as special cases.

Model 2 has been frequently adopted in empirical research; see Mátyás (1997), Goldhaber, Brewer, and Anderson (1999), Egger (2000), Davis (2002), Egger and Pfaffermayr (2003), among others. Mátyás (1997) applies Model 2 to estimate a gravity equation where the dependent variable is the logarithm of the trade (exports) from country i to country j at time t. Egger (2000) considers the Hausman test for random effects versus fixed effects in Model 2 for the gravity equation considered by Mátyás (1997) and provides arguments for the superiority of a fixed effects specification. Goldhaber, Brewer, and Anderson (1999) apply Model 2 with random effects to determine how much of the achievement on a 10th-grade standardized test can be explained by observable schooling resources and unobservable school, teacher, and class effects. Davis (2002) considers both fixed effects and random effects estimation of Model 2 using data from a retail market where the three dimensions of data variation are products sold in various locations over time.

Egger and Pfaffermayr (2003, EP) extend Model 2 to include the exporter-by-importer (bilateral) interaction effects γ_{ij} and the time effect as in Model 4. EP find evidence that suggests that Model 4 is preferred to the three-way error component specification in Model 2. Cheng and Wall (2005) estimate the gravity equation for bilateral trade flows by using Model 4 with fixed effects and compare with the results from using Model 2. They also find Model 4 is preferred to Model 2. Baltagi, Egger, and Pfaffermayr (2003, BEP) consider fixed effects estimation of various models for bilateral trade data, including Models 5, 6, and 7. See also Baldwin and Taglioni (2006). Baier and Bergstrand (2007) estimate the panel gravity equations with bilateral fixed or/and country-and-time effects (Models 3, 7) and they consider both within transformation and first-differencing. Berthélemy and Tichit (2004) estimate a censored version of Model 5 with random effects where the dependent variable is the aid the ith recipient receives from the jth donor at time t. Samaniego and Sun (2015) apply Model 7 with fixed effects to study the growth of industry j in country i at time t.

With seven models, there are 7×7 cases of model fitting to be considered. In Table 2, we summarize all the cases for model fitting. In each row, the fitting case for one true model is presented. For example, when

Model 2 is the true model, Models 1, 3, and 5 are under-fitted and Models 4, 6, and 7 are over-fitted. In the next section, we propose a method to select the just-fitted model. In the theoretical analysis, we need to discuss over-fitted and under-fitted cases separately.

Table 2: Cases for model fitting

True model	Adopted model						
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Model 1	0	+	+	+	+	+	+
Model 2	−	0	−	+	−	+	+
Model 3	−	−	0	+	−	−	+
Model 4	−	−	−	0	−	−	+
Model 5	−	−	−	−	0	+	+
Model 6	−	−	−	−	−	0	+
Model 7	−	−	−	−	−	−	0

Note: "+", "0", and "−" represent over-fitted, just-fitted and under-fitted cases, respectively.

Note that the fixed effects parameters are not separately identified without restrictions. To unify the theory and simplify our asymptotic analysis, we impose the following identification restrictions in Models 2-7:

Model 2 : $\sum_{i=1}^{N} \alpha_i = 0, \ \sum_{j=1}^{M} \gamma_j = 0, \ \sum_{t=1}^{T} \lambda_t = 0,$

Model 3 : $\sum_{i=1}^{N} \sum_{j=1}^{M} \gamma_{ij} = 0,$

Model 4 : $\sum_{i=1}^{N} \sum_{j=1}^{M} \gamma_{ij} = 0, \ \sum_{t=1}^{T} \lambda_t = 0,$

Model 5 : $\sum_{j=1}^{M} \sum_{t=1}^{T} \alpha_{jt}^* = 0,$

Model 6 : $\sum_{i=1}^{N} \alpha_{it} = 0$ for each t, $\sum_{j=1}^{M} \sum_{t=1}^{T} \alpha_{jt}^* = 0,$

Model 7 : $\sum_{j=1}^{M} \gamma_{ij} = 0$ for each i, $\sum_{t=1}^{T} \alpha_{jt}^* = 0$ for each j, $\sum_{i=1}^{N} \alpha_{it} = 0$ for each t.

That is, there are 3 restrictions in Model 2, 1 restriction in Model 3, 2 restrictions in Model 4, 1 restriction in Model 5, $T+1$ restrictions in Model 6, and $N+M+T$ restrictions in Model 7.

We stack the observations in a way such that index i goes the slowest, then j, and finally t the fastest; e.g., $Y = (y_{111}, ..., y_{11T}, ..., y_{1M1}, ..., y_{1MT}, ..., y_{N11}, ..., y_{N1T}, ..., y_{NM1}, ..., y_{NMT})'$. Define X and U analogously. Then we can write Models 1-7 in a uniform way as

$$Y = X\beta + D_m \pi_m + U = Z_m \theta_m + U,$$

where $Z_m = (X, D_m)$ and $\theta_m = (\beta', \pi_m')'$. Here D_m's are the dummy matrices that incorporate the above identification restrictions:

$D_1 : \ \varnothing$

$D_2 : \ (D_I, D_J, D_T)$

$D_3 : \ D_{IJ}$

$D_4 : \ (D_{IJ}, D_T)$

$D_5 : \ D_{JT}$

$D_6 : \ (D_{IT}^*, D_{JT})$

$D_7 : \ (D_{IJ}^*, D_{IT}^*, D_{JT}^*)$

where

$$
D_I = \begin{bmatrix} I_{N-1} \\ -\iota'_{N-1} \end{bmatrix} \otimes \iota_M \otimes \iota_T, \ D_J = \iota_N \otimes \begin{bmatrix} I_{M-1} \\ -\iota'_{M-1} \end{bmatrix} \otimes \iota_T, \ D_T = \iota_N \otimes \iota_M \otimes \begin{bmatrix} I_{T-1} \\ -\iota'_{T-1} \end{bmatrix},
$$

$$
D_{IJ} = \begin{bmatrix} I_{NM-1} \\ -\iota'_{NM-1} \end{bmatrix} \otimes \iota_T, \ D_{JT} = \iota_N \otimes \begin{bmatrix} I_{MT-1} \\ -\iota'_{MT-1} \end{bmatrix}, \ \text{and}
$$

$$
D^*_{IT} = \begin{bmatrix} I_{N-1} \\ -\iota'_{N-1} \end{bmatrix} \otimes \iota_M \otimes I_T, \ D^*_{IJ} = I_N \otimes \begin{bmatrix} I_{M-1} \\ -\iota'_{M-1} \end{bmatrix} \otimes \iota_T, \ D^*_{JT} = \iota_N \otimes I_M \otimes \begin{bmatrix} I_{T-1} \\ -\iota'_{T-1} \end{bmatrix};
$$

and π_m's are the coefficients of the dummy variables in D_m's:

$$
\begin{aligned}
\pi_1 =& \ \varnothing, \\
\pi_2 =& \ (\alpha_1, ..., \alpha_{N-1}, \gamma_1, ..., \gamma_{M-1}, \lambda_1, ..., \lambda_{T-1})', \\
\pi_3 =& \ (\gamma_{1,1}, ..., \gamma_{1,M}, ..., \gamma_{N,1}, ..., \gamma_{N,M-1})', \\
\pi_4 =& \ (\gamma_{1,1}, ..., \gamma_{1,M}, ..., \gamma_{N,1}, ..., \gamma_{N,M-1}, \lambda_1, ..., \lambda_{T-1})', \\
\pi_5 =& \ (\alpha_{1,1}, ..., \alpha_{1,T}, ..., \alpha_{M-1,1}, ..., \alpha_{M-1,T}, \alpha_{M,1}, ..., \alpha_{M,T-1})', \\
\pi_6 =& \ (\alpha_{1,1}, ..., \alpha_{1,T}, ..., \alpha_{N-1,1}, ..., \alpha_{N-1,T}, \alpha^*_{1,1}, ..., \alpha^*_{1,T}, \ ..., \alpha^*_{M-1,1}, ..., \alpha^*_{M-1,T}, \ \alpha^*_{M,1}, ..., \alpha^*_{M,T-1})', \\
\pi_7 =& \ (\gamma_{1,1}, ..., \gamma_{1,M-1}, ..., \gamma_{N,1}, ..., \gamma_{N,M-1}, \alpha_{1,1}, ..., \alpha_{1,T}, ..., \alpha_{N-1,1}, ..., \alpha_{N-1,T}, \alpha^*_{1,1}, ..., \alpha^*_{1,T-1}, ..., \\
& \ \alpha^*_{M,1}, ..., \alpha^*_{M,T-1})'.
\end{aligned}
$$

Let $d'_{ijt,m}$ and $z'_{ijt,m} = (x'_{ijt}, d'_{ijt,m})$ denote typical rows of D_m and Z_m, respectively, for $m = 2, ..., 7$. Let $Z_1 = X$ and $z_{ijt,1} = x_{ijt}$. It is easy to verify that

$$
D_I \perp D_J \perp D_T, \ D_{IJ} \perp D_T, \ D^*_{IT} \perp D_{JT}, \ \text{and} \ D^*_{IT} \perp D^*_{IJ} \perp D^*_{JT},
$$

where $A \perp B$ means that A and B are orthogonal ($A'B = 0$ and $B'A = 0$) and $A \perp B \perp C$ means A, B, and C are mutually orthogonal to each other. With such an orthogonal property, it is easy to calculate the inverses of $D'_m D_m$ and $Z'_m Z_m$ for $m = 2, ..., 7$.

Throughout the paper, we will calculate various sample means. Define

$$
\overline{u}_{i..} = \frac{1}{MT} \sum_{j=1}^{M} \sum_{t=1}^{T} u_{ijt}, \ \overline{u}_{.j.} = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} u_{ijt}, \ \overline{u}_{..t} = \frac{1}{NM} \sum_{i=1}^{N} \sum_{j=1}^{M} u_{ijt},
$$

$$
\overline{u}_{ij.} = \frac{1}{T} \sum_{t=1}^{T} u_{ijt}, \ \overline{u}_{i.t} = \frac{1}{M} \sum_{j=1}^{M} u_{ijt}, \ \overline{u}_{.jt} = \frac{1}{N} \sum_{i=1}^{N} u_{ijt}, \ \text{and} \ \overline{u} = \frac{1}{NMT} \sum_{i=1}^{N} \sum_{j=1}^{M} \sum_{t=1}^{T} u_{ijt}.
$$

Let $\overline{x}_{i..}, \overline{x}_{.j.}, \overline{x}_{..t}, \overline{x}_{ij.}, \overline{x}_{i.t}, \overline{x}_{.jt}$ and $\overline{x}$ be defined analogously.

3 Methodology and Asymptotic Theory

In this section, we first introduce the jackknife method to determine the different types of fixed effects in 3-D panels. Then we introduce the assumptions that are needed for our asymptotic analysis and report the consistency of the jackknife method.

3.1 The jackknife method

The OLS estimator of $\theta_m = \left(\beta', \pi'_m\right)'$ in Model m based on all NMT observations is given by

$$\hat{\theta}_m = \left(\hat{\beta}'_m, \hat{\pi}'_m\right)' = (Z'_m Z_m)^{-1} Z'_m Y \text{ for } m = 1, 2, ..., 7.$$

We will also consider the leave-one-out estimator of θ_m with the (i, j, t)th observation deleted from the sample:

$$\hat{\theta}_{ijt,m} = \left(\hat{\beta}'_{ijt,m}, \hat{\pi}'_{ijt,m}\right)' = \left(Z'_m Z_m - z_{ijt,m} z'_{ijt,m}\right)^{-1} \left(Z'_m Y - z_{ijt,m} y_{ijt}\right) \text{ for } m = 1, 2, ..., 7,$$

where $i = 1, ..., N$, $j = 1, ..., M$, and $t = 1, ..., T$. The out-of-sample predicted value for y_{ijt} is defined as $\hat{y}_{ijt,m} = z'_{ijt,m}\hat{\theta}_{ijt,m} = x'_{ijt}\hat{\beta}_{ijt,m} + d'_{ijt,m}\hat{\pi}_{ijt,m}$. Our jackknife method is based on the following leave-one-out cross-validation (CV) function:

$$CV\left(m\right) = \frac{1}{NMT} \sum_{i=1}^{N} \sum_{j=1}^{M} \sum_{t=1}^{T} (y_{ijt} - \hat{y}_{ijt,m})^2 \text{ for } m = 1, 2, ..., 7.$$

We propose to choose a model such that $CV\left(m\right)$ is minimized. Define

$$\hat{m} = \operatorname*{argmin}_{1 \leq m \leq 7} CV\left(m\right).$$

We will show that under some regularity conditions, $\hat{m}$ is given by the true model with probability approaching 1 (w.p.a.1) when we assume that Models 1-7 contain the true model.

Remark 1. For certain dynamic panel models (such as Models 3, 4 and 7), bias correction can be needed for the inference purpose contingent upon the rates at which N, M, and T pass to infinity. Nevertheless, our purpose here is to determine the type of fixed effects. We show that our method can consistently select the true model without the need for bias correction. Given the selected model, one can consider bias correction as needed to make inferences.

Remark 2. In this paper, we focus on how to conduct model selection and do not consider the post-selection inference issue. In practice, one can conduct pointwise inferences based on the selected model. But it is well known that inferences based on the selected model may not be uniformly valid; see, e.g., Leeb and Pötscher (2005). It is challenging to conduct post-selection uniformly valid inferences. We leave this important issue for future research.

3.2 Asymptotic theory under weak serial and cross-sectional dependence

Let $\hat{Q}_m = \frac{1}{NMT} X' M_{D_m} X$ for $m = 2, ..., 7$ and $\hat{Q}_1 = \frac{1}{NMT} X' X$. Let $\max_{i,j,t} = \max_{1 \leqslant i \leqslant N, 1 \leqslant j \leqslant M, 1 \leqslant t \leqslant T}$, and $\sum_{i,j,t} = \sum_{i=1}^{N} \sum_{j=1}^{M} \sum_{t=1}^{T}$. Similarly, let $\sum_{i,j}, \sum_{i,t}, \sum_{j,t}, \sum_i, \sum_j$, and $\sum_t$ abbreviate $\sum_{i=1}^{N} \sum_{j=1}^{M}$, $\sum_{i=1}^{N} \sum_{t=1}^{T}, \sum_{j=1}^{M} \sum_{t=1}^{T}, \sum_{i=1}^{N}, \sum_{j=1}^{M}$, and $\sum_{t=1}^{T}$, respectively. To report the asymptotic property of the jackknife method, we introduce some assumptions.

Assumption A.1 (i) $E(u_{ijt}) = 0$, $\max_{i,j,t} E(u_{ijt}^2) < C$ for some positive constant, and $\frac{1}{NMT} \sum_{i,j,t} u_{ijt}^2 \xrightarrow{p} \overline{\sigma}_u^2 > 0$;

(ii) $\max_{i,j,t} \|x_{ijt}\| = O_p((NMT)^{1/4})$;

(iii) $\frac{1}{NMT}\sum_{i,j,t}u_{ijt}^8=O_p\left(1\right)$ and $\frac{1}{NMT}\sum_{i,j,t}\|x_{ijt}\|^8=O_p\left(1\right)$;

(iv) $\overline{u}=O_p((NMT)^{-1/2})$, and $\frac{1}{NMT}X'U=O_p((NMT)^{-1/2})$;

(v) There exist positive constants $\underline{c}_Q$ and $\overline{c}_Q$ such that $P(\underline{c}_Q\le\lambda_{\min}(\hat{Q}_m)\le\lambda_{\max}(\hat{Q}_m)\le\overline{c}_Q)\to 1$ for $m=1,...,7$.

Assumption A.2. There are finite positive constants $\overline{\sigma}_{u\ell}^2$, $\ell=1,2,...,6$, such that

(i) $\frac{MT}{N}\sum_i\overline{u}_{i\cdot\cdot}^2\overset{p}{\to}\overline{\sigma}_{u1}^2$;

(ii) $\frac{NT}{M}\sum_j\overline{u}_{\cdot j\cdot}^2\overset{p}{\to}\overline{\sigma}_{u2}^2$;

(iii) $\frac{NM}{T}\sum_t\overline{u}_{\cdot\cdot t}^2\overset{p}{\to}\overline{\sigma}_{u3}^2$;

(iv) $\frac{T}{NM}\sum_{i,j}\overline{u}_{ij\cdot}^2\overset{p}{\to}\overline{\sigma}_{u4}^2$;

(v) $\frac{M}{NT}\sum_{i,t}\overline{u}_{i\cdot t}^2\overset{p}{\to}\overline{\sigma}_{u5}^2$;

(vi) $\frac{N}{MT}\sum_{j,t}\overline{u}_{\cdot jt}^2\overset{p}{\to}\overline{\sigma}_{u6}^2$.

Assumption A.3 (i)$\frac{1}{N}\sum_i(\overline{x}_{i\cdot\cdot}-\overline{x})(\overline{x}_{i\cdot\cdot}-\overline{x})'=O_p((MT)^{-1})$;

(ii)$\frac{1}{M}\sum_j(\overline{x}_{\cdot j\cdot}-\overline{x})(\overline{x}_{\cdot j\cdot}-\overline{x})'=O_p((NT)^{-1})$;

(iii)$\frac{1}{T}\sum_t(\overline{x}_{\cdot\cdot t}-\overline{x})(\overline{x}_{\cdot\cdot t}-\overline{x})'=O_p((NM)^{-1})$;

(iv)$\frac{1}{NM}\sum_{i,j}(\overline{x}_{ij\cdot}-\overline{x})(\overline{x}_{ij\cdot}-\overline{x})'=O_p(T^{-1})$;

(v)$\frac{1}{NT}\sum_{i,t}(\overline{x}_{i\cdot t}-\overline{x})(\overline{x}_{i\cdot t}-\overline{x})'=O_p(M^{-1})$;

(vi)$\frac{1}{MT}\sum_{j,t}(\overline{x}_{\cdot jt}-\overline{x})(\overline{x}_{\cdot jt}-\overline{x})'=O_p(N^{-1})$.

Assumption A.4 (i) If Model 2 is the true model, there exist positive constants $\varphi_{2,m}$ for $m=1,3,5$ such that $\frac{1}{NMT}\sum_{i,j,t}\left(\varsigma_{2ijt,m}\right)^2\overset{p}{\to}\varphi_{2,m}$ and $\frac{1}{NMT}\sum_{i,j,t}\varsigma_{2ijt,m}u_{ijt}\overset{p}{\to}0$, where $\varsigma_{2ijt,m}=\alpha_i+\gamma_j+\lambda_t-z_{ijt,m}'(Z_m'Z_m)^{-1}Z_m'D_2\pi_2$;

(ii) If Model 3 is the true model, there exist positive constants $\varphi_{3,m}$ for $m=1,2,5,6$ such that $\frac{1}{NMT}\sum_{i,j,t}\left(\varsigma_{3ijt,m}\right)^2\overset{p}{\to}\varphi_{3,m}$ and $\frac{1}{NMT}\sum_{i,j,t}\varsigma_{3ijt,m}u_{ijt}\overset{p}{\to}0$, where $\varsigma_{3ijt,m}=\gamma_{ij}-z_{ijt,m}'(Z_m'Z_m)^{-1}Z_m'D_3\pi_3$;

(iii) If Model 4 is the true model, there exist positive constants $\varphi_{4,m}$ for $m=1,2,3,5,6$ such that $\frac{1}{NMT}\sum_{i,j,t}\left(\varsigma_{4ijt,m}\right)^2\overset{p}{\to}\varphi_{4,m}$ and $\frac{1}{NMT}\sum_{i,j,t}\varsigma_{4ijt,m}u_{ijt}\overset{p}{\to}0$, where $\varsigma_{4ijt,m}=\gamma_{ij}+\lambda_t-z_{ijt,m}'(Z_m'Z_m)^{-1}Z_m'D_4\pi_4$;

(iv) If Model 5 is the true model, there exist positive constants $\varphi_{5,m}$ for $m=1,2,3,4$ such that $\frac{1}{NMT}\sum_{i,j,t}\left(\varsigma_{5ijt,m}\right)^2\overset{p}{\to}\varphi_{5,m}$ and $\frac{1}{NMT}\sum_{i,j,t}\varsigma_{5ijt,m}u_{ijt}\overset{p}{\to}0$, where $\varsigma_{5ijt,m}=\alpha_{jt}^*-z_{ijt,m}'(Z_m'Z_m)^{-1}Z_m'\times D_5\pi_5$;

(v) If Model 6 is the true model, there exist positive constants $\varphi_{6,m}$ for $m=1,2,3,4,5$ such that $\frac{1}{NMT}\sum_{i,j,t}\left(\varsigma_{6ijt,m}\right)^2\overset{p}{\to}\varphi_{6,m}>0$ and $\frac{1}{NMT}\sum_{i,j,t}\varsigma_{6ijt,m}u_{ijt}\overset{p}{\to}0$, where $\varsigma_{6ijt,m}=\alpha_{it}+\alpha_{jt}^*-z_{ijt,m}'(Z_m'Z_m)^{-1}\times Z_m'D_6\pi_6$;

(vi) If Model 7 is the true model, there exist positive constants $\varphi_{7,m}$ for $m=1,2,3,4,5,6$ such that $\frac{1}{NMT}\sum_{i,j,t}\left(\varsigma_{7ijt,m}\right)^2\overset{p}{\to}\varphi_{7,m}$ and $\frac{1}{NMT}\sum_{i,j,t}\varsigma_{7ijt,m}u_{ijt}\overset{p}{\to}0$, where $\varsigma_{7ijt,m}=\gamma_{ij}+\alpha_{it}+\alpha_{jt}^*-z_{ijt,m}'(Z_m'Z_m)^{-1}\times Z_m'D_7\pi_7$.

Assumption A.1(i)-(ii) imposes weak conditions on $\{u_{ijt}\}$ and $\{x_{ijt}\}$, which can be verified under various primitive conditions. For example, a sufficient condition for A.1(ii) is that $\max_{i,j,t}E\|x_{ijt}\|^4\le C<\infty$. Assumption A.1(iii) is imposed to ease the proof of Lemmas A.13 and A.14 in Appendix A and can be relaxed at the cost of more lengthy arguments. Assumption A.1(iv) is weak and commonly assumed in panel data models in the absence of endogeneity. Note that we permit x_{ijt} to contain lagged dependent variables so that dynamic panel data models are allowed. Assumption A.1(v) specifies the usual identification conditions

for the fixed effects (FE) estimation of Models 1-7. Using Lemmas A.2-A.3 in Appendix A, we can readily verify that $\hat{Q}_m = \frac{1}{MMT} \sum_{i,j,t} \breve{x}_{ijt}^{(m)} \breve{x}_{ijt}^{(m)\prime}$ where

$$
\begin{aligned}
\breve{x}_{ijt}^{(2)} &= x_{ijt} - (\overline{x}_{i\cdot\cdot} - \overline{x}) - (\overline{x}_{\cdot j\cdot} - \overline{x}) - (\overline{x}_{\cdot\cdot t} - \overline{x}), \\
\breve{x}_{ijt}^{(3)} &= x_{ijt} - (\overline{x}_{ij\cdot} - \overline{x}), \\
\breve{x}_{ijt}^{(4)} &= x_{ijt} - (\overline{x}_{ij\cdot\cdot} - \overline{x}) - (\overline{x}_{\cdot\cdot t} - \overline{x}), \\
\breve{x}_{ijt}^{(5)} &= x_{ijt} - (\overline{x}_{\cdot jt} - \overline{x}), \\
\breve{x}_{ijt}^{(6)} &= x_{ijt} - (\overline{x}_{i\cdot t} - \overline{x}_{\cdot\cdot t}) - (\overline{x}_{\cdot jt} - \overline{x}), \\
\breve{x}_{ijt}^{(7)} &= x_{ijt} - (\overline{x}_{ij\cdot} - \overline{x}_{i\cdot\cdot}) - (\overline{x}_{i\cdot t} - \overline{x}_{\cdot\cdot t}) - (\overline{x}_{\cdot jt} - \overline{x}_{\cdot j\cdot}).
\end{aligned}
$$

Apparently, it is fine to allow x_{ijt} to contain the constant term because of the location identification restrictions imposed in Models 2-7. On the surface, when all seven models are under consideration, x_{ijt} cannot contain a nonconstant term that is only varying over two of the three indices. In other words, x_{ijt} needs to vary over all three dimensions. Otherwise, it can be absorbed into fixed effects and its slope coefficient cannot be estimated using the fixed effect regression. One simple example of x_{ijt} that only varies over two dimensions is the geographic distance (d_{ij}) between country i and country j, which is typically time-invariant. If x_{ijt} contains such regressors, we could consider two approaches. First, one can consider a small subset of the seven models to incorporate certain regressors that have variations only along one or two dimensions. Second, we can incorporate such regressors into the fixed effects in the corresponding model, perform the model selection as usual and then consider the estimation of the marginal effect of such regressors in the estimation step. For example, if other than the usual regressors in x_{ijt} that vary over i, j and t in Models 1–7, we also want to include d_{ij} as a regressor in these models. Now we are considering the seven models as follows:

$$
\begin{aligned}
\text{Model 1}: \quad & y_{ijt} = x'_{ijt}\beta + d_{ij}\vartheta + u_{ijt}, \\
\text{Model 2}: \quad & y_{ijt} = x'_{ijt}\beta + d_{ij}\vartheta + \alpha_i + \gamma_j + \lambda_t + u_{ijt}, \\
\text{Model 3}: \quad & y_{ijt} = x'_{ijt}\beta + d_{ij}\vartheta + \gamma_{ij} + u_{ijt}, \\
\text{Model 4}: \quad & y_{ijt} = x'_{ijt}\beta + d_{ij}\vartheta + \gamma_{ij} + \lambda_t + u_{ijt}, \\
\text{Model 5}: \quad & y_{ijt} = x'_{ijt}\beta + d_{ij}\vartheta + \alpha_{jt}^* + u_{ijt}, \\
\text{Model 6}: \quad & y_{ijt} = x'_{ijt}\beta + d_{ij}\vartheta + \alpha_{it} + \alpha_{jt}^* + u_{ijt}, \\
\text{Model 7}: \quad & y_{ijt} = x'_{ijt}\beta + d_{ij}\vartheta + \gamma_{ij} + \alpha_{it} + \alpha_{jt}^* + u_{ijt},
\end{aligned}
$$

In this case, we have an identification problem in Models 3, 4 and 7 because it is impossible to separately identify $d_{ij}\vartheta$ and γ_{ij} without further restrictions. Nevertheless, effectively, we can rewrite Models 3, 4 and 7 respectively as

$$
\begin{aligned}
\text{Model } 3': \quad & y_{ijt} = x'_{ijt}\beta + \tilde{\gamma}_{ij} + u_{ijt}, \\
\text{Model } 4': \quad & y_{ijt} = x'_{ijt}\beta + \tilde{\gamma}_{ij} + \lambda_t + u_{ijt}, \\
\text{Model } 7': \quad & y_{ijt} = x'_{ijt}\beta + \tilde{\gamma}_{ij} + \alpha_{it} + \alpha_{jt}^* + u_{ijt},
\end{aligned}
$$

where $\tilde{\gamma}_{ij} = d_{ij}\vartheta + \gamma_{ij}$. One can continue to apply our jackknife method to select among Models 1, 2, $3'$, $4'$, 5, 6 and $7'$ by comparing the out-of-sample predictability.[3]

---

[3] If any one of Models 1, 2, 5, and 6 is selected, then there is no problem to identify ϑ along with β. However, we cannot identify ϑ directly in Models 3, 4 and 7 from the usual fixed-effects estimation procedure when Model $3'$, $4'$ or $7'$ is selected.

Assumption A.2 requires that $\{u_{ijt}\}$ be weakly dependent along either one of the three dimensions. For example, Assumption A.2(iv) essentially requires that

$$\frac{T}{NM}\sum_{i,j}E(\overline{u}_{ij\cdot}^2) = \frac{1}{NMT}\sum_{i,j}\sum_{t=1}^{T}\sum_{s=1}^{T}E\left(u_{ijt}u_{ijs}\right)$$

should have a finite limit $\overline{\sigma}_{u4}^2$. The latter condition is satisfied by the Davydov inequality if $\{u_{ijt}, t\geq 1\}$ is strong mixing with finite $(2+\delta)$-th moment and mixing coefficients $\alpha_{ij}\left(\cdot\right)$ such that $\alpha_{ij}\left(\tau\right)=\tau^{-\gamma_{ij}}$ for some $\gamma_{ij}>(2+\delta)/\delta$; see, e.g., Bosq (1998, pp.19-20) or the online supplement of Su, Shi, and Phillips (2016). For another example, Assumption A.2(i) requires that $\{u_{ijt}\}$ be weakly cross-sectionally dependent along the j-dimension and weakly serially dependent as well such that $\frac{MT}{N}\sum_{i=1}^{N}E(\overline{u}_{i\cdot\cdot}^2) = \frac{1}{NMT}\sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{l=1}^{M}\sum_{t=1}^{T}\sum_{s=1}^{T}E\left(u_{ijt}u_{ils}\right)$ has a finite limit $\overline{\sigma}_{u1}^2$. In the special case where u_{ijt} is not correlated along either one of the three dimensions, we can easily verify that $\overline{\sigma}_{u\ell}^2 = \overline{\sigma}_{u}^2$ for $\ell=1,...,6$. In the presence of serial or cross-sectional correlations, $\overline{\sigma}_{u\ell}^2$'s are generally different from $\overline{\sigma}_{u}^2$.

Similarly, Assumptions A.3 requires that $\{x_{ijt}\}$ be weakly dependent along either one of the three dimensions. The conditions in this assumption can be verified via the Chebyshev or Markov inequality under some conditions to ensure such weak dependence. For example, to verify Assumption A.3(iv), by the Cauchy-Schwarz inequality it is sufficient to verify each diagonal element of $\frac{1}{NM}\sum_{i,j}(\overline{x}_{ij\cdot}-\overline{x})(\overline{x}_{ij\cdot}-\overline{x})'$ is $O_p(T^{-1})$. Let ν_l be a $k\times 1$ vector that contains 1 in its lth place and zeros elsewhere where $l=1,...,k$. Then

$$
\begin{aligned}
S_l &= \frac{1}{NM}\sum_{i,j}\nu_l'(\overline{x}_{ij\cdot}-\overline{x})(\overline{x}_{ij\cdot}-\overline{x})'\nu_l \\
&= \frac{1}{NM}\sum_{i,j}\nu_l'[\overline{x}_{ij\cdot}-E\left(\overline{x}_{ij\cdot}\right)][\overline{x}_{ij\cdot}-E\left(\overline{x}_{ij\cdot}\right)]'\nu_l - \frac{2}{NM}\sum_{i,j}\nu_l'[\overline{x}_{ij\cdot}-E\left(\overline{x}_{ij\cdot}\right)][\overline{x}-E\left(\overline{x}_{ij\cdot}\right)]'\nu_l \\
&\quad + \frac{1}{NM}\sum_{i,j}\nu_l'\left[\overline{x}-E\left(\overline{x}_{ij\cdot}\right)\right]\left[\overline{x}-E\left(\overline{x}_{ij\cdot}\right)\right]'\nu_l \\
&\equiv S_l\left(1\right)-2S_l\left(2\right)+S_l\left(3\right),\ \text{say.}
\end{aligned}
$$

Then $S_l=O_p(T^{-1})$ provided $E\left[S_l\left(\ell\right)\right]=O(T^{-1})$ for $\ell=1,2,3$ by the Markov inequality. Again, the latter is true under some weak dependence conditions. For example, if $\{x_{ijt}, t\geq 1\}$ is strong mixing satisfying certain mixing rate and moment conditions, then $E\left[S_l\left(1\right)\right]=\frac{1}{NMT^2}\sum_{i,j}\sum_{t=1}^{T}\sum_{s=1}^{T}\text{Cov}(\nu_l'x_{ijt},\nu_l'x_{ijs})=O\left(T^{-1}\right)$. Similar claims hold for $S_l\left(2\right)$ and $S_l\left(3\right)$. Note that Assumptions A.1(iii)-(iv), A.2, and A.3 imply the following results:

(i) $\frac{1}{N}\sum_i \overline{x}_{i\cdot\cdot}\overline{u}_{i\cdot\cdot} = O_p((NMT)^{-1/2}+(MT)^{-1})$;

(ii) $\frac{1}{M}\sum_j \overline{x}_{\cdot j\cdot}\overline{u}_{\cdot j\cdot} = O_p((NMT)^{-1/2}+(NT)^{-1})$;

(iii) $\frac{1}{T}\sum_t \overline{x}_{\cdot\cdot t}\overline{u}_{\cdot\cdot t} = O_p((NMT)^{-1/2}+(NM)^{-1})$;

(iv) $\frac{1}{NM}\sum_{i,j} \overline{x}_{ij\cdot}\overline{u}_{ij\cdot} = O_p((NMT)^{-1/2}+T^{-1})$;

(v) $\frac{1}{NT}\sum_{i,t} \overline{x}_{i\cdot t}\overline{u}_{i\cdot t} = O_p((NMT)^{-1/2}+M^{-1})$;

(vi) $\frac{1}{MT}\sum_{j,t} \overline{x}_{\cdot jt}\overline{u}_{\cdot jt} = O_p((NMT)^{-1/2}+N^{-1})$.

In this case, we could consider the following two-step post-selection procedure to estimate ϑ if needed: 1) In the first step, we obtain the consistent estimator $\hat{\gamma}_{ij}$ of $\tilde{\gamma}_{ij}$ based on Model 3′, 4′ or 7′, whichever is selected; 2) In the second step, we run a linear regression of $\hat{\gamma}_{ij}$ on d_{ij} to estimate ϑ under the additional identification restriction that d_{ij} and γ_{ij} are uncorrelated. Of course, one must take into account the estimation error from the first stage when making inference on ϑ. We leave the systematic study of this issue to future research.

For example, (i) holds because by the triangle and Cauchy-Schwarz inequalities and Assumptions A.1(iii)-(iv), A.2(i) and A.3(i) we have

$$
\begin{aligned}
\left| \frac{1}{N} \sum_i \nu_l' \overline{x}_{i..} \overline{u}_{i..} \right| &= \left| \frac{1}{N} \sum_i \nu_l' \left(\overline{x}_{i..} - \overline{x} \right) \overline{u}_{i..} + \nu_l' \overline{xu} \right| \\
&\leq \left\{ \nu_l' \frac{1}{N} \sum_i \left(\overline{x}_{i..} - \overline{x} \right) \left(\overline{x}_{i..} - \overline{x} \right) \nu_l' \right\}^{1/2} \left\{ \frac{1}{N} \sum_i \overline{u}_{i..}^2 \right\}^{1/2} + \| \nu_l' \overline{x} \| \, |\overline{u}| \\
&= O_p((MT)^{-1/2}) O_p((MT)^{-1/2}) + O_p((NMT)^{-1/2}).
\end{aligned}
$$

Assumption A.4 specifies conditions to ensure that the under-fitted models will never be chosen asymptotically. The interpretations of the positive probability limit conditions in Assumption A.4 are easy. For example, when Model 2 is the true model, Models 1, 3, and 5 are under-fitted. In this case, the positiveness of $\varphi_{2,m}$ requires that the additive fixed effects $\alpha_i + \gamma_j + \lambda_t$, when stacked into an $NMT \times 1$ vector, should not lie in the space spanned by the columns of the regressor matrix Z_m in Model m for $m = 1, 3,$ and 5, where we recall that $Z_1 = X$. Similarly, the zero probability limit conditions in Assumption A.4 require that the interactions between the idiosyncratic error terms and the fixed effects in the under-fitted models are asymptotically negligible.

Note that we allow for both weak cross-sectional and serial dependence of unknown form in $\{(x_{ijt}, u_{ijt})\}$ even though some results derived below need further constraints. We do not need identical distributions or homoskedasticity along either one of the three dimensions, neither do we need to assume mean or covariance stationarity along either dimension. In this sense, we say our results below apply to a variety of 3-D linear panel data models in practice.

Given Assumptions A.1-A.4, we are ready to state our first main result.

Theorem 3.1 *Suppose that Assumptions A.1-A.4 hold. Suppose that $\max_{1 \leq m \leq 6} \{ \overline{\sigma}_{u,m}^2 \} < 2\overline{\sigma}_u^2$, where $\overline{\sigma}_u^2$ and $\overline{\sigma}_{u,m}^2$ are defined in Assumptions A.1(i) and A.2, respectively. Then as $(N, M, T) \to \infty$*

$$
P(\hat{m} = m^* | \text{ Model } m^* \text{ is the true model}) \to 1 \text{ for } m^* = 1, ..., 7.
$$

Theorem 3.1 indicates that we can choose the correct model w.p.a.1 as $(N, M, T) \to \infty$ under some additional side conditions on $\overline{\sigma}_{u,m}^2$'s. Despite the complication in the asymptotic analysis of general 3-D models, the idea that outlines the proof of the above theorem is simple. When Model 1 is the true model (which is unlikely in practice), all the other models are over-fitted; when Model 7 is the true model, all other models are under-fitted. For $m^* \in \{2, 3, 4, 5, 6\}$, when Model m^* is the true model, we need to classify other models into either the under-fitted category or the over-fitted category. If we use $CV_{m^*,m}$ to denote $CV(m)$ when Model m^* is the true model and Model m is used for the cross-validation, we can show that

$$
CV_{m^*,m} - CV_{m^*,m^*} \xrightarrow{p} \varphi_{m^*,m} > 0
$$

for Model m that is under-fitted with respect to Model m^*. The limits $\varphi_{m^*,m}$ are defined in Assumption A.4. On the other hand, when Model m is over-fitted with respect to Model m^*, unsurprisingly $CV_{m^*,m} - CV_{m^*,m^*}$ converges to 0 in probability, and we need to blow it up by a term that is divergent with (N, M, T) and depends

on (m^*, m) in order to obtain a positive probability limit. That is, for some $\kappa_{m^*,m} \equiv \kappa_{m^*,m}(N, M, T)$, we have

$$\kappa_{m^*,m}[CV_{m^*,m} - CV_{m^*,m^*}] \overset{p}{\to} \psi_{m^*,m} > 0,$$

where $\kappa_{m^*,m} \to \infty$ as $(N, M, T) \to \infty$, and $\psi_{m^*,m}$ are constants that are always positive when $\max_{1 \le m \le 6} \{\overline{\sigma}_{u,m}^2\}$ $< 2\overline{\sigma}_u^2$ is satisfied. For example, when Model 2 is the true Model, it is easy to see that Models 1, 3, and 5 are under-fitted and Models 4, 6, and 7 are over-fitted. In this case, we have

$$
\begin{aligned}
CV_{2,m} - CV_{2,2} &\overset{p}{\to} \varphi_{2,m} > 0 \text{ for } m = 1, 3, 5, \\
T\left(CV_{2,4} - CV_{2,2}\right) &\overset{p}{\to} 2\overline{\sigma}_u^2 - \overline{\sigma}_{u4}^2 > 0, \\
(N \wedge M)\left(CV_{2,6} - CV_{2,2}\right) &\overset{p}{\to} q_6(2\overline{\sigma}_u^2 - \overline{\sigma}_{u6}^2) + q_7(2\overline{\sigma}_u^2 - \overline{\sigma}_{u5}^2) > 0, \\
(N \wedge M \wedge T)(CV_{2,7} - CV_{2,2}) &\overset{p}{\to} q_8(2\overline{\sigma}_u^2 - \overline{\sigma}_{u4}^2) + q_9(2\overline{\sigma}_u^2 - \overline{\sigma}_{u5}^2) + q_{10}(2\overline{\sigma}_u^2 - \overline{\sigma}_{u6}^2) > 0,
\end{aligned}
$$

where $N \wedge M = \min(N, M)$, $q_6 = \lim_{(N,M) \to \infty}(1 \wedge \frac{M}{N})$, $q_7 = \lim_{(N,M) \to \infty}(1 \wedge \frac{N}{M})$, $q_8 = \lim_{(N,M,T) \to \infty}$ $(1 \wedge \frac{N}{T} \wedge \frac{M}{T})$, $q_9 = \lim_{(N,M,T) \to \infty}(1 \wedge \frac{N}{M} \wedge \frac{T}{M})$ and $q_{10} = \lim_{(N,M,T) \to \infty}(1 \wedge \frac{M}{N} \wedge \frac{T}{N})$. As a result, we have $P(\hat{m} = 2 \mid \text{Model 2 is the true model}) \to 1$ as $(N, M, T) \to \infty$.

The side condition on $\overline{\sigma}_u^2$ and $\overline{\sigma}_{u,m}^2$ in Theorem 3.1 essentially says that we cannot have too much serial or cross-sectional correlation among the error terms. It is automatically satisfied if u_{ijt}'s are uncorrelated across all the i, j, t dimensions. When $\{u_{ijt}, t \ge 1\}$ follows an AR(1) process, we can follow Lu and Su (2019a) and demonstrate that this side condition requires that the AR(1) coefficient should lie in the interval $(-1, \frac{1}{3})$. If one doubts that strong serial correlations might be present, then we can consider the modified jackknife method in the next section. Similarly, if the cross-sectional dependence along the i and j dimensions is weak, such a side condition would be satisfied. When one suspects of strong cross-sectional dependence, one can model it, say, by extending the analysis of 2-D panels with multifactor error structure in Pesaran (2006), Bai (2009), and Lu and Su (2016) to that of 3-D panels. But this is certainly beyond the scope of the current paper.

Note that we do not need any relative rate conditions on how N, M, and T pass to infinity. Our theory works even if T is proportional to $\log N$ or $\log M$, and vice versa. Of course, the proof of the above theorem can be greatly simplified if one would like to impose conditions such that $T/(NM)^2 \to 0$, $M/(NT)^2 \to 0$ and $N/(MT)^2 \to 0$ as $(N, M, T) \to \infty$.

4　Methodology and Theory in the Presence of Strong Serially Correlated Errors

In this section we propose a modified jackknife method to choose different types of fixed effects when the error terms exhibit strong serial correlation, and then justify its consistency.

4.1　The modified jackknife method

To allow strong serial correlation among the error terms, we assume that $\{u_{ijt}, t \ge 1\}$ can be approximated by an AR(p) process:

$$u_{ijt} = \rho_1 u_{ij,t-1} + \rho_2 u_{ij,t-2} + \ldots + \rho_p u_{ij,t-p} + v_{ijt} = \rho' \underline{u}_{ij,t-1} + v_{ijt}, \tag{4.1}$$

14

where $i = 1, ..., N$, $j = 1, ..., M$, $t = p + 1, ..., T$, $\rho = (\rho_1, ..., \rho_p)'$ is a vector of unknown parameters, $\underline{u}_{ij,t-1} = (u_{ij,t-1}, ..., u_{ij,t-p})'$, and v_{ijt} is a zero mean innovation term.

We propose to obtain a consistent estimate of ρ based on the OLS residuals from the largest model under consideration: $\hat{u}_{ijt}^{(7)} = y_{ijt} - z'_{ijt,7}\hat{\theta}_7$. Given $\hat{u}_{ijt}^{(7)}$, we run the following AR(p) regression to estimate ρ :

$$\hat{u}_{ijt}^{(7)} = \rho_1 \hat{u}_{ij,t-1}^{(7)} + \rho_2 \hat{u}_{ij,t-2}^{(7)} + ... + \rho_p \hat{u}_{ij,t-p}^{(7)} + v_{ijt}^* = \rho' \underline{\hat{u}}_{ij,t-1}^{(7)} + v_{ijt}^*,$$

where $i = 1, ..., N$, $j = 1, ..., M$, $t = p+1, ..., T$, $\underline{\hat{u}}_{ij,t-1} = (\hat{u}_{ij,t-1}, ..., \hat{u}_{ij,t-p})'$ and $v_{ijt}^* = \rho'(\underline{u}_{ij,t-1} - \underline{\hat{u}}_{ij,t-1}^{(7)}) + v_{ijt}$. Let the $\hat{\rho} = (\hat{\rho}_1, ..., \hat{\rho}_p)'$ be the OLS estimator from the above regression. Let $\underline{y}_{ij,t-1} = (y_{ij,t-1}, ..., y_{ij,t-p})'$ and $\underline{\hat{y}}_{ij,t-1}^{(m)} = (\hat{y}_{ij,t-1}^{(m)}, ..., \hat{y}_{ij,t-p}^{(m)})'$. Then we consider the following modified CV function:

$$CV^*(m) = \frac{1}{NMT_p} \sum_{i=1}^{N} \sum_{j=1}^{M} \sum_{t=p+1}^{T} \left[(y_{ijt} - \hat{\rho}'\underline{y}_{ij,t-1}) - (\hat{y}_{ijt}^{(m)} - \hat{\rho}'\underline{\hat{y}}_{ij,t-1}^{(m)}) \right]^2,$$

where $T_p = T - p$. Define

$$\tilde{m} = \underset{1 \leq m \leq 7}{\arg\min}\ CV^*(m).$$

When Model m is the true model, we expect that $(y_{ijt} - \hat{\rho}'\underline{y}_{ij,t-1}) - (\hat{y}_{ijt}^{(m)} - \hat{\rho}'\underline{\hat{y}}_{ij,t-1}^{(m)})$ will approximate the true innovation term v_{ijt} and $P(\tilde{m} = m) \to 1$ as $(N, M, T) \to \infty$ as long as the correlation among $\{v_{ijt}\}$ is weak.

To proceed, define

$$\Phi(L) = 1 - \rho_1 L - \rho_2 L^2 - ... - \rho_p L^p,$$

where L is the lag operator. Let $\tilde{z}_{ijt}^{(m)} = \Phi(L)z_{ijt}^{(m)}$ for $t = p + 1, ..., T$, and $m = 1, ..., 7$. Let $\overline{v}_{i\cdot\cdot} = \frac{1}{MT_p}\sum_{j=1}^{M}\sum_{t=p+1}^{T} v_{ijt}, \overline{v}_{\cdot j\cdot} = \frac{1}{NT_p}\sum_{i=1}^{N}\sum_{t=p+1}^{T} v_{ijt}, \overline{v}_{ij\cdot} = \frac{1}{T_p}\sum_{t=p+1}^{T} v_{ijt}, \overline{v} = \frac{1}{NMT_p}\sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{t=p+1}^{T} v_{ijt}$. Define $\overline{v}_{i\cdot t}, \overline{v}_{\cdot jt}$, and $\overline{v}_{\cdot\cdot t}$ analogously to $\overline{u}_{i\cdot t}, \overline{u}_{\cdot jt}, \overline{u}_{\cdot\cdot t}$. For notational simplicity, we will write $\sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{t=p+1}^{T}$ and $\max_{1\leq i \leq N, 1\leq j\leq M, p+1\leq t\leq T}$ as $\sum_{i,j,t}$ and $\max_{i,j,t}$, respectively, in this section.

4.2 Asymptotic theory under strong serial dependence

To state the next result, we add the following set of assumptions.

Assumption A.5 (i) All the roots of $\Phi(z)$ lie outside the unit circle;
(ii) $E(v_{ijt}) = 0$, $\max_{i,j,t} E(v_{ijt}^2) \leq C$, and $\frac{1}{NMT_p}\sum_{i,j,t} v_{ijt}^2 \xrightarrow{p} \bar{\sigma}_v^2 > 0$;
(iii) $\frac{1}{NMT_p}\sum_{i,j,t} E(\|x_{ijt}\|^2 v_{ijt}^2) = O(1)$;
(iv) $\frac{1}{NMT_p}\sum_{i,j,t} \varsigma_{ijt} v_{ijt} = O_p((NMT)^{-1/2})$ for $\varsigma_{ijt} = 1, x_{ijt}, x_{ij,t-j}, u_{ij,t-j}$ for $j = 1, ..., p$.

Assumption A.6 There are positive finite numbers $\bar{\sigma}_{v\ell}^2$, $\ell = 1, 2, ..., 6$, such that
(i) $\frac{MT_p}{N}\sum_i \overline{v}_{i\cdot\cdot}^2 \xrightarrow{p} \bar{\sigma}_{v1}^2$;
(ii) $\frac{NT_p}{M}\sum_j \overline{v}_{\cdot j\cdot}^2 \xrightarrow{p} \bar{\sigma}_{v2}^2$;
(iii) $\frac{NM}{T_p}\sum_{t=p+1}^{T} \overline{v}_{\cdot\cdot t}^2 \xrightarrow{p} \bar{\sigma}_{v3}^2$;
(iv) $\frac{T_p}{NM}\sum_{i,j} \overline{v}_{ij\cdot}^2 \xrightarrow{p} \bar{\sigma}_{v4}^2$;
(v) $\frac{M}{NT_p}\sum_i \sum_{t=p+1}^{T} \overline{v}_{i\cdot t}^2 \xrightarrow{p} \bar{\sigma}_{v5}^2$;
(vi) $\frac{N}{MT_p}\sum_j \sum_{t=p+1}^{T} \overline{v}_{\cdot jt}^2 \xrightarrow{p} \bar{\sigma}_{v6}^2$.

Assumption A.7 (i) If Model 2 is the true model, there exist positive constants $\varphi_{2,m}^*$ for $m = 1, 3, 5$ such that $\frac{1}{NMT} \sum_{i,j,t} (\varsigma_{2ijt,m}^*)^2 \xrightarrow{p} \varphi_{2,m}^*$ and $\frac{1}{NMT} \sum_{i,j,t} \varsigma_{2ijt,m}^* v_{ijt} \xrightarrow{p} 0$, where $\varsigma_{2ijt,m}^* = \Phi(1)(\alpha_i + \gamma_j) + \Phi(L)\lambda_t - \breve{z}_{ijt,m}'(Z_m' Z_m)^{-1} Z_m' D_2 \underline{\pi}_2$;

(ii) If Model 3 is the true model, there exist positive constants $\varphi_{3,m}^*$ for $m = 1, 2, 5, 6$ such that $\frac{1}{NMT} \sum_{i,j,t} (\varsigma_{3ijt,m}^*)^2$ $\xrightarrow{p} \varphi_{3,m}^*$ and $\frac{1}{NMT} \sum_{i,j,t} \varsigma_{3ijt,m}^* u_{ijt} \xrightarrow{p} 0$, where $\varsigma_{3ijt,m}^* = \Phi(1)\gamma_{ij} - \breve{z}_{ijt,m}'(Z_m' Z_m)^{-1} Z_m' D_3 \underline{\pi}_3$;

(iii) If Model 4 is the true model, there exist positive constants $\varphi_{4,m}^*$ for $m = 1, 2, 3, 5, 6$ such that $\frac{1}{NMT} \sum_{i,j,t} (\varsigma_{4ijt,m}^*)^2 \xrightarrow{p} \varphi_{4,m}^*$ and $\frac{1}{NMT} \sum_{i,j,t} \varsigma_{4ijt,m}^* u_{ijt} \xrightarrow{p} 0$, where $\varsigma_{4ijt,m}^* = \Phi(1)\gamma_{ij} + \Phi(L)\lambda_t - \breve{z}_{ijt,m}'(Z_m' Z_m)^{-1} Z_m' D_4 \underline{\pi}_4$;

(iv) If Model 5 is the true model, there exist positive constants $\varphi_{5,m}^*$ for $m = 1, 2, 3, 4$ such that $\frac{1}{NMT} \sum_{i,j,t} (\varsigma_{5ijt,m}^*)^2 \xrightarrow{p} \varphi_{5,m}^*$ and $\frac{1}{NMT} \sum_{i,j,t} \varsigma_{5ijt,m}^* u_{ijt} \xrightarrow{p} 0$, where $\varsigma_{5ijt,m}^* = \Phi(L)\alpha_{jt}^* - \breve{z}_{ijt,m}'(Z_m' Z_m)^{-1} Z_m' D_5 \underline{\pi}_5$;

(v) If Model 6 is the true model, there exist positive constants $\varphi_{6,m}^*$ for $m = 1, 2, 3, 4, 5$ such that $\frac{1}{NMT} \sum_{i,j,t} (\varsigma_{6ijt,m}^*)^2 \xrightarrow{p} \varphi_{6,m}^* > 0$ and $\frac{1}{NMT} \sum_{i,j,t} \varsigma_{6ijt,m}^* u_{ijt} \xrightarrow{p} 0$, where $\varsigma_{6ijt,m}^* = \Phi(L)(\alpha_{it} + \alpha_{jt}^*) - \breve{z}_{ijt,m}'(Z_m' Z_m)^{-1} Z_m' D_6 \underline{\pi}_6$;

(vi) If Model 7 is the true model, there exist positive constants $\varphi_{7,m}^*$ for $m = 1, 2, 3, 4, 5, 6$ such that $\frac{1}{NMT} \sum_{i,j,t} (\varsigma_{7ijt,m}^*)^2 \xrightarrow{p} \varphi_{7,m}^*$ and $\frac{1}{NMT} \sum_{i,j,t} \varsigma_{7ijt,m}^* u_{ijt} \xrightarrow{p} 0$, where $\varsigma_{7ijt,m}^* = \Phi(1)\gamma_{ij} + \Phi(L)(\alpha_{it} + \alpha_{jt}^*) - \breve{z}_{ijt,m}'(Z_m' Z_m)^{-1} Z_m' D_7 \underline{\pi}_7$.

Assumptions A.5-A.6 and A.7 parallel Assumptions A.1-A.2 and A.4, respectively. Note that under Assumptions A.1(iii)-(iv), A.3, and A.6, we also have the following relationships:

(i) $\frac{1}{N} \sum_i \overline{x}_{i\cdot\cdot} \overline{v}_{i\cdot\cdot} = O_p((NMT)^{-1/2} + (MT)^{-1})$;

(ii) $\frac{1}{M} \sum_j \overline{x}_{\cdot j\cdot} \overline{v}_{\cdot j\cdot} = O_p((NMT)^{-1/2} + (NT)^{-1})$;

(iii) $\frac{1}{T_p} \sum_{t=p+1}^T \overline{x}_{\cdot\cdot t} \overline{v}_{\cdot\cdot t} = O_p((NMT)^{-1/2} + (NM)^{-1})$;

(iv) $\frac{1}{NM} \sum_{i,j} \overline{x}_{ij\cdot} \overline{v}_{ij\cdot} = O_p((NMT)^{-1/2} + T^{-1})$;

(v) $\frac{1}{NT_p} \sum_i \sum_{t=p+1}^T \overline{x}_{i\cdot t} \overline{v}_{i\cdot t} = O_p((NMT)^{-1/2} + M^{-1})$;

(vi) $\frac{1}{MT_p} \sum_j \sum_{t=p+1}^T \overline{x}_{\cdot jt} \overline{v}_{\cdot jt} = O_p((NMT)^{-1/2} + N^{-1})$.

The following theorem states the main result in this section.

Theorem 4.1 *Suppose Assumption A.1-A.3 and A.5-A.7 hold. Suppose that* $\max(\overline{\sigma}_{v1}^2, ..., \overline{\sigma}_{v6}^2) < 2\overline{\sigma}_v^2$. *Then as* $(N, M, T) \to \infty$,

$$P(\tilde{m} = m^* \mid Model \ m^* \ is \ the \ true \ model) \to 1 \ for \ m^* = 1, ..., 7.$$

Theorem 4.1 indicates that the modified jackknife method helps to choose the correct model under the weaker side condition $\max(\overline{\sigma}_{v1}^2, ..., \overline{\sigma}_{v6}^2) < 2\overline{\sigma}_v^2$. When there is no serial correlation among $\{u_{ijt}\}$ such that $\Phi(1) = \Phi(L) = 1$ and $u_{ijt} = v_{ijt}$, then $\overline{\sigma}_{v\ell}^2 = \overline{\sigma}_{u\ell}^2$ and $\overline{\sigma}_v^2 = \overline{\sigma}_u^2$. That is, the result in Theorem 4.1 now coincides with that in Theorem 3.1.

Note that we do not require $\{u_{ijt}, t \geq 1\}$ to exactly follow the AR(p) process. Essentially we prewhiten the error process via the AR(p) filtering with the expectation that the serial correlation among $\{u_{ijt}, t \geq 1\}$ will be sufficiently reduced after this procedure.

5   Monte Carlo Simulations

In this section, we examine the finite sample performance of our jackknife and modified jackknife methods. We compare them with the commonly used information criteria: AIC and BIC.[4] Specifically, let $\hat{u}_{ijt,m} = y_{ijt} - z'_{ijt,m}\hat{\theta}_m = y_{ijt} - (x'_{ijt}\hat{\beta}_m + d'_{ijt,m}\hat{\pi}_m)$ be the in-sample residual for Model m, where $m = 1, ..., 7$. Then AIC and BIC are defined respectively as

$$
\begin{aligned}
AIC\,(m) &= \ln\left((\hat{\sigma}^{(m)})^2\right) + \frac{2k_m}{NMT}, \\
BIC\,(m) &= \ln\left((\hat{\sigma}^{(m)})^2\right) + \frac{\log(NMT)\,k_m}{NMT},
\end{aligned}
$$

where $(\hat{\sigma}^{(m)})^2 = \frac{1}{NMT}\sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{t=1}^{T}(\hat{u}_{ijt,m})^2$ and k_m is the dimension of $(x'_{ijt}, d'_{ijt,m})'$. In the simulations, we find that BIC performs poorly, so we also modify BIC slightly as

$$
BIC_2\,(m) = \ln\left((\hat{\sigma}^{(m)})^2\right) + \frac{\log(\log(NMT))k_m}{NMT}.
$$

We consider the three different types of DGPs: (i) static panels, (ii) dynamic panels without exogenous regressors, and (iii) dynamic panels with exogenous regressors. For static panels, we allow serial correlation in the error terms. We consider the different combinations of $(N, M, T) = (10, 10, 5), (20, 20, 5), (10, 10, 10),$ $(10, 10, 20)$ and $(20, 20, 20).$ The number of replications is 1000.

5.1   Static panels

We consider seven static panel DGPs that correspond to Models 1-7 in (1.1), where x_{ijt} contains a constant and a scalar random variable, say, $\tilde{x}_{ijt}$ and the corresponding true β is $[1, 1]'$. All fixed effects, namely, $\alpha_i, \gamma_j, \lambda_t, \gamma_{ij}, \alpha_{it},$ and $\alpha^*_{jt},$ are IID $N(0, 1)$ random variables. To allow the correlation between $\tilde{x}_{ijt}$ and fixed effects, $\tilde{x}_{ijt}$'s are generated in DGPs 1-7 respectively as

$$
\begin{aligned}
&\text{DGP 1:} \quad \tilde{x}_{ijt} = 1 + \eta_{ijt}, &\qquad &\text{DGP 2:} \quad \tilde{x}_{ijt} = 1 + \alpha_i + \gamma_j + \lambda_t + \eta_{ijt}, \\
&\text{DGP 3:} \quad \tilde{x}_{ijt} = 1 + \gamma_{ij} + \eta_{ijt}, &\qquad &\text{DGP 4:} \quad \tilde{x}_{ijt} = 1 + \gamma_{ij} + \lambda_t + \eta_{ijt}, \\
&\text{DGP 5:} \quad \tilde{x}_{ijt} = 1 + \alpha^*_{jt} + \eta_{ijt}, &\qquad &\text{DGP 6:} \quad \tilde{x}_{ijt} = 1 + \alpha_{it} + \alpha^*_{jt} + \eta_{ijt}, \\
&\text{DGP 7:} \quad \tilde{x}_{ijt} = 1 + \gamma_{ij} + \alpha_{it} + \alpha^*_{jt} + \eta_{ijt},
\end{aligned}
\tag{5.1}
$$

where η_{ijt}'s are IID $N(0, 1)$. To allow serial correlation in the error term, u_{ijt} is generated as

$$
u_{ijt} = \rho u_{ij,t-1} + v_{ijt},
$$

where v_{ijt}'s are IID $N(0, 1)$. We consider $\rho = 0, \frac{1}{4},$ and $\frac{3}{4},$ which correspond to no, weak and strong serial correlations, respectively. As discussed above, if u_{ijt} follows an AR(1) process, our jackknife method only works for $\rho \in (-1, \frac{1}{3})$. Hence, $\rho = \frac{1}{3}$ corresponds to the cut-off point for our jackknife method to work, so we also consider $\rho = \frac{1}{3}$ in the simulation. The simulation results for $\rho = 0, \frac{1}{4}, \frac{1}{3}$ and $\frac{3}{4}$ are reported in Tables 3A-3D, respectively.

---

[4] To the best of our knowledge, there is no theoretical justification for AIC and BIC in the context of determining fixed effects in 3-D panels. In fact, we are not aware of any systematic study of alternative approaches in our context.

We first consider $\rho = 0$, i.e., there is no serial correlation. Even in this case, BIC breaks down apparently. For example, when the true model is Model 6 or 7 (M6 or M7), BIC chooses M1 with probability 1. For the other four methods, namely, the AIC, modified BIC, jackknife, and modified jackknife (labeled as AIC, BIC_2, CV, and CV* respectively in the table), we first consider the large T case ($T > 5$). All four methods work well when $T > 5$. When the true model is M1, M2, M4, M5, M6, or M7, the four methods all select the correct model with probability 1. When the true model is M3, the four methods all choose the correct model with a probability larger than 0.9. Among the four methods, our jackknife (CV) method performs slightly better than others. When T is small ($T = 5$), CV* performs poorly. This suggests that a large sample is required for CV* to work.

We next consider $\rho = \frac{1}{4}$. We first examine the large T case when $T > 5$. Our modified jackknife (CV*) performs best and jackknife (CV) performs slightly worse. But both outperform AIC and BIC_2. For example, when the true model is M6 and $(N, M, T) = (10, 10, 10)$, CV*, CV, AIC and BIC_2 select the correct model with probabilities of 1, 0.93, 0.36 and 0.28, respectively. Again, in this case, BIC breaks down. When T is small ($T = 5$), again we find that the performance of CV* is poor.

When $\rho = \frac{1}{3}$, only our modified jackknife (CV*) works for the large T case ($T > 5$). For example, when the true model is M6 and $(N, M, T) = (10, 10, 10)$, CV*, CV, AIC, BIC_2, and BIC select the correct model with probabilities of 1, 0.61, 0.07, 0.04 and 0, respectively. In general, the simulation results confirm our asymptotic theories. For example, as we discuss after Theorem 3.1, when the true model is M2, $T\left(CV_{2,4} - CV_{2,2}\right) \xrightarrow{p} 2\overline{\sigma}_u^2 - \overline{\sigma}_{u4}^2$, which equals 0 when $\rho = \frac{1}{3}$. This suggests that our jackknife (CV) method cannot distinguish Models 2 and 4 in this case. Correspondingly, in the simulations, we find that CV selects the M2 and M4 with probabilities of 0.56 and 0.44, respectively when the true model is M2 and $(N, M, T) = (20, 20, 20)$. When T is small ($T = 5$), no method works well.

When $\rho = \frac{3}{4}$, again only the modified jackknife (CV*) performs well when $T > 5$. For example, when the true model is M5, CV* selects the correct model with probability 1, while all other methods with probability 0. When $T = 5$, no method seems to work.

In sum, for these static panel DGPs, our jackknife performs the best in the absence of serial correlation. Our modified jackknife performs the best in the presence of serial correlation when T is relatively large. When T is small ($T = 5$) and serial correlation is strong, no method works well.

5.2 Dynamic panels without exogenous regressors

We consider seven AR(1) dynamic panel DGPs. In this case, we cannot allow for serial correlation in the error terms as it will result in the endogeneity issue so that the FE estimates are biased and the IV/GMM estimates are generally needed. Specifically, we consider seven DGPs as Models 1-7 in (1.1) where x_{ijt} contains a constant and the lagged dependent variable $y_{ij,t-1}$ and the corresponding true β is $(1, 0.75)'$. All the fixed effects $(\alpha_i, \gamma_j, \lambda_t, \gamma_{ij}, \alpha_{it}$, and $\alpha_{jt}^*)$ and u_{ijt}'s are IID $N(0, 1)$ random variables.

The simulation results are reported in Table 4. It shows that our jackknife method performs the best, followed by AIC and BIC_2. BIC performs the worst. For example, when the true model is M6 and $(N, M, T) = (10, 10, 10)$, CV, AIC, BIC, and BIC_2 select the correct model with probabilities of 0.98, 0.46, 0.42 and 0.37, respectively. When T is large ($T > 5$), our jackknife method can select the true model with probabilities larger than 0.90.

5.3 Dynamic panels with exogenous regressors

We consider seven dynamic panel DGPs with multiple exogenous regressors. The DGPs are the same as described in Models 1-7 in (1.1) where x_{ijt} is a 7×1 vector. The first element of x_{ijt} is constant. The second is the lagged dependent variable $y_{ij,t-1}$. The third is a random variable as in (5.1) and the rest four elements are IID $N(0,1)$ random variables. The corresponding true β is $(1, 0.75, 0.2, ..., 0.2)'$. All fixed effects and u_{ijt} are IID $N(0,1)$ random variables.

Table 5 presents the simulation results. Again, in general, our jackknife dominates other methods. It can select the true model with probabilities larger than 0.90 when T is large $(T > 5)$. The performance of AIC is similar to that of the jackknife except when the true model is M6, in which case the jackknife outperforms AIC significantly. BIC_2 is worse than the jackknife and AIC, but still better than BIC.

6 Empirical Applications

In this section, we provide two empirical applications of our new methods.

6.1 Technology and contractions

We apply our new method to study how technological characteristics interact with business cycles as in Samaniego and Sun (2015, SS hereafter). Specifically, SS are interested in examining which technological characteristics lead industries to experience the most difficulty during the recession period. Their main estimation equation corresponds to our Model 7 (using our notation):

$$Growth_{ijt} = \beta_0 + \beta_1 (Contraction_{it} \times X_j) + \beta_2 Controls_{ijt} + \gamma_{ij} + \alpha_{it} + \alpha_{jt}^* + u_{ijt},$$

where $Growth_{ijt}$ is a measure of growth in industry j in country i at year t, $Contraction_{it}$ is a binary variable, which equals 1 if country i is in a contraction in year t, X_j is an industry technological characteristic, and $Controls_{ijt}$ is a control variable.

SS consider three measures of the growth variable, $Growth_{ijt}$: (i) value added (the log changes in industry value added), (ii) output (the log changes in gross output) and (iii) output index (the log changes in the Laspeyres production index). There are ten measures of the industry characteristic, X_j, (i) external finance dependence (EFD), (ii) depreciation (DEP), (iii) investment-specific technical change (ISTC), (iv) R&D intensity (RND), (v) human capital intensity (HC), (vi) labor intensity (LAB), (vii) fixity (FIX), $(viii)$ investment lumpiness (LMP), (ix) relationship-specific investment (SPEC), and (v) intermediate inputs intensity (INT). The control variable is the share of the industry value added out of the manufacturing industry at year $t-1$. For the detailed explanations of these variables, see SS (Section 3). The dataset covers 139 countries and 28 manufacturing industries from 1970 to 2007. Hence, $(N, M, T) = (139, 28, 38)$. There are a large number of missing values. The exact total sample size depends on the dependent variable. For example, there are 57,115 observations for the value added growth. SS adopt the largest model (Model 7). Using a too large model can result in substantial estimation efficiency loss. Here it is an interesting question to decide which model is the most appropriate among the seven models considered above.

As in SS's Table 5, we first run the growth regression using one measure of growth ($Growth_{ijt}$) as the dependent variable and the interaction term between one measure of industry characteristic (X_j) and contraction as the key regressor. Therefore, the dimension of the regressors is $k = 3$ (including the constant and control variable). We consider a total of 30 different combinations of $Growth_{ijt}$ and X_j. It is interesting to find that for all the 30 regressions, the jackknife, modified jackknife, AIC, and BIC_2 all select Model 6, while BIC selects Model 2. Table 6A also contains the estimates of β_1 and its 95% confidence intervals (CI) for all seven models. Based on the selected Model 6, the estimate of β_1 is -0.013 with a 95% CI of [-0.0242, -0.0018]. To save space, we only report the numerical results for $Growth_{ijt}$ being value added and X_j being EFD in Table 6A. The results for the other 29 regressions are available upon request.

We also run three regressions by including all the ten industry characteristics for the three dependent variables. Hence, the number of the regressors is $k = 12$ (including the constant and control variable). Again, for all three regressions, the jackknife, modified jackknife, AIC and BIC_2 all select Model 6, while BIC selects Model 2. The numerical results for $Growth_{ijt}$ being value added are reported in Table 6B. Based on the selected Model 6, the estimate of β_1 is -0.0163 with a 95% CI of [-0.0439, 0.0113]. The results for other two regressions are available upon request.

Considering the poor performance of BIC in the simulations, we conclude that Model 6 is an appropriate model for this application. Recall that Model 6 only includes α_{it} and α_{jt}^* as fixed effects. This suggests that after including the country-time and industry-time effects, it is redundant to include country-industry effects.

6.2 Gravity equations in international trade

Gravity equations are widely used to model bilateral trade. It is basically assumed that the bilateral trade volumes depend on the economic sizes (often using GDP measurements) and distance between two economies, which mirrors the physical gravity equation. For a review on gravity models, see Head and Mayer's (2014) chapter in the Handbook of International Economics. Different fixed effect models have been applied to estimate gravity equations, as we have discussed in Section 2 above.

We apply our new method to determine the fixed effect specifications in bilateral trade data. We first consider one basic gravity equation:

$$\ln\left(Export_{ijt}\right) = \beta_0 + \beta_1 \ln\left(GDP_{it} + GDP_{jt}\right) + \text{fixed effects} + u_{ijt},$$

where $Export_{ijt}$ is the export of country i to country j in year t, and GDP_{it} and GDP_{jt} are the GDPs of countries i and j, respectively at year t, fixed effects are specified as in our Models 1-7 in (1.1). Here $GDP_{it} + GDP_{jt}$ represents the total economic size of country i and country j. Baltagi et al. (2003) also consider the same form of regressors. Note that we do not include the distance between country i and country j as a regressor, as the distance is time-invariant and its effect is not identified under our Models 3, 4 and 7. The sample includes 35 OECD countries over 58 years (1949-2006). Thus, here $N = 35$, $M = 34$, and $T = 58$. With missing values, the total sample size is 48,403. The data are obtained from the companion website of Head and Mayer (2014). For this regression, we find that the jackknife, modified jackknife, AIC, and BIC_2 all select Model 7 as the correct model, while BIC selects Model 4. The numerical results are shown in Table 7A. We also report the estimate and 95% CI for β_1 for all seven models. Based on the selected model, the estimate of β_1 is 0.657 and its 95% CI is [0.577, 0.738].

We also modify the equation above by adding the population variables, i.e., we consider

$$\ln\left(Export_{ijt}\right) = \beta_0 + \beta_1 \ln\left(GDP_{it} + GDP_{jt}\right) + \beta_2 \ln\left(POP_{it} + POP_{jt}\right) + \text{fixed effects} + u_{ijt},$$

where POP_{it} and POP_{jt} are populations of countries i and j in year t. Again, the jackknife, modified jackknife, AIC, and BIC$_2$ all select Model 7, while BIC selects Model 4, as shown in Table 7B. Based on the selected Model 7, the effects of GDP and population are both positive and statistically significant.

We conclude that for gravity models, all fixed effects γ_{ij}, α_{it}, and α_{jt}^* are important, which suggests that there is substantial unobservable heterogeneity in the bilateral traded data.

7 Conclusion

In this paper, we propose a jackknife method to choose between a subset of 3-D panel data models with fixed effects that are widely used in the literature. We show that the method can consistently select the true model when the serial or cross-sectional correlations in the error terms are not strong. In the case where the error terms exhibit the strong serial correlation, we propose a modified jackknife method. Simulations are conducted to evaluate the finite sample performance of our methods. We apply our methods to two datasets to study the interaction between technological characteristics and business cycles and the gravity equations in international trades.

There are several interesting issues for future research. First, we can consider a broader class of 3-D panel models and conjecture that our theory continues to hold under some regularity conditions. Second, even though we only focus on balanced panels for notational simplicity, we remark that our theories for the unbalanced panels are still valid with obvious modifications. In particular, we now need that N, $\min_{1 \leq i \leq N} M_i$, and $\min_{1 \leq i \leq N, 1 \leq j \leq M_i} T_{ij}$ pass to infinity jointly. Third, we only propose a modified jackknife method to handle strong serial correlations and it is not clear how to take into account strong cross-sectional correlations. If one believes that the strong cross-sectional correlation may be present in the error terms, we may consider the use of a multi-factor error model from the scratch. The problem is that there exist many ways to model cross-sectional dependence in 3-D models. Recently, Lu and Su (2019b) have studied a 3-D factor model that contains one type of global factors and two types of local factors. As one can imagine, it is highly challenging to determine the number of global and local factors in such a factor model, not to mention the estimation and inference on the factors and factor loadings. When we include regressors in the above 3-D factor model, the issue becomes even harder and we are not sure whether one can apply the jackknife idea to determine what kinds of fixed effects should be included in the model. We leave these topics for future research.

REFERENCES

Allen, D. M. (1974). The relationship between variable selection and data augmentation and a method for prediction. *Technometrics* 16, 125-127.

Andrews, D. W. K. (1991). Asymptotic optimality of generalized C$_L$, cross-validation, and generalized cross-validation in regression with heteroskedastic errors. *Journal of Econometrics* 47, 359-377.

Anderson, J. and E. van Wincoop (2003). Gravity with gravitas: a solution to the border puzzle. *American Economic Review* 93, 170-192.

Bai, J. (2009). Panel data models with interactive fixed effects. *Econometrica* 77, 1229-1279.

Baier, S. L., and J. H. Bergstrand (2007). Do free trade agreements actually increase members' international trade? *Journal of international Economics* 71, 72-95.

Balazsi, L., Mátyás, L., and Wansbeek, T., (2017). Fixed effects models. In L. Matyas (eds), *The Econometrics of Multidimensional Panels: Theory and Applications*, pp. 1-34, Springer, Switzerland.

Balazsi, L., Mátyás, L. and Wansbeek, T., (2018). The estimation of multidimensional fixed effects panel data models. *Econometric Reviews* 37, 212-227.

Baltagi, B. H. (2013). *Econometric Analysis of Panel Data*, 5th edition. John Wiley & Sons, West Sussex.

Baltagi, B. H., P. Egger, and M. Pfaffermayr (2003). A generalized design for bilateral trade flow models. *Economics Letters* 80, 391-397.

Baldwin, R. and D. Taglioni (2006). Gravity for dummies and dummies for gravity equations. *Working Paper No. 12516*, National Bureau of Economic Research.

Berthélemy, J-C., and A. Tichit (2004). Bilateral donors' aid allocation decisions– a three-dimensional panel analysis. *International Review of Economics and Finance* 13, 253-274.

Bosq, D. (1998). *Nonparametric Statistics for Stochastic Processes*. Springer, New York.

Cheng, I. H. and H. J. Wall (2005). Controlling for Heterogeneity in Gravity Models of Trade and Integration. *Federal Reserve Bank of St. Louis* 87, 49-63.

Davis, P. (2002). Estimating multi-way error components models with unbalanced data structures. *Journal of Econometrics* 106, 67-85.

Dhaene, G. and K. Jochmans (2015). Split-panel jackknife estimation of fixed-effect models. *Review of Economic Studies* 82, 991-1030.

Efron, B. (1983). Estimating the error rate of a prediction rule: improvement on cross-validation, *Journal of the American Statistical Association* 78, 316-331.

Efron, B. (1986). How biased is the apparent error rate of a prediction rule? *Journal of the American Statistical Association* 81, 461-470.

Egger, P. (2000). A note on the proper econometric specification of the gravity equation. *Economics Letters* 66 25-31.

Egger, P. and M. Pfaffermayr (2003). The proper panel econometric specification of the gravity equation: A three-way model with bilateral interaction effects. *Empirical Economics* 28, 571-580.

Feenstra, R. C. (2016). *Advanced International Trade: Theory and Evidence*. Princeton University Press, Princeton, New Jersey.

Geisser, S. (1974). The predictive sample reuse method with applications. *Journal of the American Statistical Association* 70, 320-328.

Goldhaber, D. D., D. J. Brewer, and D. J. Anderson (1999). A three-way error components analysis of educational productivity. *Education Economics* 7, 199-208.

Hansen, B. E. and J. Racine (2012). Jackknife model averaging. *Journal of Econometrics* 167, 38-46.

Head, K. and T. Mayer (2014). Gravity equations: Workhorse, toolkit, and cookbook. In G. Gopinath, E. Helpman, & K. Rogoff (Eds.), *Handbook of International Economics*, Vol. 4, Elsevier-North Holland Amsterdam, pp. 131-195.

Hsiao, C. (2014). *Analysis of Panel Data*, 3rd edition. Cambridge University Press, Cambridge.

Leeb, H. and B. Pötscher (2005). Model selection and inference: Facts and fiction. *Econometric Theory* 21, 21-59.

Li, K.-C. (1987). Asymptotic optimality for C_p, C_L, cross-validation and generalized cross-validation: discrete index set. *Annals of Statistics* 15, 958-975.

Lu, X. and L. Su (2015). Jackknife model averaging for quantile regressions. *Journal of Econometrics*, 188, 40-58.

Lu, X. and L. Su (2016). Shrinkage estimation of dynamic panel data models with interactive fixed effects. *Journal of Econometrics* 190, 148-175.

Lu, X. and L. Su (2019a). Determining individual or time fixed effects in panel data models. *Journal of Econometrics*, forthcoming.

Lu, X., and L. Su (2019b). Estimation and inference for three-dimensional factor models. *Working Paper*, Singapore Management University.

Mátyás, L. (1997). Proper econometric specification of the gravity model. *The World Economy* 20, 363-369.

Mátyás, L. (2017). *The Econometrics of Multi-dimensional Panels: Theory and Applications*. Springer, Switzerland.

Pesaran, M. H. (2006). Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica* 74, 967-1012.

Picard, R. R., and R. D. Cook (1984). Cross-validation of regression models. *Journal of the American Statistical Association* 79, 575-583.

Quenouille, M.H. (1956). Notes on bias in estimation. *Biometrika* 43, 353-360.

Samaniego, R. M. and J. Y. Sun (2015). Technology and contractions: evidence from manufacturing. *European Economic Review* 79, 172-195.

Shao, J. (1993). Linear model selection by cross-validation. *Journal of the American Statistical Association* 88, 486-494.

Shimer, R. and L. Smith (2000). Assortative matching and search. *Econometrica* 68, 343–369.

Stone, C. (1974). Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society, Series B* 36, 111-147.

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

Su, L., Z. Shi, and P. C. B. Phillips (2016). Identifying latent structures in panel data. *Econometrica* 84, 2215-2264.

Tukey, J. W. (1958). Bias and confidence in not-quite large sample. *Annals of Mathematical Statistics* 29, 614.

Wahba, G. and S. Wold (1975). A completely automatic french curve: fitting spline functions by cross-validation. *Communications in Statistics* 4, 1-17.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*, 2nd edition. MIT Press, Massachusetts.

Wu, J. and Li, G. (2014). Moment-based tests for individual and time effects in panel data models. *Journal of Econometrics* 178, 569-581.

Appendix

A Proof of Theorem 3.1

In this appendix, we first state some technical lemmas that are used in the proof of Theorem 3.1. The proof of Theorem 4.1 is relegated to Appendix B in the online supplement. The proofs of all technical lemmas are given in Appendix C of the online supplement. Let $\sum_{i,j,t} = \sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{t=1}^{T}$ and similarly for $\sum_{i,j}$, $\sum_{i,t}$, $\sum_{j,t}$, $\sum_{i}$, $\sum_{j}$ and $\sum_{t}$.

Lemma A.1 *Let $Z = (X, D)$ and $M_D = I_{NMT} - D(D'D)^{-1}D'$. If both $D'D$ and $X'M_DX$ are nonsingular, then,*

$$(Z'Z)^{-1} = \begin{pmatrix} X_D^* & -X_D^*X'D(D'D)^{-1} \\ -(D'D)^{-1}D'XX_D^* & (D'D)^{-1} + (D'D)^{-1}D'XX_D^*X'D(D'D)^{-1} \end{pmatrix},$$

where $X_D^* = (X'M_DX)^{-1}$

Lemma A.2 *Let D_I, D_J, D_T, D_{IJ}, D_{JT}, D_{IJ}^*, D_{IT}^*, and D_{JT}^* be as defined in Section 2. Then*
  (i) $D_I \perp D_J \perp D_T$, $D_{IJ} \perp D_T$, $D_{IT}^* \perp D_{JT}$, $D_{IT}^* \perp D_{IJ}^* \perp D_{JT}^*$;
  (ii) $P_{D_2} = P_I + P_J + P_T$, $P_{D_4} = P_{IJ} + P_T$, $P_{D_6} = P_{IT}^* + P_{JT}$, $P_{D_7} = P_{IT}^* + P_{IJ}^* + P_{JT}^*$;
  (iii) $P_I = (I_N - \frac{J_N}{N}) \otimes \frac{J_M}{M} \otimes \frac{J_T}{T}$, $P_J = \frac{J_N}{N} \otimes (I_M - \frac{J_M}{M}) \otimes \frac{J_T}{T}$, $P_T = \frac{J_N}{N} \otimes \frac{J_M}{M} \otimes (I_T - \frac{J_T}{T})$;
  (iv) $P_{IJ} = (I_{NM} - \frac{J_{NM}}{NM}) \otimes \frac{J_T}{T}$, $P_{JT} = \frac{J_N}{N} \otimes (I_{MT} - \frac{J_{MT}}{MT})$;
  (v) $P_{IT}^* = (I_N - \frac{J_N}{N}) \otimes \frac{J_M}{M} \otimes I_T$, $P_{IJ}^* = I_N \otimes (I_M - \frac{J_M}{M}) \otimes \frac{J_T}{T}$, $P_{JT}^* = \frac{J_N}{N} \otimes I_M \otimes (I_T - \frac{J_T}{T})$.

Lemma A.3 *Let $A = (a_{ijt})$ and $B = (b_{ijt})$ be either X or U. Let $\overline{a}$, $\overline{a}_{i..}$, $\overline{a}_{.j.}$, $\overline{a}_{..t}$, $\overline{a}_{ij.}$, $\overline{a}_{jt}$, and $\overline{a}_{i.t}$ be defined in the same way as $\overline{x}$, $\overline{x}_{i..}$, $\overline{x}_{.j.}$, $\overline{x}_{..t}$, $\overline{x}_{ij.}$, $\overline{x}_{jt}$, and $\overline{x}_{i.t}$, and similarly for $\overline{b}$, $\overline{b}_{i..}$, $\overline{b}_{.j.}$, $\overline{b}_{..t}$, $\overline{b}_{ij.}$, $\overline{b}_{jt}$, and $\overline{b}_{i.t}$. Then*
  (i) $\frac{1}{NMT}A'P_IB = \frac{1}{N}\sum_i \overline{a}_{i..}\overline{b}_{i..}' - \overline{a}\overline{b}'$,
  (ii) $\frac{1}{NMT}A'P_JB = \frac{1}{M}\sum_j \overline{a}_{.j.}\overline{b}_{.j.}' - \overline{a}\overline{b}'$,
  (iii) $\frac{1}{NMT}A'P_TB = \frac{1}{T}\sum_t \overline{a}_{..t}\overline{b}_{..t}' - \overline{a}\overline{b}'$,
  (iv) $\frac{1}{NMT}A'P_{IJ}B = \frac{1}{NM}\sum_{i,j} \overline{a}_{ij.}\overline{b}_{ij.}' - \overline{a}\overline{b}'$,
  (v) $\frac{1}{NMT}A'P_{JT}B = \frac{1}{MT}\sum_{j,t} \overline{a}_{jt}\overline{b}_{jt}' - \overline{a}\overline{b}'$,
  (vi) $\frac{1}{NMT}A'P_{IJ}^*B = \frac{1}{NM}\sum_{i,j} \overline{a}_{ij.}\overline{b}_{ij.}' - \frac{1}{N}\sum_i \overline{a}_{i..}\overline{b}_{i..}'$,
  (vii) $\frac{1}{NMT}A'P_{IT}^*B = \frac{1}{NT}\sum_{i,t} \overline{a}_{i.t}\overline{b}_{i.t}' - \frac{1}{T}\sum_t \overline{a}_{..t}\overline{b}_{..t}'$,
  (viii) $\frac{1}{NMT}A'P_{JT}^*B = \frac{1}{MT}\sum_{j,t} \overline{a}_{jt}\overline{b}_{jt}' - \frac{1}{M}\sum_j \overline{a}_{.j.}\overline{b}_{.j.}'$.

Lemma A.4 *Suppose that Assumptions A.1(iv), A.2, and A.3 hold. Then*
  (i) $\frac{1}{NMT}X'M_{D_2}U = O_p((NM)^{-1} + (MT)^{-1} + (NT)^{-1} + (NMT)^{-1/2})$,
  (ii) $\frac{1}{NMT}X'M_{D_3}U = O_p(T^{-1} + (NMT)^{-1/2})$,
  (iii) $\frac{1}{NMT}X'M_{D_4}U = O_p(T^{-1} + (NM)^{-1} + (NMT)^{-1/2})$,
  (iv) $\frac{1}{NMT}X'M_{D_5}U = O_p(N^{-1} + (NMT)^{-1/2})$,
  (v) $\frac{1}{NMT}X'M_{D_6}U = O_p(N^{-1} + M^{-1} + (NMT)^{-1/2})$,
  (vi) $\frac{1}{NMT}X'M_{D_7}U = O_p(N^{-1} + M^{-1} + T^{-1} + (NMT)^{-1/2})$.

Lemma A.5 *Let $d_{ijt,m}$ be a typical element of D_m for $m = 2, ..., 7$. Then*
  (i) $d_{ijt,2}'(D_2'D_2)^{-1}d_{ijt,2} = \frac{1}{NT} + \frac{1}{MT} + \frac{1}{NM} - \frac{3}{NMT}$,
  (ii) $d_{ijt,3}'(D_3'D_3)^{-1}d_{ijt,3} = \frac{1}{T} - \frac{1}{NMT}$,
  (iii) $d_{ijt,4}'(D_4'D_4)^{-1}d_{ijt,4} = \frac{1}{T} + \frac{1}{NM} - \frac{2}{NMT}$,
  (iv) $d_{ijt,5}'(D_5'D_5)^{-1}d_{ijt,5} = \frac{1}{N} - \frac{1}{NMT}$,
  (v) $d_{ijt,6}'(D_6'D_6)^{-1}d_{ijt,6} = \frac{1}{N} + \frac{1}{M} - \frac{1}{NM} - \frac{1}{NMT}$,
  (vi) $d_{ijt,7}'(D_7'D_7)^{-1}d_{ijt,7} = \frac{1}{N} + \frac{1}{M} + \frac{1}{T} - \frac{1}{NM} - \frac{1}{NT} - \frac{1}{MT}$.

Lemma A.6 Let $h_{ijt,m} = z'_{ijt,m}(Z'_m Z_m)^{-1} z_{ijt,m}$ and $c_{ijt,m} = \frac{1}{1-h_{ijt,m}}$ for $m = 1, ..., 7$. Let $h^*_{ijt,m} = [x_{ijt} - X'D_m(D'_m D_m)^{-1} d_{ijt,m}]' X^*_{D_m}[x_{ijt} - X'D_m(D'_m D_m)^{-1} d_{ijt,m}]$ for $m = 2, ..., 7$ where $X^*_{D_m} = (X'M_{D_m}X)^{-1}$. Let $\bar{d}_m = d'_{ijt,m}(D'_m D_m)^{-1} d_{ijt,m}$ for $m = 2, ..., 7$, which does not vary over (i, j, t) by Lemma A.5. Let $\bar{d}_1 = 0$ and $h^*_{ijt,1} = x'_{ijt}(X'X)^{-1} x_{ijt}$. Suppose Assumption A.1(ii) and (v) holds. Then for $m = 1, ..., 7$ we have

(i) $h_{ijt,m} = \bar{d}_m + h^*_{ijt,m}$;

(ii) $\max_{i,j,t} h_{ijt,m} = O_p((NMT)^{-1/2} + \bar{d}_m) = o_p(1)$;

(iii) $\max_{i,j,t} |c_{ijt,1} - 1| = o_p(1)$;

(iv) $c^2_{ijt,m} - 1 - 2h_{ijt,m} = \frac{3h_{ijt,m} - 2h^2_{ijt,m}}{(1-h_{ijt,m})^2} \cdot h_{ijt,m}$ and $\max_{i,j,t} \left| \frac{3h_{ijt,m} - 2h^2_{ijt,m}}{(1-h_{ijt,m})^2} \right| = o_p(1)$;

(v) $c^2_{ijt,m} - 1 - 2h_{ijt,m} - 3h^2_{ijt,m} = \left[\frac{3-2h_{ijt,m}}{(1-h_{ijt,m})^2} - 3 \right] h^2_{ijt,m}$ and $\max_{i,j,t} \left| \frac{3-2h_{ijt,m}}{(1-h_{ijt,m})^2} - 3 \right| = o_p(1)$.

Lemma A.7 Suppose that the true model is $y_{ijt} = x'_{ijt}\beta + d^{*'}_{i,j,t}\pi^* + u_{ijt}$, with dummy matrix $D^* = D_{m^*} = \{d^*_{i,j,t}\}$ and coefficient vector $\pi^* = \pi_{m^*}$. For the leave-one-out prediction $\hat{y}_{ijt,m}$ using model $m \in \{1, 2, ..., 7\}$, we have[5]

(i) $y_{ijt} - \hat{y}_{ijt,m} = \frac{e_{ijt,m}}{1-h_{ijt,m}}$ where $h_{ijt,m}$ is defined in Lemma A.6 and $e_{ijt,m} = y_{ijt} - z'_{ijt,m}\hat{\theta}_m$;

(ii) $e_{ijt,m} = A_{ijt,m} + B_{ijt,m} + C_{ijt,m}$, where $A_{ijt,m} = u_{ijt} - d'_{ijt,m}(D'_m D_m)^{-1} D'_m U$, $B_{ijt,m} = d^{*'}_{ijt}\pi^* - x'_{ijt}X^*_{D_m}X'M_{D_m}D^*\pi^* - d'_{ijt,m}(D'_m D_m)^{-1} D'_m D^*\pi^* + d'_{ijt,m}(D'_m D_m)^{-1} D'_m X X^*_{D_m}X'M_{D_m}D^*\pi^*$, and $C_{ijt,m} = d'_{ijt,m}(D'_m D_m)^{-1} D'_m X X^*_{D_m}X'M_{D_m}U - x'_{ijt}X^*_{D_m}X'M_{D_m}U$,

(iii) $\sum_{i,j,t} A^2_{ijt,m} = U'M_{D_m}U$, $\sum_{i,j,t} B^2_{ijt,m} = \pi^{*'}D^{*'}M_{D_m}(I-P_{M_{D_m}X})M_{D_m}D^*\pi^*$, $\sum_{i,j,t} C^2_{ijt,m} = U'M_{D_m} \times XX^*_{D_m}X'M_{D_m}U$, $\sum_{i,j,t} A_{ijt,m}B_{ijt,m} = U'M_{D_m}(I-P_{M_{D_m}X})M_{D_m}D^*\pi^*$, $\sum_{i,j,t} A_{ijt,m}C_{ijt,m} = -U'M_{D_m}XX^*_{D_m} \times X'M_{D_m}U$, and $\sum_{i,j,t} B_{ijt,m}C_{ijt,m} = 0$.

Lemma A.8 Let $E_m = \frac{1}{NMT}\sum_{i,j,t} c^2_{ijt,m}(C^2_{ijt,m} + 2A_{ijt,m}C_{ijt,m})$ for $m = 1, 2, ..., 7$. Suppose that Assumptions A.1-A.3 hold. Then

(i) $E_1 = O_p((NMT)^{-1})$,

(ii) $E_2 = O_p((NM)^{-2} + (NT)^{-2} + (MT)^{-2} + (NMT)^{-1})$,

(iii) $E_3 = O_p(T^{-2} + (NMT)^{-1})$,

(iv) $E_4 = O_p((NM)^{-2} + T^{-2} + (NMT)^{-1})$,

(v) $E_5 = O_p(N^{-2} + (NMT)^{-1})$,

(vi) $E_6 = O_p(N^{-2} + M^{-2} + (NMT)^{-1})$,

(vii) $E_7 = O_p(N^{-2} + M^{-2} + T^{-2} + (NMT)^{-1})$.

Lemma A.9 Let $F_m = \frac{1}{NMT}\sum_{i,j,t} c^2_{ijt,m}(B^2_{ijt,m} + 2A_{ijt,m}B_{ijt,m} + 2B_{ijt,m}C_{ijt,m})$ for $m = 1, 2, ..., 7$. Suppose Assumption A.1(v) holds. If Model m is just- or over-fitted, then $F_m = 0$.

Lemma A.10 Suppose that Assumptions A.1 and A.2 hold. Let $H_m = \frac{1}{NMT}\sum_{i,j,t} c^2_{ijt,m}A^2_{ijt,m}$ for $m = 1, 2, ..., 7$. Then

(i) $H_1 = \frac{1}{NMT}\sum_{i,j,t} u^2_{ijt} + O_p((NMT)^{-1})$,

(ii) $H_2 = (1 + \frac{2}{NM} + \frac{2}{MT} + \frac{2}{NT})\frac{1}{NMT}\sum_{i,j,t} u^2_{ijt} - \frac{1}{N}\sum_i \bar{u}^2_{i\cdot\cdot} - \frac{1}{M}\sum_j \bar{u}^2_{\cdot j\cdot} - \frac{1}{T}\sum_t \bar{u}^2_{\cdot\cdot t} + O_p((NM)^{-2} + (NT)^{-2} + (MT)^{-2} + (NMT)^{-1})$,

(iii) $H_3 = (1 + \frac{2}{T})\frac{1}{NMT}\sum_{i,j,t} u^2_{ijt} - \frac{1}{NM}\sum_{i,j} \bar{u}^2_{ij\cdot} + O_p(T^{-2} + (NMT)^{-1})$,

(iv) $H_4 = (1 + \frac{2}{T} + \frac{2}{NM})\frac{1}{NMT}\sum_{i,j,t} u^2_{ijt} - \frac{1}{NM}\sum_{i,j} \bar{u}^2_{ij\cdot} - \frac{1}{T}\sum_t \bar{u}^2_{\cdot\cdot t} + O_p((NM)^{-2} + T^{-2} + (NMT)^{-1})$,

(v) $H_5 = (1 + \frac{2}{N})\frac{1}{NMT}\sum_{i,j,t} u^2_{ijt} - \frac{1}{MT}\sum_{j,t} \bar{u}^2_{\cdot jt} + O_p(N^{-2} + (NMT)^{-1})$,

(vi) $H_6 = (1 + \frac{2}{N} + \frac{2}{M})\frac{1}{NMT}\sum_{i,j,t} u^2_{ijt} - \frac{1}{MT}\sum_{j,t} \bar{u}^2_{\cdot jt} - \frac{1}{NT}\sum_{i,t} \bar{u}^2_{i\cdot t} + O_p(N^{-2} + M^{-2} + (NMT)^{-1})$,

(vii) $H_7 = (1 + \frac{2}{N} + \frac{2}{M} + \frac{2}{T})\frac{1}{NMT}\sum_{i,j,t} u^2_{ijt} - \frac{1}{MT}\sum_{j,t} \bar{u}^2_{\cdot jt} - \frac{1}{NT}\sum_{i,t} \bar{u}^2_{i\cdot t} - \frac{1}{NM}\sum_{i,j} \bar{u}^2_{ij\cdot} + O_p(N^{-2} + M^{-2} + T^{-2} + (NMT)^{-1})$.

---

[5]For Model 1, noting that $D_1 = \varnothing$, we implicitly define $d_{ijt,1} = 0$, $M_{D_1} = I_{NMT}$, and $X^*_{D_m} = (X'X)^{-1}$. In this case, $A_{ijt,1} = u_{ijt}$, $B_{ijt,1} = d^{*'}_{ijt}\pi^* - x'_{ijt}(X'X)^{-1} X'D^*\pi^*$, and $C_{ijt,1} = -x'_{ijt}(X'X)^{-1} X'U$.

Lemma A.11 Let $G_m = \frac{2}{NMT} \sum_{i,j,t} c_{ijt,m}^2 (A_{ijt,m} B_{ijt,m} + B_{ijt,m} C_{ijt,m})$ and $K_{m^*,m} = \frac{1}{NMT} \sum_{i,j,t} c_{ijt,m}^2$ $\times B_{ijt,m}^2$ for $m, m^* = 1, 2, ..., 7$, where model m^* is the true model and model m is a fitted model. Suppose that Assumptions A.1 and A.4 hold. If model m is under-fitted with respect to the true model m^*, then

(i) $G_m = o_p(1)$,

(ii) $K_{m^*,m} \xrightarrow{p} \varphi_{m^*,m}$.

Lemma A.12 Let $A = (a_{ijt})$ and $B = (b_{ijt})$ be either X or U. Suppose that Assumptions A.1-A.3 hold. Then

(i) $\frac{1}{NMT} A' (M_{D_2} - M_{D_4}) B = \frac{1}{NM} \sum_{i,j} (\overline{a}_{ij\cdot} - \overline{a}) (\overline{b}_{ij\cdot} - \overline{b})' + O_p((NT)^{-1} + (MT)^{-1}) = O_p(T^{-1})$,

(ii) $\frac{1}{NMT} A' (M_{D_3} - M_{D_4}) B = \frac{1}{T} \sum_t (\overline{a}_{\cdot\cdot t} - \overline{a}) (\overline{b}_{\cdot\cdot t} - \overline{b}\cdot)' = O_p((NM)^{-1})$,

(iii) $\frac{1}{NMT} A' (M_{D_3} - M_{D_7}) B = \frac{1}{NT} \sum_{i,j} (\overline{a}_{i\cdot t} - \overline{a})(\overline{b}_{i\cdot t} - \overline{b})' + \frac{1}{MT} \sum_{j,t} (\overline{a}_{\cdot jt} - \overline{a}) (\overline{b}_{\cdot jt} - \overline{b})' + O_p((NM)^{-1} + (MT)^{-1} + (NT)^{-1}) = O_p(M^{-1} + N^{-1})$,

(iv) $\frac{1}{NMT} A' (M_{D_4} - M_{D_7}) B = \frac{1}{NT} \sum_{i,t} (\overline{a}_{i\cdot t} - \overline{a})(\overline{b}_{i\cdot t} - \overline{b})' + \frac{1}{MT} \sum_{j,t} (\overline{a}_{\cdot jt} - \overline{a}) (\overline{b}_{\cdot jt} - \overline{b})' + O_p((NM)^{-1} + (MT)^{-1} + (NT)^{-1}) = O_p(M^{-1} + N^{-1})$,

(v) $\frac{1}{NMT} A' (M_{D_5} - M_{D_6}) B = \frac{1}{NT} \sum_{i,t} (\overline{a}_{i\cdot t} - \overline{a})(\overline{b}_{i\cdot t} - \overline{b})' + O_p((NM)^{-1}) = O_p(M^{-1})$,

(vi) $\frac{1}{NMT} A' (M_{D_5} - M_{D_7}) B = \frac{1}{NT} \sum_{i,t} (\overline{a}_{i\cdot t} - \overline{a})(\overline{b}_{i\cdot t} - \overline{b})' + \frac{1}{NM} \sum_{i,j} (\overline{a}_{ij\cdot} - \overline{a})(\overline{b}_{ij\cdot} - \overline{b})' + O_p((NM)^{-1} + (MT)^{-1} + (NT)^{-1}) = O_p(M^{-1} + T^{-1})$,

(vii) $\frac{1}{NMT} A' (M_{D_6} - M_{D_7}) B = \frac{1}{NM} \sum_{i,j} (\overline{a}_{ij\cdot} - \overline{a})(\overline{b}_{ij\cdot} - \overline{b})' + O_p((MT)^{-1} + (NT)^{-1}) = O_p(T^{-1})$.

Lemma A.13 Suppose that Assumptions A.1-A.3 hold. Then

(i) If Model 2 is the true model, $\frac{1}{NMT} \sum_{i,j,t} h_{ijt,4}^* \left| e_{ijt,4}^2 - e_{ijt,2}^2 \right| = o_p(T^{-1})$,

(ii) If Model 3 is the true model, $\frac{1}{NMT} \sum_{i,j,t} h_{ijt,4}^* \left| e_{ijt,4}^2 - e_{ijt,3}^2 \right| = o_p((NM)^{-1})$,

(iii) If Model 3 is the true model, $\frac{1}{NMT} \sum_{i,j,t} h_{ijt,7}^* \left| e_{ijt,7}^2 - e_{ijt,3}^2 \right| = o_p(N^{-1} + M^{-1})$,

(iv) If Model 4 is the true model, $\frac{1}{NMT} \sum_{i,j,t} h_{ijt,7}^* \left| e_{ijt,7}^2 - e_{ijt,4}^2 \right| = o_p(N^{-1} + M^{-1})$,

(v) If Model 5 is the true model, $\frac{1}{NMT} \sum_{i,j,t} h_{ijt,6}^* \left| e_{ijt,6}^2 - e_{ijt,5}^2 \right| = o_p(M^{-1})$,

(vi) If Model 5 is the true model, $\frac{1}{NMT} \sum_{i,j,t} h_{ijt,7}^* \left| e_{ijt,7}^2 - e_{ijt,5}^2 \right| = o_p(M^{-1} + T^{-1})$,

(vii) If Model 6 is the true model, $\frac{1}{NMT} \sum_{i,j,t} h_{ijt,7}^* \left| e_{ijt,7}^2 - e_{ijt,6}^2 \right| = o_p(T^{-1})$.

Lemma A.14 Suppose that Assumptions A.1-A.3 hold. Let $\bar{D}_{ijt,m} = D_m (D_m' D_m)^{-1} d_{ijt,m}$ for $m = 1, 2, ..., 7$. Then

(i) If Model 2 is the true model, $L_{2,4} \equiv \frac{1}{NMT} \sum_{i,j,t} \left| h_{ijt,4}^* - h_{ijt,2}^* \right| e_{ijt,2}^2 = o_p(T^{-1})$,

(ii) If Model 3 is the true model, $L_{3,4} \equiv \frac{1}{NMT} \sum_{i,j,t} \left| h_{ijt,4}^* - h_{ijt,3}^* \right| e_{ijt,3}^2 = o_p((NM)^{-1})$,

(iii) If Model 3 is the true model, $L_{3,7} \equiv \frac{1}{NMT} \sum_{i,j,t} \left| h_{ijt,7}^* - h_{ijt,3}^* \right| e_{ijt,3}^2 = o_p(N^{-1} + M^{-1})$,

(iv) If Model 4 is the true model, $L_{4,7} \equiv \frac{1}{NMT} \sum_{i,j,t} \left| h_{ijt,7}^* - h_{ijt,4}^* \right| e_{ijt,4}^2 = o_p(N^{-1} + M^{-1})$,

(v) If Model 5 is the true model, $L_{5,6} \equiv \frac{1}{NMT} \sum_{i,j,t} \left| h_{ijt,6}^* - h_{ijt,5}^* \right| e_{ijt,5}^2 = o_p(M^{-1})$,

(vi) If Model 5 is the true model, $L_{5,7} \equiv \frac{1}{NMT} \sum_{i,j,t} \left| h_{ijt,7}^* - h_{ijt,5}^* \right| e_{ijt,5}^2 = o_p(M^{-1} + T^{-1})$,

(vii) If Model 6 is the true model, $L_{6,7} \equiv \frac{1}{NMT} \sum_{i,j,t} \left| h_{ijt,7}^* - h_{ijt,6}^* \right| e_{ijt,6}^2 = o_p(T^{-1})$.

Proof of Theorem 3.1. We use $CV_{m^*,m}$ to denote $CV(m)$ when Model m^* is the true model. By Lemma A.7, $CV_{m^*,m} = \frac{1}{NMT} \sum_{i,j,t} c_{ijt,m}^2 e_{ijt,m}^2 = \frac{1}{NMT} \sum_{i,j,t} c_{ijt,m}^2 (A_{ijt,m}^2 + B_{ijt,m}^2 + C_{ijt,m}^2 + 2A_{ijt,m} B_{ijt,m} + 2A_{ijt,m} C_{ijt,m} + 2B_{ijt,m} C_{ijt,m})$, where $e_{ijt,m} = A_{ijt,m} + B_{ijt,m} + C_{ijt,m}$. When Model m is just- or over-fitted with respect to the true Model m^*, we have, by Lemmas A.8-A.10,

$$CV_{m^*,m} = E_m + F_m + H_m \xrightarrow{p} \overline{\sigma}_u^2.$$

We will show that in this case $\kappa_{m^*,m} (CV_{m^*,m} - CV_{m^*,m^*}) \to$ constant> 0 as long as $m \neq m^*$, where $\kappa_{m^*,m} = \kappa_{m^*,m} (N, M, T) \to \infty$ as $(N, M, T) \to \infty$ and it depends on the underlying true model (Model m^*) and the fitted model (Model m).

On the other hand, when Model m is under-fitted with respect to Model m^*, by Lemmas A.8 and A.10-A.11 we have

$$CV_{m^*,m} = E_m + G_m + H_m + K_{m^*,m},$$

where $E_m + G_m = o_p(1)$ and $H_m \xrightarrow{p} \overline{\sigma}_u^2$ for any m, and $K_{m^*,m^*} > 0$. It follows that

$$CV_{m^*,m} - CV_{m^*,m^*} \to \lim_{(N,M,T)\to\infty} K_{m^*,m} > 0.$$

The details are given below.

Case 1: Model 1 is the true model. In this case, Models 2-7 are all over-fitted and we have by Lemmas A.8-A.10

$$
\begin{aligned}
CV_{1,1} &= H_1 + O_p((NMT)^{-1}), \\
CV_{1,2} &= H_2 + O_p((NM)^{-2} + (NT)^{-2} + (MT)^{-2} + (NMT)^{-1}), \\
CV_{1,3} &= H_3 + O_p(T^{-2} + (NMT)^{-1}), \\
CV_{1,4} &= H_4 + O_p((NM)^{-2} + T^{-2} + (NMT)^{-1}), \\
CV_{1,5} &= H_5 + O_p(N^{-2} + (NMT)^{-1}), \\
CV_{1,6} &= H_6 + O_p(N^{-2} + M^{-2} + (NMT)^{-1}), \\
CV_{1,7} &= H_7 + O_p(N^{-2} + M^{-2} + T^{-2} + (NMT)^{-1}).
\end{aligned}
$$

Subcase 1a. For $CV_{1,2} - CV_{1,1}$, we have

$$
\begin{aligned}
CV_{1,2} - CV_{1,1} = {} & 2\left(\frac{1}{NT} + \frac{1}{MT} + \frac{1}{NM}\right)\frac{1}{NMT}\sum_{i,j,t} u_{ijt}^2 - \left(\frac{1}{N}\sum_i \overline{u}_{i..}^2 + \frac{1}{M}\sum_j \overline{u}_{.j.}^2 + \frac{1}{T}\sum_t \overline{u}_{..t}^2\right) \\
& + O_p((NM)^{-2} + (NT)^{-2} + (MT)^{-2} + (NMT)^{-1}),
\end{aligned}
$$

and

$$(NM \wedge NT \wedge MT)(CV_{1,2} - CV_{1,1}) \xrightarrow{p} q_1(2\overline{\sigma}_u^2 - \overline{\sigma}_{u1}^2) + q_2(2\overline{\sigma}_u^2 - \overline{\sigma}_{u2}^2) + q_3(2\overline{\sigma}_u^2 - \overline{\sigma}_{u3}^2),$$

where $q_1 = \lim_{(N,M,T)\to\infty}(1 \wedge \frac{N}{T} \wedge \frac{N}{M})$, $q_2 = \lim_{(N,M,T)\to\infty}(1 \wedge \frac{M}{T} \wedge \frac{M}{N})$, and $q_3 = \lim_{(N,M,T)\to\infty}(1 \wedge \frac{T}{N} \wedge \frac{T}{M})$.

Subcase 1b. For $CV_{1,3} - CV_{1,1}$, we have

$$T(CV_{1,3} - CV_{1,1}) = \frac{2}{NMT}\sum_{i,j,t} u_{ijt}^2 - \frac{T}{NM}\sum_{i,j}\overline{u}_{ij.}^2 + o_p(1) \xrightarrow{p} 2\overline{\sigma}_u^2 - \overline{\sigma}_{u4}^2.$$

Subcase 1c. For $CV_{1,4} - CV_{1,1}$, we have

$$
\begin{aligned}
(NM \wedge T)(CV_{1,4} - CV_{1,1}) = {} & (NM \wedge T)(\frac{1}{T} + \frac{1}{NM})\frac{2}{NMT}\sum_{i,j,t} u_{ijt}^2 - (NM \wedge T)\frac{1}{NM}\sum_{i,j}\overline{u}_{ij.}^2 \\
& - (NM \wedge T)\frac{1}{T}\sum_t \overline{u}_{..t}^2 + o_p(1) \\
& \xrightarrow{p} q_4(2\overline{\sigma}_u^2 - \overline{\sigma}_{u4}^2) + q_5(2\overline{\sigma}_u^2 - \overline{\sigma}_{u3}^2),
\end{aligned}
$$

where $q_4 = \lim_{(N,M,T)\to\infty}(1 \wedge \frac{T}{NM})$ and $q_5 = \lim_{(N,M,T)\to\infty}(1 \wedge \frac{NM}{T})$.

Subcase 1d. For $CV_{1,5} - CV_{1,1}$, we have

$$N(CV_{1,5} - CV_{1,1}) = \frac{2}{NMT}\sum_{i,j,t} u_{ijt}^2 - \frac{N}{MT}\sum_{j,t}\overline{u}_{.jt}^2 \xrightarrow{p} 2\overline{\sigma}_u^2 - \overline{\sigma}_{u6}^2.$$

Subcase 1e. For $CV_{1,6} - CV_{1,1}$, we have

$$
\begin{aligned}
(N \wedge M)(CV_{1,6} - CV_{1,1}) = {} & ((1 \wedge \frac{M}{N}) + (1 \wedge \frac{N}{M}))\frac{2}{NMT}\sum_{i,j,t} u_{ijt}^2 - (1 \wedge \frac{M}{N})\frac{N}{MT}\sum_{j,t}\overline{u}_{.jt}^2 \\
& - (1 \wedge \frac{N}{M})\frac{M}{NT}\sum_{i,t}\overline{u}_{i.t}^2 + o_p(1) \\
& \xrightarrow{p} q_6(2\overline{\sigma}_u^2 - \overline{\sigma}_{u6}^2) + q_7(2\overline{\sigma}_u^2 - \overline{\sigma}_{u5}^2),
\end{aligned}
$$

where $q_6 = \lim_{(N,M)\to\infty}(1 \wedge \frac{M}{N})$, and $q_7 = \lim_{(N,M)\to\infty}(1 \wedge \frac{N}{M})$.

Subcase 1f. For $CV_{1,7} - CV_{1,1}$, we have

$$
\begin{aligned}
(N \wedge M \wedge T)(CV_{1,7} - CV_{1,1}) &= (1 \wedge \frac{N}{T} \wedge \frac{M}{T})\left(\frac{2}{NMT}\sum_{i,j,t} u_{ijt}^2 - \frac{T}{NM}\sum_{i,j}\overline{u}_{ij\cdot}^2\right) \\
&\quad + (1 \wedge \frac{N}{M} \wedge \frac{T}{M})\left(\frac{2}{NMT}\sum_{i,j,t} u_{ijt}^2 - \frac{M}{NT}\sum_{i,t}\overline{u}_{i\cdot t}^2\right) \\
&\quad + (1 \wedge \frac{M}{N} \wedge \frac{T}{N})\left(\frac{2}{NMT}\sum_{i,j,t} u_{ijt}^2 - \frac{N}{MT}\sum_{j,t}\overline{u}_{\cdot jt}^2\right) + o_p(1) \\
&\xrightarrow{p} q_8(2\overline{\sigma}_u^2 - \overline{\sigma}_{u4}^2) + q_9(2\overline{\sigma}_u^2 - \overline{\sigma}_{u5}^2) + q_{10}(2\overline{\sigma}_u^2 - \overline{\sigma}_{u6}^2),
\end{aligned}
$$

where $q_8 = \lim_{(NMT)\to\infty}(1 \wedge \frac{N}{T} \wedge \frac{M}{T})$, $q_9 = \lim_{(NMT)\to\infty}(1 \wedge \frac{N}{M} \wedge \frac{T}{M})$ and $q_{10} = \lim_{(NMT)\to\infty}(1 \wedge \frac{M}{N} \wedge \frac{T}{N})$.

Case 2: Model 2 is the true model. In this case, Models 4, 6 and 7 are over-fitted and Models 1, 3 and 5 are under-fitted. By Lemmas A.8-A.11, we have

$$
\begin{aligned}
CV_{2,m} &= H_m + K_{2,m} + o_p(1) \text{ for } m = 1, 3, 5, \\
CV_{2,2} &= H_2 + O_p((NM)^{-2} + (NT)^{-2} + (MT)^{-2} + (NMT)^{-1}), \\
CV_{2,4} &= H_4 + O_p((NM)^{-2} + T^{-2} + (NMT)^{-1}), \\
CV_{2,6} &= H_6 + O_p(N^{-2} + M^{-2} + (NMT)^{-1}), \\
CV_{2,7} &= H_7 + O_p(N^{-2} + M^{-2} + T^{-2} + (NMT)^{-1}).
\end{aligned}
$$

For the under-fitted models, we have $CV_{2,m} - CV_{2,2} \xrightarrow{p} \varphi_{2,m} > 0$ where $m = 1, 3, 5$. For the over-fitted cases, we have to be careful in the discussion.

Subcase 2a. For $CV_{2,4} - CV_{2,2}$, if we assume that N, M, and T pass to infinity at the restrictive rates such that $T/(NM)^2 \to 0$,[6] then analogously to Case 1, we can easily show that these conditions will ensure $T(CV_{2,4} - CV_{2,2}) = T(H_4 - H_2) + o_p(1) \xrightarrow{p} 2\overline{\sigma}_u^2 - \overline{\sigma}_{u4}^2$. But as emphasized in the text, we do not want to impose such a rate condition. In this case, we need to keep track of all terms in the expression of $CV_{2,4}$ and $CV_{2,2}$ that are not $o_p(T^{-1})$. To unify notation, we make the following decomposition

$$
\begin{aligned}
CV_{2,4} - CV_{2,2} &= \frac{1}{NMT}\sum_{i,j,t}(c_{ijt,4}^2 e_{ijt,4}^2 - c_{ijt,2}^2 e_{ijt,2}^2) \\
&= \frac{1}{NMT}\sum_{i,j,t}\left[(e_{ijt,4}^2 - e_{ijt,2}^2) + (c_{ijt,4}^2 - 1)(e_{ijt,4}^2 - e_{ijt,2}^2) + (c_{ijt,4}^2 - c_{ijt,2}^2)e_{ijt,2}^2\right] \\
&\equiv L_1^{(2,4)} + L_2^{(2,4)} + L_3^{(2,4)}, \text{ say.}
\end{aligned}
$$

Note that

$$
\begin{aligned}
L_1^{(2,4)} &= \frac{1}{NMT}U'(M_{D_4} - M_{D_2})U + \frac{1}{NMT}(U'M_{D_2}XX_{D_2}^*X'M_{D_2}U - U'M_{D_4}XX_{D_4}^*X'M_{D_4}U) \\
&\equiv L_{1,1}^{(2,4)} + L_{1,2}^{(2,4)}, \text{ say,}
\end{aligned}
$$

---

[6] Admittedly, this rate requirement does not appear very restrictive and looks quite reasonable in many applications.

where we recall $X_{D_m}^* = (X'M_{D_m}X)^{-1}$. $L_{1,1}^{(2,4)} = -\frac{1}{NM}\sum_{i,j}\overline{u}_{ij.}^2 + O_p((MT)^{-1} + (NT)^{-1}) = O_p(T^{-1})$ by Lemmas A.12(i). For $L_{1,2}^{(2,4)}$, we make further decomposition

$$
\begin{aligned}
L_{1,2}^{(2,4)} &= \frac{1}{NMT}U'(M_{D_2} - M_{D_4})XX_{D_4}^*X'(M_{D_2} - M_{D_4})U + \frac{2}{NMT}U'(M_{D_2} - M_{D_4})XX_{D_4}^*X'M_{D_4}U \\
&\quad + \frac{1}{NMT}U'M_{D_2}XX_{D_2}^*X'(M_{D_4} - M_{D_2})XX_{D_4}^*X'M_{D_2}U \\
&\equiv L_{1,2,1}^{(2,4)} + 2L_{1,2,2}^{(2,4)} + L_{1,2,3}^{(2,4)}, \text{ say.}
\end{aligned}
$$

By Lemmas A.4(i) and (iii) and A.12(i), and Assumption A.1(v), $L_{1,2,1}^{(2,4)} = O_p(T^{-2})$, $L_{1,2,2}^{(2,4)} = O_p(T^{-1})O_p(T^{-1} + (NM)^{-1} + (NMT)^{-1/2})$, and $L_{1,2,3}^{(2,4)} = O_p((NM)^{-2} + (MT)^{-2} + (NT)^{-2} + (NMT)^{-1})O_p(T^{-1})$. It follows that $L_{1,2}^{(2,4)} = o_p(T^{-1})$ and $L_1^{(2,4)} = -\frac{1}{NM}\sum_{i,j}\overline{u}_{ij.}^2 + o_p(T^{-1}) = O_p(T^{-1})$.

For $L_2^{(2,4)}$, we use the fact that $h_{ijt,m} = \overline{d}_m + h_{ijt,m}^*$ by Lemma A.6(i). Let $\overline{c}_m = \frac{1}{1-\overline{d}_m}$. Then

$$
c_{ijt,m}^2 - \overline{c}_m^2 = \frac{1}{(1 - \overline{d}_m - h_{ijt,m}^*)^2} - \frac{1}{(1 - \overline{d}_m)^2} = \overline{c}_m^2 h_{ijt,m}^* r_{ijt,m},
$$

where $r_{ijt,m} = (2 - \overline{d}_m - h_{ijt,m}^*)/(1 - \overline{d}_m - h_{ijt,m}^*)^2$. Noting that $h_{ijt,m}^*$ is $O_p(\overline{d}_m)$ as shown in the proof of Lemma A.6(ii). One can readily show that $\max_{i,j,t,m}|r_{ijt,m}| = 2 + o_p(1)$. We make the following decomposition

$$
L_2^{(2,4)} = \frac{\overline{c}_4^2 - 1}{NMT}\sum_{i,j,t}(e_{ijt,4}^2 - e_{ijt,2}^2) + \frac{1}{NMT}\sum_{i,j,t}(c_{ijt,4}^2 - \overline{c}_4^2)(e_{ijt,4}^2 - e_{ijt,2}^2) \equiv L_{2,1}^{(2,4)} + L_{2,2}^{(2,4)}, \text{ say.}
$$

$L_{2,1}^{(2,4)} = (\overline{c}_4^2 - 1)L_1^{(2,4)} = o(1)L_1^{(2,4)} = o_p(T^{-1})$. For $L_{2,2}^{(2,4)}$, we apply Lemma A.13(i) to obtain

$$
\left|L_{2,2}^{(2,4)}\right| = \left|\frac{\overline{c}_4^2}{NMT}\sum_{i,j,t}h_{ijt,4}^* r_{ijt,4}(e_{ijt,4}^2 - e_{ijt,2}^2)\right| \leq \frac{\overline{c}_4^2 \max_{i,j,t}|r_{ijt,4}|}{NMT}\sum_{i,j,t}h_{ijt,4}^*|e_{ijt,4}^2 - e_{ijt,2}^2| = o_p(T^{-1}).
$$

For $L_3^{(2,4)}$, we have

$$
\begin{aligned}
L_3^{(2,4)} &= \frac{1}{NMT}\sum_{i,j,t}(c_{ijt,4}^2 - c_{ijt,2}^2)e_{ijt,2}^2 \\
&= \frac{1}{NMT}\sum_{i,j,t}2(h_{ijt,4} - h_{ijt,2})e_{ijt,2}^2 + \frac{1}{NMT}\sum_{i,j,t}(c_{ijt,4}^2 - c_{ijt,2}^2 - 2h_{ijt,4} + 2h_{ijt,2})e_{ijt,2}^2 \\
&\equiv L_{3,1}^{(2,4)} + L_{3,2}^{(2,4)}, \text{ say.}
\end{aligned}
$$

For $L_{3,1}^{(2,4)}$, we apply Lemmas A.5, A.6 and A.14(i) to obtain

$$
\begin{aligned}
L_{3,1}^{(2,4)} &= 2[T^{-1} - (MT)^{-1} + (NT)^{-1}]\frac{1}{NMT}\sum_{i,j,t}e_{ijt,2}^2 + \frac{2}{NMT}\sum_{i,j,t}(h_{ijt,4}^* - h_{ijt,2}^*)e_{ijt,2}^2 \\
&= 2[T^{-1} - (MT)^{-1} + (NT)^{-1}]\frac{1}{NMT}\sum_{i,j,t}e_{ijt,2}^2 + o_p(T^{-1}) \\
&= \frac{2}{T}\frac{1}{NMT}\sum_{i,j,t}u_{ijt}^2 + o_p(T^{-1}),
\end{aligned}
$$

as we can readily show that $\frac{1}{NMT}\sum_{i,j,t} e_{ijt,m}^2 = \frac{1}{NMT}\sum_{i,j,t} u_{ijt}^2 + o_p(1)$ when Model m is the true model. For $L_{3,2}^{(2,4)}$, we can apply Lemmas A.5-A.6 and A.14(i) to show that

$$
\begin{aligned}
\left| L_{3,2}^{(2,4)} \right| &= \frac{1}{NMT}\sum_{i,j,t}(c_{ijt,4}^2 - c_{ijt,2}^2 - 2h_{ijt,4} + 2h_{ijt,2})e_{ijt,2}^2 = \frac{o_p(1)}{NMT}\sum_{i,j,t}|h_{ijt,4} - h_{ijt,2}|\, e_{ijt,2}^2 \\
&\leq o_p(1)\left[\frac{|\bar{d}_4 - \bar{d}_2|}{NMT}\sum_{i,j,t} e_{ijt,2}^2 + \frac{1}{NMT}\sum_{i,j,t}|h_{ijt,4}^* - h_{ijt,2}^*|e_{ijt,2}^2 \right] \\
&= o_p(1)\left[O_p(T^{-1}) + o_p(T^{-1}) \right] = o_p(T^{-1}).
\end{aligned}
$$

It follows that $L_3^{(2,4)} = 2T^{-1}\frac{1}{NMT}\sum_{i,j,t} u_{ijt}^2 + o_p(T^{-1})$. In sum, we have proved that

$$
CV_{2,4} - CV_{2,2} = 2T^{-1}\frac{1}{NMT}\sum_{i,j,t} u_{ijt}^2 - \frac{1}{NM}\sum_{i,j}\overline{u}_{ij\cdot}^2 + o_p(T^{-1})
$$

and then $T\left(CV_{2,4} - CV_{2,2}\right) \xrightarrow{p} 2\overline{\sigma}_u^2 - \overline{\sigma}_{u4}^2$ by Assumptions A.1(i) and A.2(iv).

Subcase 2b. For $CV_{2,6} - CV_{2,2}$, noting that

$$
\begin{aligned}
CV_{2,6} - CV_{2,2} &= H_6 - H_2 + o_p(N^{-1} + M^{-1}) \\
&= \left(\frac{1}{N} + \frac{1}{M}\right)\frac{2}{NMT}\sum_{i,j,t} u_{ijt}^2 - \frac{1}{MT}\sum_{j,t}\overline{u}_{\cdot jt}^2 - \frac{1}{NT}\sum_{i,t}\overline{u}_{i\cdot t}^2 + o_p(N^{-1} + M^{-1}),
\end{aligned}
$$

as in Subcase 1e, we have

$$
(N \wedge M)\left(CV_{2,6} - CV_{2,2}\right) \xrightarrow{p} q_6(2\overline{\sigma}_u^2 - \overline{\sigma}_{u6}^2) + q_7(2\overline{\sigma}_u^2 - \overline{\sigma}_{u5}^2).
$$

Subcase 2c. For $CV_{2,7} - CV_{2,2}$, noting that

$$
\begin{aligned}
CV_{2,7} - CV_{2,2} &= H_7 - H_2 + o_p(N^{-1} + M^{-1} + T^{-1}) \\
&= \left(\frac{1}{N} + \frac{1}{M} + \frac{1}{T}\right)\frac{2}{NMT}\sum_{i,j,t} u_{ijt}^2 - \frac{1}{NM}\sum_{i,j}\overline{u}_{ij\cdot}^2 - \frac{1}{NT}\sum_{i,t}\overline{u}_{i\cdot t}^2 - \frac{1}{MT}\sum_{j,t}\overline{u}_{\cdot jt}^2 \\
&\quad + o_p(N^{-1} + M^{-1} + T^{-1}),
\end{aligned}
$$

as in Subcase 1f we have

$$
(N \wedge M \wedge T)(CV_{2,7} - CV_{2,2}) \xrightarrow{p} q_8(2\overline{\sigma}_u^2 - \overline{\sigma}_{u4}^2) + q_9(2\overline{\sigma}_u^2 - \overline{\sigma}_{u5}^2) + q_{10}(2\overline{\sigma}_u^2 - \overline{\sigma}_{u6}^2).
$$

Case 3: Model 3 is the true model. In this case, Models 1, 2, 5, and 6 are under-fitted and Models 4 and 7 are over-fitted. By Lemmas A.8-A.11, we have

$$
\begin{aligned}
CV_{3,m} &= H_m + K_{3,m} + o_p(1) \text{ for } m = 1, 2, 5, 6, \\
CV_{3,3} &= H_3 + O_p(T^{-2} + (NMT)^{-1}), \\
CV_{3,4} &= H_4 + O_p((NM)^{-2} + T^{-2} + (NMT)^{-1}), \\
CV_{3,7} &= H_7 + O_p(N^{-2} + M^{-2} + T^{-2} + (NMT)^{-1}).
\end{aligned}
$$

For the under-fitted cases, by Lemma A.8 and A.10-A.11 we have

$$
CV_{3,m} - CV_{3,3} \xrightarrow{p} \varphi_{3,m} > 0 \text{ for } m = 1, 2, 5, 6.
$$

We study the over-fitted cases in order.

Subcase 3a. For $CV_{3,4} - CV_{3,3}$, the proof parallels to the analysis of $CV_{2,4} - CV_{2,2}$ and we only sketch the main steps. Note that $CV_{3,4} - CV_{3,3} = \frac{1}{NMT} \sum_{i,j,t} (c_{ijt,4}^2 e_{ijt,4}^2 - c_{ijt,3}^2 e_{ijt,3}^2) \equiv L_1^{(3,4)} + L_2^{(3,4)} + L_3^{(3,4)}$, where $L_\ell^{(3,4)}$ is defined analogously to $L_\ell^{(2,4)}$ for $\ell = 1, 2, 3$. For $L_1^{(3,4)}$, we have by Lemmas A.4(ii)-(iii) and A.12(ii)

$$
\begin{aligned}
L_1^{(3,4)} &= \frac{1}{NMT} U'(M_{D_4} - M_{D_3}) U + \frac{1}{NMT} (U' M_{D_3} X X_{D_3}^* X' M_{D_3} U - U' M_{D_4} X X_{D_4}^* X' M_{D_4} U) \\
&= -\frac{1}{T} \sum_t (\overline{u}_{..t} - \overline{u})^2 + o_p((NM)^{-1}) = -\frac{1}{T} \sum_t \overline{u}_{..t}^2 + o_p((NM)^{-1}) = O_p((NM)^{-1}).
\end{aligned}
$$

For $L_2^{(3,4)}$, we follow the analysis of $L_2^{(2,4)}$ and consider the following decomposition

$$
L_2^{(3,4)} = \frac{\overline{c}_4^2 - 1}{NMT} \sum_{i,j,t} (e_{ijt,4}^2 - e_{ijt,3}^2) + \frac{1}{NMT} \sum_{i,j,t} (c_{ijt,4}^2 - \overline{c}_4^2)(e_{ijt,4}^2 - e_{ijt,3}^2) \equiv L_{2,1}^{(3,4)} + L_{2,2}^{(3,4)}, \text{ say.}
$$

$L_{2,1}^{(3,4)} = (\overline{c}_4^2 - 1) L_1^{(3,4)} = o(1) L_1^{(3,4)} = o_p((NM)^{-1})$. For $L_{2,2}^{(3,4)}$, we apply Lemma A.13

$$
\left| L_{2,2}^{(3,4)} \right| = \left| \frac{\overline{c}_4^2}{NMT} \sum_{i,j,t} h_{ijt,4}^* r_{ijt,4} (e_{ijt,4}^2 - e_{ijt,3}^2) \right| \le \frac{\overline{c}_4^2 \max_{i,j,t} |r_{ijt,4}|}{NMT} \sum_{i,j,t} h_{ijt,4}^* |e_{ijt,4}^2 - e_{ijt,3}^2| = o_p((NM)^{-1}).
$$

Next, we apply Lemmas A.5, A.6 and A.14(ii) to obtain

$$
\begin{aligned}
L_3^{(3,4)} &= \frac{1 + o_p(1)}{NMT} \sum_{i,j,t} 2 (h_{ijt,4} - h_{ijt,3}) e_{ijt,3}^2 \\
&= 2[(NM)^{-1} - (NMT)^{-1}] \frac{1 + o_p(1)}{NMT} \sum_{i,j,t} e_{ijt,3}^2 + \frac{2}{NMT} \sum_{i,j,t} (h_{ijt,4}^* - h_{ijt,3}^*) e_{ijt,3}^2 \\
&= \frac{2}{NM} \frac{1}{NMT} \sum_{i,j,t} u_{ijt}^2 + o_p((NM)^{-1}).
\end{aligned}
$$

In sum, we have proved $CV_{3,4} - CV_{3,3} = 2(NM)^{-1} \frac{1}{NMT} \sum_{i,j,t} u_{ijt}^2 - \frac{1}{T} \sum_{t=1}^T \overline{u}_{..t}^2 + o_p((NM)^{-1})$. Then $NM(CV_{3,4} - CV_{3,3}) \xrightarrow{p} 2\overline{\sigma}_u^2 - \overline{\sigma}_{u3}^2$ by Assumptions A.1(i) and A.2(iii).

Subcase 3b. For $CV_{3,7} - CV_{3,3}$, the proof parallels the analysis of $CV_{2,4} - CV_{2,2}$ and we only sketch the main steps. Note that $CV_{3,7} - CV_{3,3} = \frac{1}{NMT} \sum_{i,j,t} (c_{ijt,7}^2 e_{ijt,7}^2 - c_{ijt,3}^2 e_{ijt,3}^2) \equiv L_1^{(3,7)} + L_2^{(3,7)} + L_3^{(3,7)}$, where $L_\ell^{(3,7)}$ is defined analogously to $L_\ell^{(2,4)}$ for $\ell = 1, 2, 3$. For $L_1^{(3,7)}$, we have by Lemmas A.4(ii),(vi) and A.12(iii),

$$
\begin{aligned}
L_1^{(3,7)} &= \frac{1}{NMT} U'(M_{D_7} - M_{D_3}) U + \frac{1}{NMT} (U' M_{D_3} X X_{D_3}^* X' M_{D_3} U - U' M_{D_7} X X_{D_7}^* X' M_{D_7} U) \\
&= -\frac{1}{NT} \sum_{i,t} \overline{u}_{i\cdot t}^2 - \frac{1}{MT} \sum_{j,t} \overline{u}_{\cdot jt}^2 + o_p(M^{-1} + N^{-1}) = O_p(M^{-1} + N^{-1}).
\end{aligned}
$$

Following the analysis of $L_2^{(2,4)}$, we can show that $L_2^{(3,7)} = o_p(M^{-1} + N^{-1})$. For $L_3^{(3,7)}$, we can apply Lemmas A.5, A.6 and A.14(iii) to obtain

$$
\begin{aligned}
L_3^{(3,7)} &= \frac{1 + o_p(1)}{NMT} \sum_{i,j,t} 2 (h_{ijt,7} - h_{ijt,3}) e_{ijt,3}^2 \\
&= 2[N^{-1} + M^{-1} - (NM)^{-1} - (NT)^{-1} - (MT)^{-1} + (NMT)^{-1}] \frac{1 + o_p(1)}{NMT} \sum_{i,j,t} e_{ijt,3}^2 \\
&\quad + \frac{2}{NMT} \sum_{i,j,t} (h_{ijt,7}^* - h_{ijt,3}^*) e_{ijt,3}^2 \\
&= 2(N^{-1} + M^{-1}) \frac{1}{NMT} \sum_{i,j,t} u_{ijt}^2 + o_p(M^{-1} + N^{-1}).
\end{aligned}
$$

31

In sum, we have proved

$$CV_{3,7} - CV_{3,3} = 2\left(N^{-1} + M^{-1}\right)\frac{1}{NMT}\sum_{i,j,t}u_{ijt}^2 - \frac{1}{NT}\sum_{i,t}\overline{u}_{i\cdot t}^2 - \frac{1}{MT}\sum_{j,t}\overline{u}_{\cdot jt}^2 + o_p(M^{-1} + N^{-1})$$

and then $(N \wedge M)(CV_{3,7} - CV_{3,3}) \xrightarrow{p} q_6(2\overline{\sigma}_u^2 - \overline{\sigma}_{u6}^2) + q_7(2\overline{\sigma}_u^2 - \overline{\sigma}_{u5}^2)$ by Assumptions A.1(i) and A.2(v)-(vi).

Case 4: Model 4 is the true model. In this case, Models 1, 2, 3, 5 and 6 are under-fitted and Model 7 is over-fitted. By Lemmas A.8-A.11, we have

$$\begin{aligned}
CV_{4,m} &= H_m + K_{4,m} + o_p(1) \text{ for } m = 1,2,3,5,6, \\
CV_{4,4} &= H_4 + O_p((NM)^{-2} + T^{-2} + (NMT)^{-1}), \text{ and} \\
CV_{4,7} &= H_7 + O_p(N^{-2} + M^{-2} + T^{-2} + (NMT)^{-1}).
\end{aligned}$$

For the under-fitted cases, by Lemmas A.8 and A.10-A.11 we have

$$CV_{4,m} - CV_{4,4} \xrightarrow{p} \varphi_{4,m} > 0 \text{ for } m = 1,2,3,5,6.$$

For the over-fitted case, we need to study $CV_{4,7} - CV_{4,4}$. Note that $CV_{4,7} - CV_{4,4} = \frac{1}{NMT}\sum_{i,j,t}(c_{ijt,7}^2 e_{ijt,7}^2 - c_{ijt,4}^2 e_{ijt,4}^2) \equiv L_1^{(4,7)} + L_2^{(4,7)} + L_3^{(4,7)}$, where $L_\ell^{(4,7)}$ is defined analogously to $L_\ell^{(2,4)}$ for $\ell = 1,2,3$. For $L_1^{(4,7)}$, we have by Lemmas A.4(iii), (vi) and A.12(iv),

$$\begin{aligned}
L_1^{(4,7)} &= \frac{1}{NMT}U'(M_{D_7} - M_{D_4})U + \frac{1}{NMT}(U'M_{D_4}XX_{D_4}^*X'M_{D_4}U - U'M_{D_7}XX_{D_7}^*X'M_{D_7}U) \\
&= -\frac{1}{NT}\sum_{i,t}\overline{u}_{i\cdot t}^2 - \frac{1}{MT}\sum_{j,t}\overline{u}_{\cdot jt}^2 + o_p(M^{-1} + N^{-1}) = O_p(M^{-1} + N^{-1}).
\end{aligned}$$

Following the analysis of $L_2^{(2,3)}$, we can show that $L_2^{(4,7)} = o_p(M^{-1} + N^{-1})$. For $L_3^{(4,7)}$, we can apply Lemmas A.5, A.6 and A.14(iv) to obtain

$$L_3^{(4,7)} = \frac{1 + o_p(1)}{NMT}\sum_{i,j,t}2\left(h_{ijt,7} - h_{ijt,4}\right)e_{ijt,4}^2 = 2\left(N^{-1} + M^{-1}\right)\frac{1}{NMT}\sum_{i,j,t}u_{ijt}^2 + o_p\left(M^{-1} + N^{-1}\right).$$

In sum, we have proved

$$CV_{4,7} - CV_{4,4} = 2\left(N^{-1} + M^{-1}\right)\frac{1}{NMT}\sum_{i,j,t}u_{ijt}^2 - \frac{1}{NT}\sum_{i,t}\overline{u}_{i\cdot t}^2 - \frac{1}{MT}\sum_{j,t}\overline{u}_{\cdot jt}^2 + o_p(M^{-1} + N^{-1})$$

and then $(N \wedge M)(CV_{4,7} - CV_{4,4}) \xrightarrow{p} q_6(2\overline{\sigma}_u^2 - \overline{\sigma}_{u6}^2) + q_7(2\overline{\sigma}_u^2 - \overline{\sigma}_{u5}^2)$ by Assumptions A.1(i) and A.2(v)-(vi).

Case 5: Model 5 is the true model. In this cases, Models 1, 2, 3, and 4 are under-fitted and Models 6 and 7 are over-fitted. By Lemmas A.8-A.11, we have

$$\begin{aligned}
CV_{5,m} &= H_m + K_{5,m} + o_p(1), \text{ for } m = 1,2,3,4, \\
CV_{5,5} &= H_5 + O_p(N^{-2} + (NMT)^{-1}), \\
CV_{5,6} &= H_6 + O_p(N^{-2} + M^{-2} + (NMT)^{-1}), \\
CV_{5,7} &= H_7 + O_p(N^{-2} + M^{-2} + T^{-2} + (NMT)^{-1}).
\end{aligned}$$

For the under-fitted cases, by Lemmas A.8 and A.10-A.11 we have

$$CV_{5,m} - CV_{5,5} \xrightarrow{p} \varphi_{5,m} > 0 \text{ for } m = 1,2,3,4.$$

We study the two over-fitted cases in order.

Note that $CV_{5,6} - CV_{5,5} = \frac{1}{NMT}\sum_{i,j,t}(c_{ijt,6}^2 e_{ijt,6}^2 - c_{ijt,5}^2 e_{ijt,5}^2) \equiv L_1^{(5,6)} + L_2^{(5,6)} + L_3^{(5,6)}$, where $L_\ell^{(5,6)}$ is defined analogously to $L_\ell^{(2,4)}$ for $\ell = 1,2,3$. By Lemmas A.4(iv)-(v) and A.12(v),

$$
\begin{aligned}
L_1^{(5,6)} &= \frac{1}{NMT}U'\left(M_{D_6} - M_{D_5}\right)U + \frac{1}{NMT}\left(U'M_{D_5}XX_{D_5}^*X'M_{D_5}U - U'M_{D_6}XX_{D_6}^*X'M_{D_6}U\right) \\
&= -\frac{1}{NT}\sum_{i,t}\overline{u}_{i\cdot t}^2 + o_p(M^{-1}) = O_p(M^{-1}).
\end{aligned}
$$

Following the analysis of $L_2^{(2,4)}$, we can show that $L_2^{(5,6)} = o_p(M^{-1})$. For $L_3^{(5,6)}$, we can apply Lemmas A.5, A.6 and A.14(v) to obtain

$$
L_3^{(5,6)} = \frac{1 + o_p(1)}{NMT}\sum_{i,j,t} 2\left(h_{ijt,6} - h_{ijt,5}\right)e_{ijt,5}^2 = 2M^{-1}\frac{1}{NMT}\sum_{i,j,t}u_{ijt}^2 + o_p\left(M^{-1}\right).
$$

In sum, we have proved

$$
CV_{5,6} - CV_{5,5} = 2M^{-1}\frac{1}{NMT}\sum_{i,j,t}u_{ijt}^2 - \frac{1}{NT}\sum_{i,t}\overline{u}_{i\cdot t}^2 + o_p(M^{-1})
$$

and then $M\left(CV_{5,6} - CV_{5,5}\right) \xrightarrow{p} 2\overline{\sigma}_u^2 - \overline{\sigma}_{u5}^2$ by Assumptions A.1(i) and A.2(v).

Note that $CV_{5,7} - CV_{5,5} = \frac{1}{NMT}\sum_{i,j,t}(c_{ijt,7}^2 e_{ijt,7}^2 - c_{ijt,5}^2 e_{ijt,5}^2) \equiv L_1^{(5,7)} + L_2^{(5,7)} + L_3^{(5,7)}$, where $L_\ell^{(5,7)}$ is defined analogously to $L_\ell^{(2,4)}$ for $\ell = 1,2,3$. By Lemmas A.4(iv)-(v) and A.12(vi),

$$
\begin{aligned}
L_1^{(5,7)} &= \frac{1}{NMT}U'\left(M_{D_7} - M_{D_5}\right)U + \frac{1}{NMT}\left(U'M_{D_5}XX_{D_5}^*X'M_{D_5}U - U'M_{D_7}XX_{D_7}^*X'M_{D_7}U\right) \\
&= -\frac{1}{NT}\sum_{i,t}\overline{u}_{i\cdot t}^2 - \frac{1}{NM}\sum_{i,j}\overline{u}_{ij\cdot}^2 + o_p(M^{-1} + T^{-1}) = O_p(M^{-1} + T^{-1}).
\end{aligned}
$$

Following the analysis of $L_2^{(2,4)}$, we can show that $L_2^{(5,7)} = o_p(M^{-1} + T^{-1})$. For $L_3^{(5,7)}$, we can apply Lemmas A.5, A.6 and A.14(vi) to obtain

$$
L_3^{(5,7)} = \frac{1 + o_p(1)}{NMT}\sum_{i,j,t} 2\left(h_{ijt,7} - h_{ijt,5}\right)e_{ijt,5}^2 = 2(M^{-1} + T^{-1})\frac{1}{NMT}\sum_{i,j,t}u_{ijt}^2 + o_p\left(M^{-1} + T^{-1}\right).
$$

In sum, we have proved

$$
CV_{5,7} - CV_{5,5} = 2\left(M^{-1} + T^{-1}\right)\frac{1}{NMT}\sum_{i,j,t}u_{ijt}^2 - \frac{1}{NT}\sum_{i,t}\overline{u}_{i\cdot t}^2 - \frac{1}{NM}\sum_{i,j}\overline{u}_{ij\cdot}^2 + o_p(M^{-1} + T^{-1})
$$

and then $(M \wedge T)\left(CV_{5,7} - CV_{5,5}\right) \xrightarrow{p} q_{11}(2\overline{\sigma}_u^2 - \overline{\sigma}_{u4}^2) + q_{12}(2\overline{\sigma}_u^2 - \overline{\sigma}_{u5}^2)$ by Assumptions A.1(i) and A.2(iv)-(v), where $q_{11} = \lim_{(M,T)\to\infty} 1 \wedge \frac{M}{T}$ and $q_{12} = \lim_{(M,T)\to\infty} 1 \wedge \frac{T}{M}$.

Case 6: Model 6 is the true model. In this case, Models 1, 2, 3, 4 and 5 are under-fitted and Model 7 is over-fitted. By Lemmas A.8-A.11, we have

$$
\begin{aligned}
CV_{6,m} &= H_m + K_{6,m} + o_p(1) \text{ for } m = 1,2,3,4 \text{ and } 5, \\
CV_{6,6} &= H_6 + O_p(N^{-2} + M^{-2} + (NMT)^{-1}), \\
CV_{6,7} &= H_7 + O_p(N^{-2} + M^{-2} + T^{-2} + (NMT)^{-1}).
\end{aligned}
$$

For the under-fitted cases, by Lemmas A.8 and A.10-A.11 we have

$$
CV_{6,m} - CV_{6,6} \xrightarrow{p} \varphi_{6,m} > 0 \text{ for } m = 1,2,3,4 \text{ and } 5.
$$

33

For $CV_{6,7} - CV_{6,6}$, we have $CV_{6,7} - CV_{6,6} = \frac{1}{NMT} \sum_{i,j,t} (c_{ijt,7}^2 e_{ijt,7}^2 - c_{ijt,6}^2 e_{ijt,6}^2) \equiv L_1^{(6,7)} + L_2^{(6,7)} + L_3^{(6,7)}$, where $L_\ell^{(6,7)}$ is defined analogously to $L_\ell^{(2,4)}$ for $\ell = 1, 2, 3$. By Lemmas A.4(iv)-(v) and A.12(vii),

$$
\begin{aligned}
L_1^{(6,7)} &= \frac{1}{NMT} U' (M_{D_7} - M_{D_6}) U + \frac{1}{NMT} \left[U' M_{D_6} X X_{D_6}^* X' M_{D_6} U - U' M_{D_7} X X_{D_7}^* X' M_{D_7} U \right] \\
&= -\frac{1}{NM} \sum_{i,j} \overline{u}_{ij.}^2 + o_p(T^{-1}) = O_p(T^{-1}).
\end{aligned}
$$

Following the analysis of $L_2^{(2,4)}$, we can show that $L_2^{(6,7)} = o_p(T^{-1})$. For $L_3^{(6,7)}$, we can apply Lemmas A.5, A.6 and A.14(vii) to obtain

$$
L_3^{(6,7)} = \frac{1 + o_p(1)}{NMT} \sum_{i,j,t} 2 \left(h_{ijt,7} - h_{ijt,6} \right) e_{ijt,6}^2 = 2T^{-1} \frac{1}{NMT} \sum_{i,j,t} u_{ijt}^2 + o_p\left(T^{-1}\right).
$$

In sum, we have proved

$$
CV_{6,7} - CV_{6,6} = 2T^{-1} \frac{1}{NMT} \sum_{i,j,t} u_{ijt}^2 - \frac{1}{NM} \sum_{i,j} \overline{u}_{ij.}^2 + o_p(T^{-1})
$$

and then $T(CV_{6,7} - CV_{6,6}) \xrightarrow{p} 2\overline{\sigma}_u^2 - \overline{\sigma}_{u4}^2$ by Assumptions A.1(i) and A.2(iv).

Case 7: Model 7 is the true model. In this case, Models 1-6 are all under-fitted. By Lemmas A.8-A.11, we have

$$
CV_{7,m} - CV_{7,7} \xrightarrow{p} \varphi_{7,m} > 0 \text{ for } m = 1, 2, ..., 6.
$$

This completes the proof of the theorem. ∎

Table 3A: Frequency of the model selected: static panels, $\rho = 0$

	(N,M,T)	True model: M1 Selected models							True model: M2 Selected models							True model: M3 Selected models							True model: M4 Selected models						
		M1	M2	M3	M4	M5	M6	M7	M1	**M2**	M3	M4	M5	M6	M7	M1	M2	**M3**	M4	M5	M6	M7	M1	M2	M3	**M4**	M5	M6	M7
AIC	(10,10,5)	**.99**	.01	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**.85**	.15	0	0	0	0	0	0	**.99**	0	0	.02
	(20,20,5)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**.83**	.17	0	0	0	0	0	0	**1**	0	0	0
	(10,10,10)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**.94**	.06	0	0	0	0	0	0	**1**	0	0	0
	(10,10,20)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**.99**	.01	0	0	0	0	0	0	**1**	0	0	0
	(20,20,20)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**.99**	.01	0	0	0	0	0	0	**1**	0	0	0
BIC	(10,10,5)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	1	0	**0**	0	0	0	0	.98	0	0	**0**	0	0	0
	(20,20,5)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	1	0	**0**	0	0	0	0	.83	.18	0	**0**	0	0	0
	(10,10,10)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	1	0	**0**	0	0	0	0	.56	.44	0	**0**	0	0	0
	(10,10,20)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	.01	0	**.99**	0	0	0	0	0	0	0	**1**	0	0	0
	(20,20,20)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	.33	0	**.67**	0	0	0	0	0	.05	0	**.95**	0	0	0
BIC$_2$	(10,10,5)	**.98**	.01	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**.79**	.18	0	0	.03	0	0	0	**.96**	0	0	.04
	(20,20,5)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**.84**	.16	0	0	0	0	0	0	**1**	0	0	0
	(10,10,10)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**.94**	.06	0	0	0	0	0	0	**1**	0	0	0
	(10,10,20)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**.99**	.01	0	0	0	0	0	0	**1**	0	0	0
	(20,20,20)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	.01	0	0	0	0	0	0	**1**	0	0	0
CV	(10,10,5)	**.99**	.01	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**.93**	.07	0	0	0	0	0	0	**1**	0	0	0
	(20,20,5)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**.91**	.09	0	0	0	0	0	0	**1**	0	0	0
	(10,10,10)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**.97**	.03	0	0	0	0	0	0	**1**	0	0	0
	(10,10,20)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	.01	0	0	0	0	0	0	**1**	0	0	0
	(20,20,20)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**.99**	.01	0	0	0	0	0	0	**1**	0	0	0
CV*	(10,10,5)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	.55	.03	**.39**	.03	0	0	0	.04	.40	0	**.55**	.01	0	0
	(20,20,5)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	.58	.02	**.37**	.04	0	0	0	0	.47	0	**.53**	0	0	0
	(10,10,10)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**.96**	.04	0	0	0	0	0	0	**1**	0	0	0
	(10,10,20)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	.01	0	0	0	0	0	0	**1**	0	0	0
	(20,20,20)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**.99**	.01	0	0	0	0	0	0	**1**	0	0	0

	(N,M,T)	True model: M5 Selected models							True model: M6 Selected models							True model: M7 Selected models						
		M1	M2	M3	M4	**M5**	M6	M7	M1	M2	M3	M4	M5	**M6**	M7	M1	M2	M3	M4	M5	M6	**M7**
AIC	(10,10,5)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	.01	0	0	0	0	0	0	**1**
	(20,20,5)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**
	(10,10,10)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**
	(10,10,20)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**
	(20,20,20)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**
BIC	(10,10,5)	.94	0	0	0	**.06**	0	0	1	0	0	0	0	**0**	0	1	0	0	0	0	0	**0**
	(20,20,5)	.01	0	0	0	**.99**	0	0	1	0	0	0	0	**0**	0	1	0	0	0	0	0	**0**
	(10,10,10)	1	0	0	0	**0**	0	0	1	0	0	0	0	**0**	0	1	0	0	0	0	0	**0**
	(10,10,20)	1	0	0	0	**0**	0	0	1	0	0	0	0	**0**	0	1	0	0	0	0	0	**0**
	(20,20,20)	.34	0	0	0	**.66**	0	0	1	0	0	0	0	**0**	0	1	0	0	0	0	0	**0**
BIC$_2$	(10,10,5)	0	0	0	0	**.98**	.01	.01	0	0	0	0	0	**.98**	.02	0	0	0	0	0	0	**1**
	(20,20,5)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**
	(10,10,10)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**
	(10,10,20)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**
	(20,20,20)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**
CV	(10,10,5)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**.95**
	(20,20,5)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**
	(10,10,10)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**
	(10,10,20)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**
	(20,20,20)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**
CV*	(10,10,5)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	.25	.01	0	0	.02	.36	**.37**
	(20,20,5)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	.58	**.42**
	(10,10,10)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**
	(10,10,20)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**
	(20,20,20)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**

Table 3B: Frequency of the model selected: static panels, $\rho = 1/4$

		True model: M1							True model: M2							True model: M3							True model: M4						
		Selected models							Selected models							Selected models							Selected models						
	(N,M,T)	**M1**	M2	M3	M4	M5	M6	M7	M1	**M2**	M3	M4	M5	M6	M7	M1	M2	**M3**	M4	M5	M6	M7	M1	M2	M3	**M4**	M5	M6	M7
	(10,10,5)	**.77**	.07	.13	.03	0	0	.01	0	**.73**	0	.27	0	0	.01	0	0	**.84**	.14	0	0	.01	0	0	0	**.98**	0	0	.02
AIC	(20,20,5)	**.92**	.04	.03	.01	0	0	0	0	**.93**	0	.07	0	0	0	0	0	**.84**	.16	0	0	0	0	0	0	**1**	0	0	0
	(10,10,10)	**.83**	.06	.10	.01	0	0	0	0	**.82**	0	.18	0	0	0	0	0	**.97**	.07	0	0	0	0	0	0	**1**	0	0	0
	(10,10,20)	**.88**	.02	.09	0	0	0	0	0	**.84**	0	.16	0	0	0	0	0	**.98**	.02	0	0	0	0	0	0	**1**	0	0	0
	(20,20,20)	**.97**	.02	.01	0	0	0	0	0	**.99**	0	.01	0	0	0	0	0	**.99**	.01	0	0	0	0	0	0	**1**	0	0	0
	(10,10,5)	**1**	1	0	0	0	0	0	0	**1**	0	0	0	0	0	1	0	**0**	0	0	0	0	.97	.03	0	**0**	0	0	0
BIC	(20,20,5)	**1**	1	0	0	0	0	0	0	**1**	0	0	0	0	0	1	0	**0**	0	0	0	0	.82	.18	0	**0**	0	0	0
	(10,10,10)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	.99	0	**.01**	0	0	0	0	.77	.21	0	**.03**	0	0	0
	(10,10,20)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	.01	0	**.99**	0	0	0	0	0	0	0	**1**	0	0	0
	(20,20,20)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	.13	0	**.87**	0	0	0	0	0	.01	0	**.99**	0	0	0
	(10,10,5)	**.53**	.07	.31	.08	0	0	.02	0	**.45**	0	.52	0	0	.03	0	0	**.78**	.18	0	0	.04	0	0	0	**.95**	0	0	.05
BIC$_2$	(20,20,5)	**.94**	.04	.02	0	0	0	0	0	**.95**	0	.05	0	0	0	0	0	**.85**	.16	0	0	0	0	0	0	**1**	0	0	0
	(10,10,10)	**.78**	.07	.15	.01	0	0	0	0	**.77**	0	.23	0	0	0	0	0	**.93**	.07	0	0	0	0	0	0	**1**	0	0	0
	(10,10,20)	**.90**	.02	.08	0	0	0	0	0	**.87**	0	.13	0	0	0	0	0	**.99**	.02	0	0	0	0	0	0	**1**	0	0	0
	(20,20,20)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**.99**	.01	0	0	0	0	0	0	**1**	0	0	0
	(10,10,5)	**.90**	.07	.03	0	0	0	0	0	**.95**	0	.05	0	0	0	0	0	**.93**	.07	0	0	0	0	0	0	**1**	0	0	0
CV	(20,20,5)	**.96**	.04	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**.91**	.09	0	0	0	0	0	0	**1**	0	0	0
	(10,10,10)	**.89**	.06	.06	0	0	0	0	0	**.93**	0	.07	0	0	0	0	0	**.96**	.04	0	0	0	0	0	0	**1**	0	0	0
	(10,10,20)	**.91**	.03	.06	0	0	0	0	0	**.91**	0	.09	0	0	0	0	0	**.99**	.01	0	0	0	0	0	0	**1**	0	0	0
	(20,20,20)	**.97**	.02	.01	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**.99**	.01	0	0	0	0	0	0	**1**	0	0	0
	(10,10,5)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	.79	.06	**.14**	.02	0	0	0	.03	.74	0	**.23**	.01	0	0
CV*	(20,20,5)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	.96	.02	**.02**	0	0	0	0	.01	.95	0	**.04**	0	0	0
	(10,10,10)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**.97**	.03	0	0	0	0	0	0	**1**	0	0	0
	(10,10,20)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	.01	0	0	0	0	0	0	**1**	0	0	0
	(20,20,20)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**.99**	.01	0	0	0	0	0	0	**1**	0	0	0

		True model: M5							True model: M6							True model: M7						
		Selected models							Selected models							Selected models						
	(N,M,T)	M1	M2	M3	M4	**M5**	M6	M7	M1	M2	M3	M4	M5	**M6**	M7	M1	M2	M3	M4	M5	M6	**M7**
	(10,10,5)	0	0	0	0	**.79**	0	.20	0	0	0	0	0	**.27**	.73	0	0	0	0	0	0	**1**
AIC	(20,20,5)	0	0	0	0	**.97**	0	.03	0	0	0	0	0	**.63**	.37	0	0	0	0	0	0	**1**
	(10,10,10)	0	0	0	0	**.98**	0	.02	0	0	0	0	0	**.36**	.64	0	0	0	0	0	0	**1**
	(10,10,20)	0	0	0	0	**1**	0	0	0	0	0	0	0	**.49**	.60	0	0	0	0	0	0	**1**
	(20,20,20)	0	0	0	0	**1**	0	0	0	0	0	0	0	**.82**	.18	0	0	0	0	0	0	**1**
	(10,10,5)	.97	0	0	0	**.03**	0	0	1	0	0	0	0	**0**	0	1	0	0	0	0	0	**0**
BIC	(20,20,5)	.06	0	0	0	**.94**	0	0	1	0	0	0	0	**0**	0	1	0	0	0	0	0	**0**
	(10,10,10)	1	0	0	0	**0**	0	0	1	0	0	0	0	**0**	0	1	0	0	0	0	0	**0**
	(10,10,20)	1	0	0	0	**0**	0	0	1	0	0	0	0	**0**	0	1	0	0	0	0	0	**0**
	(20,20,20)	.79	0	0	0	**.22**	0	0	1	0	0	0	0	**0**	0	1	0	0	0	0	0	**0**
	(10,10,5)	0	0	0	0	**.50**	0	.50	0	0	0	0	0	**.12**	.88	0	0	0	0	0	0	**1**
BIC$_2$	(20,20,5)	0	0	0	0	**.98**	0	.02	0	0	0	0	0	**.70**	.30	0	0	0	0	0	0	**1**
	(10,10,10)	0	0	0	0	**.96**	0	.04	0	0	0	0	0	**.28**	.72	0	0	0	0	0	0	**1**
	(10,10,20)	0	0	0	0	**1**	0	0	0	0	0	0	0	**.43**	.57	0	0	0	0	0	0	**1**
	(20,20,20)	0	0	0	0	**1**	0	0	0	0	0	0	0	**.99**	.01	0	0	0	0	0	0	**1**
	(10,10,5)	0	0	0	0	**1**	0	0	0	0	0	0	0	**.96**	.04	0	0	0	0	0	0	**1**
CV	(20,20,5)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**
	(10,10,10)	0	0	0	0	**1**	0	0	0	0	0	0	0	**.93**	.07	0	0	0	0	0	0	**1**
	(10,10,20)	0	0	0	0	**1**	0	0	0	0	0	0	0	**.92**	.08	0	0	0	0	0	0	**1**
	(20,20,20)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**
	(10,10,5)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	.07	0	0	0	.01	.74	**.18**
CV*	(20,20,5)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	.98	**.02**
	(10,10,10)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**
	(10,10,20)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**
	(20,20,20)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**

36

Table 3C: Frequency of the model selected: static panels, $\rho = 1/3$

	(N,M,T)	True model: M1 Selected models **M1**	M2	M3	M4	M5	M6	M7	True model: M2 Selected models M1	**M2**	M3	M4	M5	M6	M7	True model: M3 Selected models M1	M2	**M3**	M4	M5	M6	M7	True model: M4 Selected models M1	M2	M3	**M4**	M5	M6	M7
AIC	(10,10,5)	**.33**	.05	.51	.10	0	0	.01	0	**.29**	0	.70	0	0	.01	0	0	**.84**	.15	0	0	.01	0	0	0	**.98**	0	0	.02
	(20,20,5)	**.28**	.03	.57	.11	0	0	0	0	**.21**	0	.79	0	0	0	0	0	**.84**	.16	0	0	0	0	0	0	**1**	0	0	0
	(10,10,10)	**.41**	.06	.50	.04	0	0	0	0	**.37**	0	.63	0	0	0	0	0	**.93**	.07	0	0	0	0	0	0	**1**	0	0	0
	(10,10,20)	**.45**	.02	.52	.01	0	0	0	0	**.43**	0	.57	0	0	0	0	0	**.98**	.02	0	0	0	0	0	0	**1**	0	0	0
	(20,20,20)	**.42**	.03	.54	.02	0	0	0	0	**.39**	0	.61	0	0	0	0	0	**.98**	.02	0	0	0	0	0	0	**1**	0	0	0
BIC	(10,10,5)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	1	0	**0**	0	0	0	0	.97	.03	0	**0**	0	0	0
	(20,20,5)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	1	0	**0**	0	0	0	0	.82	.18	0	**0**	0	0	0
	(10,10,10)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	.98	0	**.02**	0	0	0	0	.76	.19	0	**.05**	0	0	0
	(10,10,20)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	.01	0	**.99**	0	0	0	0	0	0	0	**1**	0	0	0
	(20,20,20)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	.14	0	**.86**	0	0	0	0	0	.02	0	**.99**	0	0	0
BIC$_2$	(10,10,5)	**.13**	.03	.66	.14	0	0	.04	0	**.12**	0	.83	0	0	.05	0	0	**.78**	.18	0	0	.04	0	0	0	**.95**	0	0	.06
	(20,20,5)	**.35**	.03	.52	.10	0	0	0	0	**.28**	0	.72	0	0	0	0	0	**.85**	.15	0	0	0	0	0	0	**1**	0	0	0
	(10,10,10)	**.31**	.04	.59	.06	0	0	0	0	**.30**	0	.70	0	0	0	0	0	**.92**	.08	0	0	0	0	0	0	**1**	0	0	0
	(10,10,20)	**.49**	.02	.48	.01	0	0	0	0	**.46**	0	.54	0	0	0	0	0	**.98**	.02	0	0	0	0	0	0	**1**	0	0	0
	(20,20,20)	**.86**	.03	.11	0	0	0	0	0	**.86**	0	.14	0	0	0	0	0	**.99**	.01	0	0	0	0	0	0	**1**	0	0	0
CV	(10,10,5)	**.63**	.10	.25	.02	0	0	0	0	**.66**	0	.34	0	0	0	0	0	**.92**	.08	0	0	0	0	0	0	**1**	0	0	0
	(20,20,5)	**.77**	.11	.11	.02	0	0	0	0	**.83**	0	.17	0	0	0	0	0	**.91**	.09	0	0	0	0	0	0	**1**	0	0	0
	(10,10,10)	**.55**	.08	.35	0	0	0	0	0	**.59**	0	.41	0	0	0	0	0	**.95**	.05	0	0	0	0	0	0	**1**	0	0	0
	(10,10,20)	**.53**	.02	.45	0	0	0	0	0	**.53**	0	.47	0	0	0	0	0	**.99**	.01	0	0	0	0	0	0	**1**	0	0	0
	(20,20,20)	**.56**	.05	.39	.01	0	0	0	0	**.56**	0	.44	0	0	0	0	0	**.99**	.01	0	0	0	0	0	0	**1**	0	0	0
CV*	(10,10,5)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	.81	.06	**.12**	.02	0	0	0	.03	.78	0	**.18**	.01	0	0
	(20,20,5)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	.96	.03	**.01**	0	0	0	0	.01	.97	0	**.02**	0	0	0
	(10,10,10)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**.97**	.03	0	0	0	0	0	0	**1**	0	0	0
	(10,10,20)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**.99**	.01	0	0	0	0	0	0	**1**	0	0	0
	(20,20,20)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**.99**	.01	0	0	0	0	0	0	**1**	0	0	0

	(N,M,T)	True model: M5 Selected models M1	M2	M3	M4	**M5**	M6	M7	True model: M6 Selected models M1	M2	M3	M4	M5	**M6**	M7	True model: M7 Selected models M1	M2	M3	M4	M5	M6	**M7**
AIC	(10,10,5)	0	0	0	0	**.38**	0	.62	0	0	0	0	0	**.05**	.95	0	0	0	0	0	0	**1**
	(20,20,5)	0	0	0	0	**.39**	0	.61	0	0	0	0	0	**.02**	.98	0	0	0	0	0	0	**1**
	(10,10,10)	0	0	0	0	**.85**	0	.15	0	0	0	0	0	**.07**	.93	0	0	0	0	0	0	**1**
	(10,10,20)	0	0	0	0	**1**	0	0	0	0	0	0	0	**.10**	.90	0	0	0	0	0	0	**1**
	(20,20,20)	0	0	0	0	**1**	0	0	0	0	0	0	0	**.08**	.92	0	0	0	0	0	0	**1**
BIC	(10,10,5)	.98	0	0	0	**.02**	0	0	1	0	0	0	0	**0**	0	1	0	0	0	0	0	**0**
	(20,20,5)	.13	0	0	0	**.87**	0	0	1	0	0	0	0	**0**	0	1	0	0	0	0	0	**0**
	(10,10,10)	1	0	0	0	**0**	0	0	1	0	0	0	0	**0**	0	1	0	0	0	0	0	**0**
	(10,10,20)	1	0	0	0	**0**	0	0	1	0	0	0	0	**0**	0	1	0	0	0	0	0	**0**
	(20,20,20)	.97	0	0	0	**.04**	0	0	1	0	0	0	0	**0**	0	1	0	0	0	0	0	**0**
BIC$_2$	(10,10,5)	0	0	0	0	**.14**	0	.86	0	0	0	0	0	**.02**	.98	0	0	0	0	0	0	**1**
	(20,20,5)	0	0	0	0	**.46**	0	.54	0	0	0	0	0	**.04**	.96	0	0	0	0	0	0	**1**
	(10,10,10)	0	0	0	0	**.75**	0	.25	0	0	0	0	0	**.04**	.96	0	0	0	0	0	0	**1**
	(10,10,20)	0	0	0	0	**1**	0	0	0	0	0	0	0	**.11**	.89	0	0	0	0	0	0	**1**
	(20,20,20)	0	0	0	0	**1**	0	0	0	0	0	0	0	**.39**	.61	0	0	0	0	0	0	**1**
CV	(10,10,5)	0	0	0	0	**.98**	0	.02	0	0	0	0	0	**.72**	.29	0	0	0	0	0	0	**1**
	(20,20,5)	0	0	0	0	**.99**	0	.01	0	0	0	0	0	**.85**	.15	0	0	0	0	0	0	**1**
	(10,10,10)	0	0	0	0	**1**	0	0	0	0	0	0	0	**.61**	.39	0	0	0	0	0	0	**1**
	(10,10,20)	0	0	0	0	**1**	0	0	0	0	0	0	0	**.53**	.47	0	0	0	0	0	0	**1**
	(20,20,20)	0	0	0	0	**1**	0	0	0	0	0	0	0	**.58**	.42	0	0	0	0	0	0	**1**
CV*	(10,10,5)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	.03	0	0	0	.01	.80	**.16**
	(20,20,5)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	.99	**.01**
	(10,10,10)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**
	(10,10,20)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**
	(20,20,20)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**

Table 3D: Frequency of the model selected: static panels, $\rho = 3/4$

		True model: M1							True model: M2							True model: M3							True model: M4						
		Selected models							Selected models							Selected models							Selected models						
	(N,M,T)	**M1**	M2	M3	M4	M5	M6	M7	M1	**M2**	M3	M4	M5	M6	M7	M1	M2	**M3**	M4	M5	M6	M7	M1	M2	M3	**M4**	M5	M6	M7
AIC	(10,10,5)	**0**	0	.81	.17	0	0	.02	0	**0**	0	.97	0	0	.04	0	0	**.81**	.17	0	0	.02	0	0	0	**.97**	0	0	.02
	(20,20,5)	**0**	0	.84	.16	0	0	0	0	**0**	0	1	0	0	0	0	0	**.84**	.16	0	0	0	0	0	0	**1**	0	0	0
	(10,10,10)	**0**	0	.88	.12	0	0	0	0	**0**	0	1	0	0	0	0	0	**.88**	.13	0	0	0	0	0	0	**1**	0	0	0
	(10,10,20)	**0**	0	.94	.07	0	0	0	0	**0**	0	1	0	0	0	0	0	**.94**	.07	0	0	0	0	0	0	**1**	0	0	0
	(20,20,20)	**0**	0	.93	.07	0	0	0	0	**0**	0	1	0	0	0	0	0	**.93**	.07	0	0	0	0	0	0	**1**	0	0	0
BIC	(10,10,5)	**.89**	0	.10	0	0	0	0	.01	**.62**	0	.36	0	0	0	.29	0	**.70**	.01	0	0	0	.26	.01	0	**.73**	0	0	0
	(20,20,5)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	1	0	**0**	0	0	0	0	.95	.05	0	**0**	0	0	0
	(10,10,10)	**.75**	0	.25	0	0	0	0	0	**.57**	0	.44	0	0	0	.02	0	**.98**	0	0	0	0	.02	0	0	**.97**	0	0	0
	(10,10,20)	**.80**	0	.20	0	0	0	0	0	**.69**	0	.32	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0
	(20,20,20)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0
BIC$_2$	(10,10,5)	**0**	0	.75	.18	0	0	.06	0	**0**	0	.93	0	0	.07	0	0	**.75**	.19	0	0	.06	0	0	0	**.92**	0	0	.08
	(20,20,5)	**0**	0	.84	.16	0	0	0	0	**0**	0	1	0	0	0	0	0	**.84**	.16	0	0	0	0	0	0	**1**	0	0	0
	(10,10,10)	**0**	0	.87	.13	0	0	0	0	**0**	0	1	0	0	0	0	0	**.87**	.13	0	0	0	0	0	0	**1**	0	0	0
	(10,10,20)	**0**	0	.94	.06	0	0	0	0	**0**	0	1	0	0	0	0	0	**.94**	.06	0	0	0	0	0	0	**1**	0	0	0
	(20,20,20)	**0**	0	.95	.05	0	0	0	0	**0**	0	1	0	0	0	0	0	**.95**	.05	0	0	0	0	0	0	**1**	0	0	0
CV	(10,10,5)	**0**	0	.90	.11	0	0	0	0	**0**	0	1	0	0	0	0	0	**.90**	.10	0	0	0	0	0	0	**1**	0	0	0
	(20,20,5)	**0**	0	.91	.10	0	0	0	0	**0**	0	1	0	0	0	0	0	**.90**	.10	0	0	0	0	0	0	**1**	0	0	0
	(10,10,10)	**0**	0	.91	.10	0	0	0	0	**0**	0	1	0	0	0	0	0	**.91**	.09	0	0	0	0	0	0	**1**	0	0	0
	(10,10,20)	**0**	0	.95	.06	0	0	0	0	**0**	0	1	0	0	0	0	0	**.95**	.06	0	0	0	0	0	0	**1**	0	0	0
	(20,20,20)	**0**	0	.95	.05	0	0	0	0	**0**	0	1	0	0	0	0	0	**.95**	.05	0	0	0	0	0	0	**1**	0	0	0
CV*	(10,10,5)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	.65	.16	**.18**	.02	0	0	0	.01	.67	0	**.30**	.02	0	0
	(20,20,5)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	.82	.12	**.05**	.01	0	0	0	0	.89	0	**.10**	0	0	0
	(10,10,10)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	.01	**.96**	.03	0	0	0	0	.02	0	**.98**	0	0	0
	(10,10,20)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0
	(20,20,20)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**.99**	.01	0	0	0	0	0	0	**1**	0	0	0

		True model: M5							True model: M6							True model: M7						
		Selected models							Selected models							Selected models						
	(N,M,T)	M1	M2	M3	M4	**M5**	M6	M7	M1	M2	M3	M4	M5	**M6**	M7	M1	M2	M3	M4	M5	M6	**M7**
AIC	(10,10,5)	0	0	0	0	**0**	0	1	0	0	0	0	0	**0**	1	0	0	0	0	0	0	**1**
	(20,20,5)	0	0	0	0	**0**	0	1	0	0	0	0	0	**0**	1	0	0	0	0	0	0	**1**
	(10,10,10)	0	0	0	0	**0**	0	1	0	0	0	0	0	**0**	1	0	0	0	0	0	0	**1**
	(10,10,20)	0	0	0	0	**0**	0	1	0	0	0	0	0	**0**	1	0	0	0	0	0	0	**1**
	(20,20,20)	0	0	0	0	**0**	0	1	0	0	0	0	0	**0**	1	0	0	0	0	0	0	**1**
BIC	(10,10,5)	.99	.01	0	0	**0**	0	0	.99	0	0	0	0	**0**	0	1	0	0	0	0	0	**0**
	(20,20,5)	1	0	0	0	**0**	0	0	1	0	0	0	0	**0**	0	1	0	0	0	0	0	**0**
	(10,10,10)	.99	.01	0	0	**0**	0	0	1	0	0	0	0	**0**	0	1	0	0	0	0	0	**0**
	(10,10,20)	.99	0	.01	0	**0**	0	0	1	0	0	0	0	**0**	0	.98	0	.02	0	0	0	**0**
	(20,20,20)	1	0	0	0	**0**	0	0	1	0	0	0	0	**0**	0	1	0	0	0	0	0	**0**
BIC$_2$	(10,10,5)	0	0	0	0	**0**	0	1	0	0	0	0	0	**0**	1	0	0	0	0	0	0	**1**
	(20,20,5)	0	0	0	0	**0**	0	1	0	0	0	0	0	**0**	1	0	0	0	0	0	0	**1**
	(10,10,10)	0	0	0	0	**0**	0	1	0	0	0	0	0	**0**	1	0	0	0	0	0	0	**1**
	(10,10,20)	0	0	0	0	**0**	0	1	0	0	0	0	0	**0**	1	0	0	0	0	0	0	**1**
	(20,20,20)	0	0	0	0	**0**	0	1	0	0	0	0	0	**0**	1	0	0	0	0	0	0	**1**
CV	(10,10,5)	0	0	0	0	**0**	0	.99	0	0	0	0	0	**0**	1	0	0	0	0	0	0	**.95**
	(20,20,5)	0	0	0	0	**0**	0	1	0	0	0	0	0	**0**	1	0	0	0	0	0	0	**1**
	(10,10,10)	0	0	.01	.01	**0**	0	.98	0	0	0	0	0	**0**	1	0	0	0	0	0	0	**1**
	(10,10,20)	0	0	.08	.05	**0**	0	.87	0	0	0	0	0	**0**	1	0	0	0	0	0	0	**1**
	(20,20,20)	0	0	0	0	**0**	0	1	0	0	0	0	0	**0**	1	0	0	0	0	0	0	**1**
CV*	(10,10,5)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	.25	.01	0	0	0	.74	**.25**
	(20,20,5)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	.93	**.07**
	(10,10,10)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	.03	**.98**
	(10,10,20)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**
	(20,20,20)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**

Table 4: Frequency of the model selected: dynamic panels, $\beta = (1, 3/4)'$

	(N,M,T)	True model: M1 **M1**	M2	M3	M4	M5	M6	M7	True model: M2 M1	**M2**	M3	M4	M5	M6	M7	True model: M3 M1	M2	**M3**	M4	M5	M6	M7	True model: M4 M1	M2	M3	**M4**	M5	M6	M7
	(10,10,5)	**.49**	0	.42	.08	0	0	.01	0	**.20**	0	.78	0	0	.02	0	0	**.84**	.15	0	0	.01	0	0	0	**.98**	0	0	.02
AIC	(20,20,5)	**.53**	0	.40	.07	0	0	0	0	**.21**	0	.79	0	0	0	0	0	**.85**	.15	0	0	0	0	0	0	**1**	0	0	0
	(10,10,10)	**.95**	0	.04	0	0	0	0	0	**.82**	0	.18	0	0	0	0	0	**.94**	.06	0	0	0	0	0	0	**1**	0	0	0
	(10,10,20)	**1**	0	0	0	0	0	0	0	**.99**	0	.01	0	0	0	0	0	**.99**	.01	0	0	0	0	0	0	**1**	0	0	0
	(20,20,20)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**.99**	.01	0	0	0	0	0	0	**1**	0	0	0
	(10,10,5)	**1**	0	0	0	0	0	0	.04	**.96**	0	0	0	0	0	1	0	**0**	0	0	0	0	.14	.86	0	**0**	0	0	0
BIC	(20,20,5)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	1	0	**0**	0	0	0	0	.05	.95	0	**0**	0	0	0
	(10,10,10)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	1	0	**0**	0	0	0	0	0	1	0	**0**	0	0	0
	(10,10,20)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	1	0	**0**	0	0	0	0	0	1	0	**0**	0	0	0
	(20,20,20)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	1	0	**0**	0	0	0	0	0	1	0	**0**	0	0	0
	(10,10,5)	**.20**	0	.61	.15	0	0	.04	0	**.07**	0	.88	0	0	.05	0	0	**.79**	.18	0	0	.04	0	0	0	**.95**	0	0	.05
BIC$_2$	(20,20,5)	**.62**	0	.32	.06	0	0	0	0	**.28**	0	.72	0	0	0	0	0	**.86**	.14	0	0	0	0	0	0	**1**	0	0	0
	(10,10,10)	**.91**	0	.08	.01	0	0	0	0	**.73**	0	.27	0	0	0	0	0	**.93**	.07	0	0	0	0	0	0	**1**	0	0	0
	(10,10,20)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**.99**	.01	0	0	0	0	0	0	**1**	0	0	0
	(20,20,20)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**.99**	.01	0	0	0	0	0	0	**1**	0	0	0
	(10,10,5)	**.87**	0	.11	.02	0	0	0	0	**.65**	0	.35	0	0	0	.04	0	**.89**	.08	0	0	0	0	0	0	**1**	0	0	0
CV	(20,20,5)	**.98**	0	.02	0	0	0	0	0	**.91**	0	.09	0	0	0	0	0	**.93**	.08	0	0	0	0	0	0	**1**	0	0	0
	(10,10,10)	**.99**	0	.01	0	0	0	0	0	**.94**	0	.06	0	0	0	0	0	**.96**	.04	0	0	0	0	0	0	**1**	0	0	0
	(10,10,20)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**.99**	.01	0	0	0	0	0	0	**1**	0	0	0
	(20,20,20)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**.99**	.01	0	0	0	0	0	0	**1**	0	0	0

	(N,M,T)	True model: M5 M1	M2	M3	M4	**M5**	M6	M7	True model: M6 M1	M2	M3	M4	M5	**M6**	M7	True model: M7 M1	M2	M3	M4	M5	M6	**M7**
	(10,10,5)	0	0	0	0	**.53**	0	.47	0	0	0	0	0	**.05**	.95	0	0	0	0	0	0	**1**
AIC	(20,20,5)	0	0	0	0	**.61**	0	.40	0	0	0	0	0	**.05**	.96	0	0	0	0	0	0	**1**
	(10,10,10)	0	0	0	0	**.99**	0	.01	0	0	0	0	0	**.46**	.54	0	0	0	0	0	0	**1**
	(10,10,20)	0	0	0	0	**1**	0	0	0	0	0	0	0	**.90**	.11	0	0	0	0	0	0	**1**
	(20,20,20)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**
	(10,10,5)	.06	0	0	0	**.94**	0	0	.14	.06	0	0	.01	**.79**	0	.21	0	0	0	.01	.78	**0**
BIC	(20,20,5)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	1	**0**
	(10,10,10)	.10	0	0	0	**.90**	0	0	.45	.13	0	0	0	**.42**	0	.62	.01	0	0	0	.37	**0**
	(10,10,20)	.26	0	0	0	**.74**	0	0	.85	.14	0	0	0	**.01**	0	.98	.02	0	0	0	0	**0**
	(20,20,20)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	1	**0**
	(10,10,5)	0	0	0	0	**.20**	0	.80	0	0	0	0	0	**.01**	.99	0	0	0	0	0	0	**1**
BIC$_2$	(20,20,5)	0	0	0	0	**.70**	0	.30	0	0	0	0	0	**.07**	.93	0	0	0	0	0	0	**1**
	(10,10,10)	0	0	0	0	**.99**	0	.01	0	0	0	0	0	**.37**	.63	0	0	0	0	0	0	**1**
	(10,10,20)	0	0	0	0	**1**	0	0	0	0	0	0	0	**.91**	.09	0	0	0	0	0	0	**1**
	(20,20,20)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**
	(10,10,5)	0	0	0	0	**1**	0	0	0	0	0	0	0	**.86**	.15	0	0	0	0	0	.06	**.94**
CV	(20,20,5)	0	0	0	0	**1**	0	0	0	0	0	0	0	**.98**	.02	0	0	0	0	0	0	**1**
	(10,10,10)	0	0	0	0	**1**	0	0	0	0	0	0	0	**.98**	.02	0	0	0	0	0	0	**1**
	(10,10,20)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**
	(20,20,20)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**

Table 5: Frequency of the model selected: dynamic panels with exogenous regressors

	(N,M,T)	**M1**	M2	M3	M4	M5	M6	M7	M1	**M2**	M3	M4	M5	M6	M7	M1	M2	**M3**	M4	M5	M6	M7	M1	M2	M3	**M4**	M5	M6	M7
		True model: M1							True model: M2							True model: M3							True model: M4						
AIC	(10,10,5)	**.72**	.01	.23	.04	0	0	.01	0	**.41**	0	.57	0	0	.02	.06	0	**.79**	.14	0	0	.01	0	0	0	**.98**	0	0	.02
	(20,20,5)	**.90**	0	.08	.02	0	0	0	0	**.66**	0	.34	0	0	0	0	0	**.85**	.15	0	0	0	0	0	0	**1**	0	0	0
	(10,10,10)	**.98**	.01	.02	0	0	0	0	0	**.93**	0	.08	0	0	0	0	0	**.94**	.06	0	0	0	0	0	0	**1**	0	0	0
	(10,10,20)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**.99**	.01	0	0	0	0	0	0	**1**	0	0	0
	(20,20,20)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**.99**	.01	0	0	0	0	0	0	**1**	0	0	0
BIC	(10,10,5)	**1**	0	0	0	0	0	0	.10	**.90**	0	0	0	0	0	1	0	**0**	0	0	0	0	.32	.68	0	**0**	0	0	0
	(20,20,5)	**1**	0	0	0	0	0	0	.01	**.99**	0	0	0	0	0	1	0	**0**	0	0	0	0	.11	.89	0	**0**	0	0	0
	(10,10,10)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	1	0	**0**	0	0	0	0	.03	.97	0	**0**	0	0	0
	(10,10,20)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	1	0	**0**	0	0	0	0	0	1	0	**0**	0	0	0
	(20,20,20)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	1	0	**0**	0	0	0	0	0	1	0	**0**	0	0	0
BIC_2	(10,10,5)	**.39**	0	.46	.11	0	0	.03	0	**.16**	0	.78	0	0	.06	0	0	**.78**	.17	0	0	.04	0	0	0	**.95**	0	0	.05
	(20,20,5)	**.94**	0	.05	.01	0	0	0	0	**.75**	0	.26	0	0	0	0	0	**.85**	.15	0	0	0	0	0	0	**1**	0	0	0
	(10,10,10)	**.97**	.01	.03	0	0	0	0	0	**.86**	0	.14	0	0	0	0	0	**.93**	.07	0	0	0	0	0	0	**1**	0	0	0
	(10,10,20)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**.99**	.01	0	0	0	0	0	0	**1**	0	0	0
	(20,20,20)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**.99**	.01	0	0	0	0	0	0	**1**	0	0	0
CV	(10,10,5)	**.96**	0	.04	0	0	0	0	0	**.84**	0	.16	0	0	0	.36	0	**.60**	.04	0	0	0	0	.09	0	**.91**	0	0	0
	(20,20,5)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	.22	0	**.72**	.07	0	0	0	0	.02	0	**.98**	0	0	0
	(10,10,10)	**.99**	0	0	0	0	0	0	0	**.98**	0	.02	0	0	0	.04	0	**.93**	.04	0	0	0	0	0	0	**1**	0	0	0
	(10,10,20)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**.99**	.01	0	0	0	0	0	0	**1**	0	0	0
	(20,20,20)	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**.99**	.01	0	0	0	0	0	0	**1**	0	0	0

	(N,M,T)	M1	M2	M3	M4	**M5**	M6	M7	M1	M2	M3	M4	M5	**M6**	M7	M1	M2	M3	M4	M5	M6	**M7**
		True model: M5							True model: M6							True model: M7						
AIC	(10,10,5)	0	0	0	0	**.71**	0	.30	0	0	0	0	0	**.14**	.87	0	0	0	0	0	.01	**.99**
	(20,20,5)	0	0	0	0	**.90**	0	.10	0	0	0	0	0	**.30**	.70	0	0	0	0	0	0	**1**
	(10,10,10)	0	0	0	0	**1**	0	0	0	0	0	0	0	**.63**	.37	0	0	0	0	0	0	**1**
	(10,10,20)	0	0	0	0	**1**	0	0	0	0	0	0	0	**.93**	.07	0	0	0	0	0	0	**1**
	(20,20,20)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**
BIC	(10,10,5)	.94	0	0	0	**.06**	0	0	.48	.04	0	0	.37	**.11**	0	1	0	0	0	0	0	**0**
	(20,20,5)	.02	0	0	0	**.98**	0	0	1	0	0	0	.01	**1**	0	1	0	0	0	0	0	**0**
	(10,10,10)	1	0	0	0	**0**	0	0	.82	.04	0	0	.15	**0**	0	1	0	0	0	0	0	**0**
	(10,10,20)	1	.01	0	0	**0**	0	0	.99	.01	0	0	0	**0**	0	1	0	0	0	0	0	**0**
	(20,20,20)	.35	0	0	0	**.65**	0	0	0	0	0	0	.04	**.96**	0	1	0	0	0	0	0	**0**
BIC_2	(10,10,5)	0	0	0	0	**.36**	0	.64	0	0	0	0	0	**.05**	.96	0	0	0	0	0	.01	**.99**
	(20,20,5)	0	0	0	0	**.95**	0	.05	0	0	0	0	0	**.38**	.63	0	0	0	0	0	0	**1**
	(10,10,10)	0	0	0	0	**1**	0	0	0	0	0	0	0	**.54**	.46	0	0	0	0	0	0	**1**
	(10,10,20)	0	0	0	0	**1**	0	0	0	0	0	0	0	**.94**	.06	0	0	0	0	0	0	**1**
	(20,20,20)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**
CV	(10,10,5)	0	0	0	0	**1**	0	0	0	0	0	0	0	**.95**	.05	0	0	0	0	0	.44	**.57**
	(20,20,5)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	.24	**.76**
	(10,10,10)	0	0	0	0	**1**	0	0	0	0	0	0	0	**.99**	.01	0	0	0	0	0	.08	**.92**
	(10,10,20)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**
	(20,20,20)	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**

Table 6A: Technology and contractions: Specification A

| | Model selection | | | | | Estimate and 95% CI of β_1 | | |
	AIC	BIC	BIC$_2$	CV	CV*	estimate	95% CI	
Model 1	-1.900	-1.899	-1.900	0.150	0.141	-0.0002	[-0.0098	0.0094]
Model 2	-1.937	-1.908	-1.936	0.145	0.134	-0.0279	[-0.0402	-0.0156]
Model 3	-1.892	-1.436	-1.872	0.162	0.148	-0.0314	[-0.0437	-0.0191]
Model 4	-1.904	-1.442	-1.883	0.160	0.146	-0.0304	[-0.0427	-0.0181]
Model 5	-1.904	-1.738	-1.897	0.150	0.141	-0.0267	[-0.0392	-0.0142]
Model 6	-2.098	-1.557	-2.075	0.126	0.113	-0.0130	[-0.0242	-0.0018]
Model 7	-2.082	-1.108	-2.039	0.147	0.126	-0.0141	[-0.0251	-0.0031]
Selected model	M6	M2	M6	M6	M6			

Notes: The dependent variable is value added. The independent variables include Contraction×EFD and the control variable for Specification A. β_1 is the coefficient on Contraction×EFD. "CI" stands for "confidence interval". The CI is based on the heteroskedasticity-robust standard errors. The total sample size is 57,115.

Table 6B: Technology and contractions: Specification B

| | Model selection | | | | | Estimate and 95% CI of β_1 | | |
	AIC	BIC	BIC$_2$	CV	CV*	estimate	95% CI	
Model 1	-1.903	-1.901	-1.903	0.149	0.140	0.0107	[-0.0134	0.0348]
Model 2	-1.940	-1.909	-1.938	0.144	0.134	-0.0274	[-0.0584	0.0036]
Model 3	-1.894	-1.436	-1.874	0.162	0.147	-0.0297	[-0.0601	0.0007]
Model 4	-1.906	-1.442	-1.885	0.160	0.145	-0.0296	[-0.0596	0.0004]
Model 5	-1.908	-1.740	-1.900	0.149	0.140	-0.0165	[-0.0477	0.0147]
Model 6	-2.099	-1.556	-2.075	0.126	0.113	-0.0163	[-0.0439	0.0113]
Model 7	-2.082	-1.107	-2.039	0.149	0.127	-0.0236	[-0.0510	0.0038]
Selected model	M6	M2	M6	M6	M6			

Notes: The dependent variable is value added. The independent variables include Contraction×EFD and the control variable for Specification A, and Contraction×EFD, Contraction×DEP, Contraction×ISTC, Contraction×RND, Contraction×HC, Contraction×LAB, Contraction×FIX, Contraction×LMP, Contraction×SPEC, Contraction×INT, and the control variable for Specification B. β_1 is the coefficient on Contraction×EFD. "CI" stands for "confidence interval". The CI is based on the heteroskedasticity-robust standard errors. The total sample size is 57,115.

Table 7A: Gravity equations: Specification A

| | Model selection | | | | | Estimate and 95% CI of β_1 | | |
	AIC	BIC	BIC$_2$	CV	CV*	estimate	95% CI	
Model 1	1.077	1.078	1.077	2.937	0.237	1.214	[1.206	1.223]
Model 2	0.309	0.332	0.310	1.363	0.183	0.333	[0.297	0.369]
Model 3	-0.668	-0.465	-0.660	0.515	0.165	1.252	[1.247	1.258]
Model 4	-0.710	-0.496	-0.701	0.493	0.162	1.277	[1.239	1.316]
Model 5	0.874	1.188	0.888	2.398	0.225	1.345	[1.328	1.362]
Model 6	0.240	0.856	0.266	1.281	0.184	-0.002	[-0.040	0.036]
Model 7	-1.080	-0.273	-1.045	0.347	0.156	0.657	[0.577	0.738]
Selected model	M7	M4	M7	M7	M7			

Notes: The dependent variable is ln(Export$_{ijt}$). The independent variables include ln(GDP$_{it}$+GDP$_{jt}$) for Specification A. "CI" stands for "confidence interval". The CI is based on the heteroskedasticity-robust standard errors. The total sample size is 48,403.

Table 7B: Gravity equations: Specification B

| | Model selection | | | | | Estimate and CI of β_1 | | | Estimate and 95% CI of β_2 | | |
	AIC	BIC	BIC$_2$	CV	CV*	estimate	95% CI		estimate	95% CI	
Model 1	1.075	1.076	1.075	2.931	0.236	1.251	[1.240	1.262]	-0.100	[-0.120	-0.079]
Model 2	0.287	0.310	0.288	1.333	0.182	0.657	[0.614	0.701]	-0.714	[-0.762	-0.665]
Model 3	-0.670	-0.466	-0.661	0.514	0.165	1.217	[1.208	1.227]	0.335	[0.245	0.425]
Model 4	-0.713	-0.499	-0.704	0.492	0.162	1.262	[1.224	1.300]	0.461	[0.370	0.552]
Model 5	0.826	1.139	0.839	2.284	0.224	1.951	[1.919	1.982]	-0.844	[-0.880	-0.809]
Model 6	0.228	0.845	0.255	1.267	0.185	0.340	[0.285	0.395]	-0.550	[-0.605	-0.495]
Model 7	-1.082	-0.275	-1.048	0.346	0.156	0.619	[0.539	0.700]	0.864	[0.664	1.064]
Selected model	M7	M4	M7	M7	M7						

Notes: The dependent variable is ln(Export$_{ijt}$). The independent variables include ln(GDP$_{it}$+GDP$_{jt}$) and ln(POP$_{it}$+POP$_{jt}$) for Specification B. "CI" stands for "confidence interval". The CI is based on the heteroskedasticity-robust standard errors. The total sample size is 48,403.

Supplementary Appendix to
"Determination of Different Types of Fixed Effects in Three-Dimensional Panels"

(NOT for Publication)

Xun Lu[a], Ke Miao[b] and Liangjun Su[b]

[a] Department of Economics, Chinese University of Hong Kong
[b] School of Economics, Singapore Management University, Singapore

This supplement is composed of three Appendices. Appendix B contains the proof of Theorem 4.1; Appendix C contains the proofs of the technical lemmas in Appendices A and B; Appendix D contains the Nickell biases for the seven estimators of the slope coefficient in the panel AR(1) model.

B Proof of Theorem 4.1

To prove Theorem 4.1, without loss of generality and for notational simplicity we consider the AR(1) filtering for $\{u_{ijt}, t \geq 1\}$. Let $\widehat{\mathbf{u}}_{ij} = (\hat{u}_{ij,2}, ..., \hat{u}_{ij,T})'$, $\widehat{\mathbf{U}} = (\hat{u}'_{11}, ..., \hat{u}'_{1M}, ..., \hat{u}'_{N1}, ..., \hat{u}'_{NM})'$, $\hat{\mathbf{z}}_{ij} = (\hat{z}_{ij,1}, ..., \hat{z}_{ij,T-1})'$ and $\hat{\mathbf{Z}} = (\hat{z}'_{11}, ..., \hat{z}'_{1M}, ..., \hat{z}'_{N1}, ..., \hat{z}'_{NM})'$. Let $\mathbf{u}_{ij} = (u_{ij,2}, ..., u_{ij,T})'$, $\mathbf{U} = (\mathbf{u}'_{11}, ..., \mathbf{u}'_{1,M}, ..., \mathbf{u}'_{N,1}, ..., \mathbf{u}'_{NM})'$, and $\mathbf{z}_{ij} = (\ddot{u}_{ij,1}, ..., \ddot{u}_{ij,T-1})'$, where $\ddot{u}_{ijt} = u_{ijt} - \overline{u}_{i \cdot t} - \overline{u}_{ij \cdot} - \overline{u}_{\cdot jt} + \overline{u}_{i \cdot \cdot} + \overline{u}_{\cdot j \cdot} + \overline{u}_{\cdot \cdot t}$. Let $\ddot{y}_{ijt}, \overline{y}_{i \cdot \cdot}, \overline{y}_{\cdot j \cdot}, \overline{y}_{\cdot \cdot t}, \overline{y}_{i \cdot t}, \overline{y}_{ij \cdot}$, and $\overline{y}_{\cdot jt}$ be defined analogously to $\ddot{u}_{ijt}, \overline{u}_{i \cdot \cdot}, \overline{u}_{\cdot j \cdot}, \overline{u}_{\cdot \cdot t}, \overline{u}_{i \cdot t}, \overline{u}_{ij \cdot}$, and $\overline{u}_{\cdot jt}$. Hereafter let $\sum_{i,j,t} = \sum_{i=1}^{N} \sum_{j=1}^{M} \sum_{t=2}^{T}$, $\sum_{i,t} = \sum_{j=1}^{M} \sum_{t=2}^{T}$, $\sum_{j,t} = \sum_{j=1}^{M} \sum_{t=2}^{T}$, and $\sum_{t} = \sum_{t=2}^{T}$.

Lemma B.1 Let $\eta_{NMT} = (NMT)^{-1/2} + N^{-1} + M^{-1} + T^{-1}$. Suppose that the conditions in Theorem 4.1 hold. Then
   (i) $\frac{1}{NMT_1}(\hat{\mathbf{Z}}'\hat{\mathbf{Z}} - \mathbf{Z}'\mathbf{Z}) = O_p(\eta_{NMT})$;
   (ii) $\frac{1}{NMT_1}(\hat{\mathbf{Z}}'\hat{\mathbf{U}} - \mathbf{Z}'\mathbf{U}) = O_p(\eta_{NMT})$;
   (iii) $(\hat{\mathbf{Z}}'\hat{\mathbf{Z}})^{-1}\hat{\mathbf{Z}}'\hat{\mathbf{U}} - \rho = O_p(\eta_{NMT})$.

Lemma B.2 Let $\tilde{z}_{ijt,m} = z_{ijt,m} - \rho z_{ij,t-1,m}$ and $Q_m = \frac{1}{NMT_1} \sum_{i,j,t} v_{ijt} \tilde{z}'_{ijt,m} (Z'_m Z_m)^{-1} Z'_m U$ for $m = 1, 2, ..., 7$. Suppose that the conditions in Theorem 4.1 hold. Then
   (i) $Q_1 = O_p((NMT)^{-1})$;
   (ii) $Q_2 = \frac{1}{NMT_1} \sum_{i,j,t} v_{ijt} [(1-\rho)(\overline{u}_{i \cdot \cdot} + \overline{u}_{\cdot j \cdot}) + (1-\rho L)\overline{u}_{\cdot \cdot t}] + O_p((NM)^{-2} + (NT)^{-2} + (MT)^{-2} + (NMT)^{-1})$;
   (iii) $Q_3 = \frac{1-\rho}{NMT_1} \sum_{i,j,t} v_{ijt} \overline{u}_{ij \cdot} + O_p(T^{-2} + (NMT)^{-1})$;
   (iv) $Q_4 = \frac{1}{NMT_1} \sum_{i,j,t} v_{ijt} [(1-\rho)\overline{u}_{ij \cdot} + (1-\rho L)\overline{u}_{\cdot \cdot t}] + O_p((NM)^{-2} + T^{-2} + (NMT)^{-1})$;
   (v) $Q_5 = \frac{1}{NMT_1} \sum_{i,j,t} v_{ijt} (1-\rho L)\overline{u}_{\cdot jt} + O_p(N^{-2} + (NMT)^{-1})$;
   (vi) $Q_6 = \frac{1}{NMT_1} \sum_{i,j,t} v_{ijt} (1-\rho L)(\overline{u}_{\cdot jt} + \overline{u}_{i \cdot t}) + O_p(N^{-2} + M^{-2} + (NMT)^{-1})$;
   (vii) $Q_7 = \frac{1}{NMT_1} \sum_{i,j,t} v_{ijt} [(1-\rho)\overline{u}_{ij \cdot} + (1-\rho L)(\overline{u}_{\cdot jt} + \overline{u}_{i \cdot t})] + O_p(N^{-2} + M^{-2} + T^{-2} + (NMT)^{-1})$.

Lemma B.3 Let Q_m be as defined in Lemma B.2. Suppose the conditions in Theorem 4.1 hold.
   (i) If Model 2 is the true model, $Q_4 - Q_2 = \frac{1-\rho}{NM} \sum_{i,j} \overline{v}_{ij \cdot} \overline{u}_{ij \cdot} + o_p(T^{-1})$;
   (ii) If Model 3 is the true model, $Q_4 - Q_3 = \frac{1}{NMT_1} \sum_{i,j,t} v_{ijt} (1-\rho L)\overline{u}_{\cdot \cdot t} + o_p((NM)^{-1})$;
   (iii) If Model 3 is the true model, $Q_7 - Q_3 = \frac{1}{NMT_1} \sum_{i,j,t} v_{ijt} (1-\rho L)(\overline{u}_{\cdot jt} + \overline{u}_{i \cdot t}) + o_p(N^{-1} + M^{-1})$;
   (iv) If Model 4 is the true model, $Q_7 - Q_4 = \frac{1}{NMT_1} \sum_{i,j,t} v_{ijt} (1-\rho L)(\overline{u}_{\cdot jt} + \overline{u}_{i \cdot t}) + o_p(N^{-1} + M^{-1})$;
   (v) If Model 5 is the true model, $Q_6 - Q_5 = \frac{1}{NMT_1} \sum_{i,j,t} v_{ijt} (1-\rho L)\overline{u}_{i \cdot t} + o_p(M^{-1})$;
   (vi) If Model 5 is the true model, $Q_7 - Q_5 = \frac{1}{NMT_1} \sum_{i,j,t} v_{ijt} [(1-\rho)\overline{u}_{ij \cdot} + (1-\rho L)\overline{u}_{i \cdot t}] + o_p(T^{-1} + M^{-1})$;
   (vii) If Model 6 is the true model, $Q_7 - Q_6 = \frac{1-\rho}{NMT_1} \sum_{i,j,t} v_{ijt} \overline{u}_{ij \cdot} + o_p(T^{-1})$.

Lemma B.4 *Let* $L_m = \frac{1}{NMT_1} \sum_{i,j,t} [\tilde{z}'_{ijt,m}(Z'_m Z_m)^{-1} Z'_m U]^2$ *for* $m = 1, 2, ..., 7$. *Suppose that the conditions in Theorem 4.1 hold. Then*

(i) $L_1 = O_p((NMT)^{-1})$;

(ii) $L_2 = \frac{1}{NMT_1} \sum_{i,j,t} \{(1-\rho)^2 (\bar{u}_{i\cdot\cdot}^2 + \bar{u}_{\cdot j\cdot}^2) + [(1-\rho L)\bar{u}_{\cdot\cdot t}]^2\} + O_p((NMT)^{-1} + (NM)^{-2} + (NT)^{-2} + (MT)^{-2})$;

(iii) $L_3 = \frac{(1-\rho)^2}{NM} \sum_{i,j} \bar{u}_{ij\cdot}^2 + O_p(T^{-2} + (NMT)^{-1})$;

(iv) $L_4 = \frac{1}{NMT_1} \sum_{i,j,t} \{(1-\rho)^2 \bar{u}_{ij\cdot}^2 + [(1-\rho L)\bar{u}_{\cdot\cdot t}]^2\} + O_p((NM)^{-2} + T^{-2} + (NMT)^{-1})$;

(v) $L_5 = \frac{1}{MT_1} \sum_{j,t} [(1-\rho L)\bar{u}_{\cdot jt}]^2 + O_p(N^{-2} + (NMT)^{-1})$;

(vi) $L_6 = \frac{1}{NMT_1} \sum_{i,j,t} \{[(1-\rho L)\bar{u}_{i\cdot t}]^2 + [(1-\rho L)\bar{u}_{\cdot jt}]^2\} + O_p(N^{-2} + M^{-2} + (NMT)^{-1})$;

(vii) $L_7 = \frac{1}{NMT_1} \sum_{i,j,t} \{(1-\rho)^2 \bar{u}_{ij\cdot}^2 + [(1-\rho L)\bar{u}_{\cdot jt}]^2 + [(1-\rho L)\bar{u}_{i\cdot t}]^2\} + O_p(N^{-2} + M^{-2} + T^{-2} + (NMT)^{-1})$.

Lemma B.5 *Let* L_m *be as defined in Lemma B.4. Suppose that the conditions in Theorem 4.1 hold.*

(i) *If Model 2 is the true model,* $L_4 - L_2 = \frac{(1-\rho)^2}{NM} \sum_{i,j} \bar{u}_{ij\cdot}^2 + o_p(T^{-1})$;

(ii) *If Model 3 is the true model,* $L_4 - L_3 = \frac{1}{T_1} \sum_t [(1-\rho L)\bar{u}_{\cdot\cdot t}]^2 + o_p((NM)^{-1})$;

(iii) *If Model 3 is the true model,* $L_7 - L_3 = \frac{1}{NMT_1} \sum_{i,j,t} \{[(1-\rho L)\bar{u}_{\cdot jt}]^2 + [(1-\rho L)\bar{u}_{i\cdot t}]^2\} + o_p(N^{-1} + M^{-1})$;

(iv) *If Model 4 is the true model,* $L_7 - L_4 = \frac{1}{NMT_1} \sum_{i,j,t} \{[(1-\rho L)\bar{u}_{\cdot jt}]^2 + [(1-\rho L)\bar{u}_{i\cdot t}]^2\} + o_p(N^{-1} + M^{-1})$;

(v) *If Model 5 is the true model,* $L_6 - L_5 = \frac{1}{NT_1} \sum_{i,t} [(1-\rho L)\bar{u}_{i\cdot t}]^2 + o_p(M^{-1})$;

(vi) *If Model 5 is the true model,* $L_7 - L_5 = \frac{1}{NMT_1} \sum_{i,j,t} \{(1-\rho)^2 \bar{u}_{ij\cdot}^2 + [(1-\rho L)\bar{u}_{i\cdot t}]^2\} + o_p(T^{-1} + M^{-1})$;

(vii) *If Model 6 is the true model,* $L_7 - L_6 = \frac{(1-\rho)^2}{NM} \sum_{i,j} \bar{u}_{ij\cdot}^2 + o_p(T^{-1})$.

Proof of Theorem 4.1. For notational simplicity, we assume that $p = 1$. Let $CV^*_{m^*,m}$ demote $CV^*(m)$ when Model m^* is the true model. Noting that $(y_{ijt} - \hat{\rho} y_{ij,t-1}) - (\hat{y}_{ijt}^{(m)} - \hat{\rho}' \hat{y}_{ij,t-1}^{(m)}) = (y_{ijt} - \rho y_{ij,t-1}) - (\hat{y}_{ijt}^{(m)} - \rho \hat{y}_{ij,t-1}^{(m)}) + (\hat{\rho} - \rho)(\hat{y}_{ij,t-1}^{(m)} - y_{i,t-1})$, we can make the following decomposition

$$
\begin{aligned}
CV^*_{m^*,m} &= \frac{1}{NMT_1} \sum_{i,j,t} \left[(y_{ijt} - \hat{\rho} y_{ij,t-1}) - (\hat{y}_{ijt}^{(m)} - \hat{\rho} \hat{y}_{ij,t-1}^{(m)}) \right]^2 \\
&= \frac{1}{NMT_1} \sum_{i,j,t} \left[(y_{ijt} - \rho y_{ij,t-1}) - (\hat{y}_{ijt}^{(m)} - \rho \hat{y}_{ij,t-1}^{(m)}) \right]^2 + \frac{(\hat{\rho} - \rho)^2}{NMT_1} \sum_{i,j,t} (\hat{y}_{ij,t-1}^{(m)} - y_{ij,t-1})^2 \\
&\quad + \frac{2(\hat{\rho} - \rho)}{NMT_1} \sum_{i,j,t} \left[(y_{ijt} - \rho y_{ij,t-1}) - (\hat{y}_{ijt}^{(m)} - \rho \hat{y}_{ij,t-1}^{(m)}) \right] (\hat{y}_{ij,t-1}^{(m)} - y_{ij,t-1}) \\
&\equiv CV^*_{m^*,m}(1) + CV^*_{m^*,m}(2) + CV^*_{m^*,m}(3), \text{ say.} \tag{B.1}
\end{aligned}
$$

We prove the theorem by considering all seven cases where Model m^* is the true model for $m^* = 1, 2, ..., 7$.

Case 1: Model 1 is the true Model. In this case, Models 2-7 are all over-fitted and we will show that $P(CV^*_{1,m} > CV^*_{1,1}) \to 1$ for $m = 2, ..., 7$ as $(N, M, T) \to \infty$. By Lemma A.7(i), we have

$$
\begin{aligned}
&(y_{ijt} - \rho y_{ij,t-1}) - (\hat{y}_{ijt}^{(m)} - \rho \hat{y}_{ij,t-1}^{(m)}) \\
&= c_{ijt,m}[u_{ijt} - z'_{ijt,m}(Z'_m Z_m)^{-1} Z'_m U] - \rho c_{ij,t-1,m}[u_{ijt-1} - z'_{ij,t-1,m}(Z'_m Z_m)^{-1} Z'_m U] \\
&= c_{ijt,m}[v_{ijt} - \tilde{z}'_{ijt,m}(Z'_m Z_m)^{-1} Z'_m U] + \rho \varkappa_{ijt,m}[u_{ijt-1} - z'_{ij,t-1,m}(Z'_m Z_m)^{-1} Z'_m U], \tag{B.2}
\end{aligned}
$$

2

where $\varkappa_{ijt,m} = c_{ijt,m} - c_{ij,t-1,m}$, $\tilde{z}_{ijt,m} = z_{ijt,m} - \rho z_{ij,t-1,m}$. By Lemma A.6, $\max_{i,j,t} |\varkappa_{ijt,m}| = o_p(1)$. For $CV_{1,m}^*(1)$ we make the following decomposition

$$
\begin{aligned}
CV_{1,m}^*(1) &= \frac{1}{NMT_1} \sum_{i,j,t} c_{ijt,m}^2 [v_{ijt} - \tilde{z}_{ijt,m}'(Z_m'Z_m)^{-1}Z_m'U]^2 \\
&\quad + \frac{\rho^2}{NMT_1} \sum_{i,j,t} \varkappa_{ijt,m}^2 [u_{ijt-1} - z_{ij,t-1,m}'(Z_m'Z_m)^{-1}Z_m'U]^2 \\
&\quad + \frac{2\rho}{NMT_1} \sum_{i,j,t} c_{ijt,m}\varkappa_{ijt,m}[u_{ijt-1} - z_{ij,t-1,m}'(Z_m'Z_m)^{-1}Z_m'U][v_{ijt} - \tilde{z}_{ijt,m}'(Z_m'Z_m)^{-1}Z_m'U] \\
&= CV_{1,m}^*(1,1) + CV_{1,m}^*(1,2) + CV_{1,m}^*(1,3), \text{ say.}
\end{aligned}
$$

Following the study of $CV_{1,1}$ in the proof of Theorem 3.1, we can show $CV_{1,1}^*(1,1) = \frac{1}{NMT} \sum_{i,j,t} v_{ijt}^2 + O_p((NMT)^{-1})$. We shall study

$$
\begin{aligned}
\Phi_{1,2}(1,1) &= \frac{1}{NMT_1} \sum_{i,j,t} \left\{ (1-\rho)^2 (\overline{u}_{i\cdot\cdot}^2 + \overline{u}_{\cdot j\cdot}^2) + \overline{v}_{\cdot\cdot t}^2 - 2v_{ijt}[(1-\rho)(\overline{u}_{i\cdot\cdot} + \overline{u}_{\cdot j\cdot}) + \overline{v}_{\cdot\cdot t}] \right\} + o_p(\eta_{1NMT}) \\
&= -\frac{1}{N} \sum_i \overline{v}_{i\cdot\cdot}^2 - \frac{1}{M} \sum_j \overline{v}_{\cdot j\cdot}^2 - \frac{1}{T_1} \sum_t \overline{v}_{\cdot\cdot t}^2 + o_p(\eta_{1NMT}), \\
\Phi_{1,2}(1,2) &= \frac{2 + o_p(1)}{NMT_1} \sum_{i,j,t} (\overline{d}_2 + h_{ijt,2}^*)[v_{ijt} - \tilde{z}_{ijt,2}'(Z_2'Z_2)^{-1}Z_2'U]^2 \\
&= \frac{2 + o_p(1)}{NMT_1} \sum_{i,j,t} \eta_{1NMT}[v_{ijt} - \tilde{z}_{ijt,2}'(Z_2'Z_2)^{-1}Z_2'U]^2 + o_p(\eta_{1NMT}) \\
&= \frac{2\eta_{1NMT}}{NMT_1} \sum_{i,j,t} v_{ijt}^2 + o_p(\eta_{1NMT}), \text{ and} \\
\Phi_{1,2}(1,3) &= O_p((NMT)^{-1}),
\end{aligned}
$$

where $\eta_{1NMT} = (NM)^{-1} + (NT)^{-1} + (MT)^{-1}$, and we use the fact $\overline{v}_{i\cdot\cdot} = \frac{1}{MT_1} \sum_{j=1}^M \sum_{t=2}^T (u_{ijt} - \rho u_{ij,t-1}) = (1-\rho)\overline{u}_{i\cdot\cdot} + O_p((MT)^{-1})$ and that $\overline{v}_{\cdot j\cdot} = \frac{1}{NT_1} \sum_{i=1}^N \sum_{t=2}^T (u_{ijt} - \rho u_{ij,t-1}) = (1-\rho)\overline{u}_{\cdot j\cdot} + O_p((NT)^{-1})$. It follows that

$$
\begin{aligned}
&(NM \wedge NT \wedge MT)\left[CV_{1,2}^*(1,1) - CV_{1,1}^*(1,1)\right] \\
&= (NM \wedge NT \wedge MT)\left\{ \frac{2\eta_{1NMT}}{NMT_1} \sum_{i,j,t} v_{ijt}^2 - \frac{1}{N} \sum_i \overline{v}_{i\cdot\cdot}^2 - \frac{1}{M} \sum_j \overline{v}_{\cdot j\cdot}^2 - \frac{1}{T_1} \sum_t \overline{v}_{\cdot\cdot t}^2 \right\} + o_p(1) \\
&\xrightarrow{p} q_1(2\overline{\sigma}_v^2 - \overline{\sigma}_{v1}^2) + q_2(2\overline{\sigma}_v^2 - \overline{\sigma}_{v2}^2) + q_3(2\overline{\sigma}_v^2 - \overline{\sigma}_{v3}^2),
\end{aligned}
$$

where q_l's are defined as in the proof of Theorem 3.1. In addition, we can show that $CV_{1,m}^*(1,l) - CV_{1,1}^*(1,l) = o_p(\eta_{1NMT})$ for $l = 2, 3$, $CV_{1,m}^*(2) - CV_{1,1}^*(2) = (\hat{\rho} - \rho)^2 O_p(\eta_{1NMT})$, and $CV_{1,m}^*(3) - CV_{1,1}^*(3) = (\hat{\rho} - \rho) O_p(\eta_{1NMT})$. Consequently, we have

$$
(NM \wedge NT \wedge MT)(CV_{1,2}^* - CV_{1,1}^*) \xrightarrow{p} q_1(2\overline{\sigma}_v^2 - \overline{\sigma}_{v1}^2) + q_2(2\overline{\sigma}_v^2 - \overline{\sigma}_{v2}^2) + q_3(2\overline{\sigma}_v^2 - \overline{\sigma}_{v3}^2).
$$

3

Similarly, we can show that

$$\Phi_{1,3}(1,1) \;=\; \frac{1}{NM}\sum_{i,j}[(1-\rho)^2\,\overline{u}_{ij\cdot}^2 - 2\,(1-\rho)\,\overline{u}_{ij\cdot}\overline{v}_{ij\cdot}] + o_p(T^{-1}) = -\frac{1}{NM}\sum_{i,j}\overline{v}_{ij\cdot}^2 + o_p(T^{-1}),$$

$$\Phi_{1,3}(1,2) \;=\; \frac{2T^{-1}}{NMT_1}\sum_{i,j,t} v_{ijt}^2 + o_p(T^{-1}),$$

$$\Phi_{1,3}(1,3) \;=\; O_p((NMT)^{-1}),$$

$CV_{1,3}^*(1,l) - CV_{1,1}^*(1,l) = o_p\left(T^{-1}\right)$ for $l = 2,3$, $CV_{1,3}^*(2) - CV_{1,1}^*(2) = (\hat{\rho}-\rho)^2\,O_p\left(T^{-1}\right)$, and $CV_{1,3}^*(3) - CV_{1,1}^*(3) = (\hat{\rho}-\rho)\,O_p\left(T^{-1}\right)$, where we use the fact $\overline{v}_{ij\cdot} = \frac{1}{T_1}\sum_{t=2}^{T}(u_{ijt} - \rho u_{ij,t-1}) = (1-\rho)\overline{u}_{ij\cdot} + O_p(T^{-1})$. Then we have

$$T_1(CV_{1,3}^* - CV_{1,1}^*) = \frac{2}{NMT_1}\sum_{i,j,t} v_{ijt}^2 - \frac{T_1}{NM}\sum_{i,j}\overline{v}_{ij\cdot}^2 + o_p(1) \xrightarrow{p} 2\overline{\sigma}_v^2 - \overline{\sigma}_{v4}^2.$$

By the same token, we can show that

$$(NM \wedge T_1)(CV_{1,4}^* - CV_{1,1}^*) \;=\; (NM \wedge T_1)\left\{(T_1^{-1} + (NM)^{-1})\frac{2}{NMT_1}\sum_{i,j,t} v_{ijt}^2 - \frac{1}{NM}\sum_{i,j}\overline{v}_{ij\cdot}^2 - \frac{1}{T_1}\sum_{t}\overline{v}_{\cdot\cdot t}^2\right\}$$

$$+o_p(1)$$

$$\xrightarrow{p} q_4(2\overline{\sigma}_u^2 - \overline{\sigma}_{u4}^2) + q_5(2\overline{\sigma}_u^2 - \overline{\sigma}_{u3}^2),$$

$$N(CV_{1,5}^* - CV_{1,1}^*) \;=\; \frac{2}{NMT_1}\sum_{i,j,t} v_{ijt}^2 - \frac{N}{MT_1}\sum_{j,t}\overline{v}_{\cdot jt}^2 \xrightarrow{p} 2\overline{\sigma}_v^2 - \overline{\sigma}_{v6}^2,$$

$$(N \wedge M)(CV_{1,6}^* - CV_{1,1}^*) \;=\; (N \wedge M)\left\{(N^{-1} + M^{-1})\frac{2}{NMT_1}\sum_{i,j,t} v_{ijt}^2 - \frac{1}{MT_1}\sum_{j,t}\overline{v}_{\cdot jt}^2 - \frac{1}{NT_1}\sum_{i,t}\overline{v}_{i\cdot t}^2\right\}$$

$$+o_p(1)$$

$$\xrightarrow{p} q_6(2\overline{\sigma}_v^2 - \overline{\sigma}_{v6}^2) + q_7(2\overline{\sigma}_v^2 - \overline{\sigma}_{v5}^2),$$

and

$$(N \wedge M \wedge T_1)(CV_{1,7}^* - CV_{1,1}^*) \;=\; (N \wedge M \wedge T_1)\left\{(N^{-1} + M^{-1} + T_1^{-1})\frac{2}{NMT_1}\sum_{i,j,t} v_{ijt}^2 - \frac{1}{NM}\sum_{i,j}\overline{v}_{ij\cdot}^2\right.$$

$$\left. -\frac{1}{NT_1}\sum_{i,t}\overline{v}_{i\cdot t}^2 - \frac{1}{MT_1}\sum_{j,t}\overline{v}_{\cdot jt}^2\right\} + o_p(1)$$

$$\xrightarrow{p} q_8(2\overline{\sigma}_v^2 - \overline{\sigma}_{v4}^2) + q_9(2\overline{\sigma}_v^2 - \overline{\sigma}_{v5}^2) + q_{10}(2\overline{\sigma}_v^2 - \overline{\sigma}_{v6}^2).$$

It follows that $P(CV_{1,m}^* > CV_{1,1}^*) \to 1$ for $m = 2, 3, ...7$.

Case 2: Model 2 is the true model. In this case, Model 1, 3 and 5 are under-fitted and Model 4, 6 and 7 are over-fitted. We will show that $P(CV_{2,m}^* > CV_{2,2}^*) \to 1$ for $m = 1, 3, 4, 5, 6, 7$.

First, we consider the under-fitted case. We will only show that $P(CV_{2,1}^* > CV_{2,2}^*) \to 1$ as the proof for the $m = 3, 5$ case follows similar arguments. By Lemma A.7(i), we have

$$(y_{ijt} - \rho y_{ij,t-1}) - (\hat{y}_{ijt}^{(1)} - \rho\hat{y}_{ij,t-1}^{(1)}) \;=\; c_{ijt,1}[u_{ijt} + \alpha_i + \gamma_j + \lambda_t - x_{ijt}'(X'X)^{-1}X'U^{(2)}]$$

$$-\rho c_{ij,t-1,1}[u_{ij,t-1} + \alpha_i + \gamma_j + \lambda_{t-1} - x_{ij,t-1}'(X'X)^{-1}X'U^{(2)}]$$

$$=\; c_{ijt,1}[v_{ijt} + (1-\rho)(\alpha_i + \gamma_j) + (1-\rho L)\lambda_t - \tilde{x}_{ijt}'(X'X)^{-1}X'U^{(2)}]$$

$$+\rho\varkappa_{ijt,1}[u_{ijt-1} + \alpha_i + \gamma_j + \lambda_{t-1} - x_{ij,t-1}'(X'X)^{-1}X'U^{(2)}],$$

4

where $U^{(2)} = U + D_2\pi_2$. We make the following decomposition for $CV_{2,1}^*(1)$:

$$
\begin{aligned}
CV_{2,1}^*(1) &= \frac{1}{NMT_1}\sum_{i,j,t} c_{ijt,1}^2[v_{ijt} + (1-\rho)(\alpha_i+\gamma_j) + (1-\rho L)\lambda_t - \tilde{x}_{ijt}'(X'X)^{-1}X'U^{(2)}]^2 \\
&\quad + \frac{\rho^2}{NMT_1}\sum_{i,j,t} \varkappa_{ijt,1}^2[u_{ijt-1} + \alpha_i + \gamma_j + \lambda_{t-1} - x_{ij,t-1}'(X'X)^{-1}X'U^{(2)}]^2 \\
&\quad + \frac{2\rho}{NMT_1}\sum_{i,j,t} c_{ijt,1}\varkappa_{ijt,1}[v_{ijt} + (1-\rho)(\alpha_i+\gamma_j) + (1-\rho L)\lambda_t - \tilde{x}_{ijt}'(X'X)^{-1}X'U^{(2)}] \\
&\quad \times [u_{ijt-1} + \alpha_i + \gamma_j + \lambda_{t-1} - x_{ij,t-1}'(X'X)^{-1}X'U^{(2)}] \\
&\equiv CV_{2,1}^*(1,1) + CV_{2,1}^*(1,2) + CV_{2,1}^*(1,3), \text{ say.}
\end{aligned}
$$

By Assumptions A.5 and A.7, we can show

$$
\begin{aligned}
CV_{2,1}^*(1,1) &= \frac{1}{NMT_1}\sum_{i,j,t} v_{ijt}^2 + \frac{1}{NMT_1}\sum_{i,j,t}[(1-\rho)(\alpha_i+\gamma_j) + (1-\rho L)\lambda_t - \tilde{x}_{ijt}'(X'X)^{-1}X'D_2\pi_2]^2 \\
&\quad + o_p(1) \\
&\xrightarrow{p} \overline{\sigma}_v^2 + \varphi_{2,1}^*.
\end{aligned}
$$

Noting that $\max_{i,j,t}|\varkappa_{ijt,1}| = o_p(1)$, we can readily show that $CV_{2,1}^*(1,l) = o_p(1)$ for $l = 2,3$. Then $CV_{2,1}^*(1) \xrightarrow{p} \overline{\sigma}_v^2 + \varphi_{2,1}^*$. In addition, using the fact that $\hat{\rho} - \rho = o_p(1)$ and following the analysis of $CV_{2,1}$, we can readily show that $CV_{2,1}^*(l) = o_p(1)$ for $l = 2,3$. Consequently, we have shown that $CV_{2,1}^* = \overline{\sigma}_v^2 + \varphi_{2,1}^* + o_p(1)$. For $CV_{2,2}^*$, it is easy to show $CV_{2,2}^* = \overline{\sigma}_v^2 + o_p(1)$. It follows that $CV_{2,1}^* - CV_{2,2}^* \to \varphi_{2,1}^* > 0$. Analogously, we can show that
$$CV_{2,m}^* - CV_{2,2}^* \to \varphi_{2,m}^* > 0 \text{ for } m = 3,5.$$

Now, we consider the over-fitted case. We focus on showing that $P(CV_{2,4}^* > CV_{2,2}^*) \to 1$ as the other over-fitted cases are similar. By (B.1) and applying similar arguments as used in the analysis of $CV_{1,2}^* - CV_{1,1}^*$, we will show that $T_1[CV_{2,4}^*(1) - CV_{2,2}^*(1)] \to 2\overline{\sigma}_v^2 - \overline{\sigma}_{v4}^2$ and $T_1[CV_{2,4}^*(l) - CV_{2,2}^*(l)] = o_p(1)$ for $l = 2,3$. Noting that when Model 2 is the true model and Model $m = 2,4,6,7$ are used, we have

$$
y_{ijt} - \hat{y}_{ijt}^{(m)} = c_{ijt,m}[y_{ijt} - z_{ijt,m}(Z_m'Z_m)^{-1}Z_m'Y] = c_{ijt,m}[u_{ijt} - z_{ijt,m}(Z_m'Z_m)^{-1}Z_m'U],
$$

where we use the fact that $x_{ijt}'\beta + \alpha_i + \gamma_j + \lambda_t = z_{ijt,m}'\theta_m$ and $Y = Z_2\theta_2 + U = Z_m\theta_m + U$ when Model 2 is the true model and Model m is just- or over-fitted. In particular, when Model m is over-fitted, some elements in θ_m corresponding to the redundant columns in Z_m have true value zero. Then for $m = 2,4,6,7$ we have

$$
\begin{aligned}
&(y_{ijt} - \rho y_{ij,t-1}) - (\hat{y}_{ijt}^{(m)} - \hat{\rho}\hat{y}_{ij,t-1}^{(m)}) \\
&= c_{ijt,m}[u_{ijt} - z_{ijt,m}'(Z_m'Z_m)^{-1}Z_m'U] - \rho c_{ij,t-1,m}[u_{ij,t-1} - z_{ij,t-1,m}'(Z_m'Z_m)^{-1}Z_m'U] \\
&= c_{ijt,m}[v_{ijt} - \tilde{z}_{ijt,m}'(Z_m'Z_m)^{-1}Z_m'U] + \rho\varkappa_{ijt,m}[u_{ijt-1} - z_{ij,t-1,m}'(Z_m'Z_m)^{-1}Z_m'U], \quad\quad \text{(B.3)}
\end{aligned}
$$

and

$$
\begin{aligned}
CV_{2,m}^*(1) &= \frac{1}{NMT_1}\sum_{i,j,t} c_{ijt,m}^2[v_{ijt} - \tilde{z}_{ijt,m}'(Z_m'Z_m)^{-1}Z_m'U]^2 \\
&\quad + \frac{\rho^2}{NMT_1}\sum_{i,j,t} \varkappa_{ijt,m}^2[u_{ijt-1} - z_{ij,t-1,m}'(Z_m'Z_m)^{-1}Z_m'U]^2 \\
&\quad + \frac{2\rho}{NMT_1}\sum_{i,j,t} c_{ijt,m}\varkappa_{ijt,m}[v_{ijt} - \tilde{z}_{ijt,m}'(Z_m'Z_m)^{-1}Z_m'U][u_{ijt-1} - z_{ij,t-1,m}'(Z_m'Z_m)^{-1}Z_m'U] \\
&= CV_{2,m}^*(1,1) + CV_{2,m}^*(1,2) + CV_{2,m}^*(1,3), \text{ say.}
\end{aligned}
$$

5

By Lemmas A.4(i) and (iii) A.6(ii), B.3(i) and B.5(i), we have

$$
\begin{aligned}
T_1[CV_{2,4}^*(1,1) - CV_{2,2}^*(1,1)] &= \frac{2}{NMT_1}\sum_{i,j,t} v_{ijt}^2 + \frac{(1-\rho)^2 T_1}{NM}\sum_{i,j}\overline{u}_{ij\cdot}^2 - \frac{2(1-\rho)T_1}{NM}\sum_{i,j}\overline{v}_{ij\cdot}\overline{u}_{ij\cdot} \\
&\quad + o_p(1) \\
&= \frac{2}{NMT_1}\sum_{i,j,t} v_{ijt}^2 - \frac{2T_1}{NM}\sum_{i,j}\overline{v}_{ij\cdot}^2 + o_p(1) \\
&\xrightarrow{p} 2\overline{\sigma}_v^2 - \overline{\sigma}_{v4}^2,
\end{aligned}
$$

where we use the fact $\overline{v}_{ij\cdot} = \frac{1}{T_1}\sum_{t=2}^{T}(u_{ijt} - \rho u_{ij,t-1}) = (1-\rho)\overline{u}_{ij\cdot} + O_p(T^{-1})$. In addition, using the fact that $\max_{i,j,t}|\varkappa_{ijt,4}| = o_p(1)$ and that $\hat{\rho} - \rho = o_p(1)$, we can also show that $CV_{2,4}^*(1,l) - CV_{2,2}^*(1,l) = o_p(T^{-1})$ and $CV_{2,4}^*(l) - CV_{2,2}^*(l)$ for $l = 2, 3$. Thus we have

$$
T_1(CV_{2,4}^* - CV_{2,2}^*) \xrightarrow{p} 2\overline{\sigma}_v^2 - \overline{\sigma}_{v4}^2.
$$

Analogously, we can apply Lemmas A.4(i) and (v)-(vi), A.6(ii), B.2(ii) and (vi)-(vii) and B.4(ii) and (vi)-(vii) and show that

$$
\begin{aligned}
(N \wedge M)(CV_{2,6}^* - CV_{2,2}^*) &= (N \wedge M)\left(\frac{1}{N} + \frac{1}{M}\right)\frac{2}{NMT}\sum_{i,j,t} v_{ijt}^2 - \frac{1}{MT}\sum_{j,t}\overline{v}_{\cdot jt}^2 - \frac{1}{NT}\sum_{i,t}\overline{v}_{i\cdot t}^2 + o_p(1) \\
&\xrightarrow{p} q_6(2\overline{\sigma}_v^2 - \overline{\sigma}_{v6}^2) + q_7(2\overline{\sigma}_v^2 - \overline{\sigma}_{v5}^2),
\end{aligned}
$$

and

$$
\begin{aligned}
(N \wedge M \wedge T)(CV_{2,7}^* - CV_{2,2}^*) &= (N \wedge M \wedge T)\left\{\left(\frac{1}{N} + \frac{1}{M} + \frac{1}{T}\right)\frac{2}{NMT}\sum_{i,j,t} v_{ijt}^2 - \frac{1}{NM}\sum_{i,j}\overline{v}_{ij\cdot}^2 \right. \\
&\quad \left. -\frac{1}{NT}\sum_{i,t}\overline{v}_{i\cdot t}^2 - \frac{1}{MT}\sum_{j,t}\overline{v}_{\cdot jt}^2\right\} + o_p(1) \\
&\xrightarrow{p} q_8(2\overline{\sigma}_u^2 - \overline{\sigma}_{u4}^2) + q_9(2\overline{\sigma}_u^2 - \overline{\sigma}_{u5}^2) + q_{10}(2\overline{\sigma}_u^2 - \overline{\sigma}_{u6}^2).
\end{aligned}
$$

Consequently we have $P(CV_{2,m}^* > CV_{2,2}^*) \to 1$ as $(N, M, T) \to \infty$ for $m = 1, 3, 4, 5, 6, 7$.

Cases 3-6: Model 3, 4, 5, or 6 is the true model. The proof is analogous to that of Case 2 and thus omitted.

Case 7: Model 7 is the true model. In this case, Models 1-6 are all under-fitted. Noting that $y_{ijt} - \hat{y}_{ijt}^{(m)} = c_{ijt,1}[u_{ijt} + \gamma_{ij} + \alpha_{it} + \alpha_{jt}^* + x_{ijt}'\beta - z_{ijt,m}'(Z_m'Z_m)^{-1}Z_m'Y]$, we have

$$
\begin{aligned}
&(y_{ijt} - \rho y_{ij,t-1}) - (\hat{y}_{ijt}^{(m)} - \rho\hat{y}_{ij,t-1}^{(m)}) \\
&= c_{ijt,m}[v_{ijt} + (1-\rho)\gamma_{ij} + (1-\rho L)(\alpha_{it} + \alpha_{jt}^*) + \tilde{x}_{ijt}'\beta - \tilde{z}_{ijt,m}'(Z_m'Z_m)^{-1}Z_m'Y] \\
&\quad + \rho\varkappa_{ijt,m}[u_{ij,t-1} + \gamma_{ij} + \alpha_{i,t-1} + \alpha_{j,t-1}^* - x_{ij,t-1}'\beta - z_{ij,t-1,m}'(Z_m'Z_m)^{-1}Z_m'Y],
\end{aligned}
$$

and

$$CV_{7,m}^*(1)$$
$$= \frac{1}{NMT_1} \sum_{i,j,t} c_{ijt,m}^2 [v_{ijt} + (1-\rho)\gamma_{ij} + (1-\rho L)(\alpha_{it} + \alpha_{jt}^*) + \tilde{x}_{ijt}'\beta - \tilde{z}_{ijt,m}'(Z_m'Z_m)^{-1}Z_m'Y]^2$$
$$+ \frac{\rho^2}{NMT_1} \sum_{i,j,t} \varkappa_{ijt,m}^2 [u_{ij,t-1} + \gamma_{ij} + \alpha_{i,t-1} + \alpha_{j,t-1}^* - x_{ij,t-1}'\beta - z_{ij,t-1,m}'(Z_m'Z_m)^{-1}Z_m'Y]^2$$
$$+ \frac{2\rho}{NMT_1} \sum_{i,j,t} c_{ijt,m}\varkappa_{ijt,m}[v_{ijt} + (1-\rho)\gamma_{ij} + (1-\rho L)(\alpha_{it} + \alpha_{jt}^*) + \tilde{x}_{ijt}'\beta - \tilde{z}_{ijt,m}'(Z_m'Z_m)^{-1}Z_m'Y]$$
$$\times [u_{ij,t-1} + \gamma_{ij} + \alpha_{i,t-1} + \alpha_{j,t-1}^* - x_{ij,t-1}'\beta - z_{ij,t-1,m}'(Z_m'Z_m)^{-1}Z_m'Y]$$
$$\equiv CV_{7,m}^*(1,1) + CV_{7,m}^*(1,2) + CV_{2,m}^*(1,3), \text{ say.}$$

Noting that $Y = X\beta + D_7\underline{\pi}_7 + U$ when Model 7 is the true model and $z_{ijt,m}'(Z_m'Z_m)^{-1}Z_m'X\beta = d_{ijt,m}'(D_m'D_m)^{-1}$ $\times D_m'X\beta + [x_{ijt} - X'D_m(D_m'D_m)^{-1}d_{ijt,m}]' X_{D_m}^* [X'X\beta - X'D_m(D_m'D_m)^{-1}D_m'X\beta]$ with $X_{D_m}^* = (X'M_{D_m}X)^{-1}$, we can readily apply Lemma A.6 and Assumptions A.5 and A.7 to show that

$$CV_{7,m}^*(1,1) - CV_{7,7}^*(1,1) = \frac{1}{NMT_1} \sum_{i,j,t} [(1-\rho)\gamma_{ij} + (1-\rho L)(\alpha_{it} + \alpha_{jt}^*) - \tilde{z}_{ijt,m}'(Z_m'Z_m)^{-1}Z_m'D_7\underline{\pi}_7]^2$$
$$+ o_p(1)$$
$$\xrightarrow{p} \varphi_{7,m}^* > 0 \text{ for } m = 1,2,...,6.$$

Using the fact that $\max_{i,j,t} |\varkappa_{ijt,m}| = o_p(1)$ and that $\hat{\rho} - \rho = o_p(1)$, we can also show that $CV_{7,m}^*(1,l) - CV_{7,7}^*(1,l) = o_p(1)$ and $CV_{7,m}^*(l) - CV_{7,7}^*(l)$ for $l = 2,3$ and $m = 1,2,...,6$. It follows that $CV_{7,m}^* - CV_{7,7}^* \xrightarrow{p} \varphi_{7,m}^* > 0$ and $P(CV_{7,m}^* > CV_{7,7}^*) \to 1$ as $(N,M,T) \to \infty$ for $m = 1,2,...,6$. ∎

C Proofs of the technical lemmas in Appendices A and B

Proof of Lemma A.1. Noting that $X_D'X_D = \begin{pmatrix} X'X & X'D \\ D'X & D'D \end{pmatrix}$, the result follows from the inversion formula for a 2×2 partitioned matrix. See, e.g., Bernstein (2005, p.45). One can also verify the result by definition. ∎

Proof of Lemma A.2. (i) Noting that $\iota_a' \begin{bmatrix} I_{a-1} \\ -\iota_{a-1}' \end{bmatrix} = 0$ for $a = N, M, T$, we can readily show that $D_I'D_J = 0$, $D_I'D_T = 0$, and $D_J'D_T = 0$ by using the property that $(A_1 \otimes A_2 \otimes A_3)'(B_1 \otimes B_2 \otimes B_3) = A_1'B_1 \otimes A_2'A_2 \otimes A_3'B_3$ for conformable matrices A_1, A_2, A_3, B_1, B_2, and B_3. Noting that $\iota_N \otimes \iota_M = \iota_{NM}$, we have

$$D_{IJ}'D_T = \left(\begin{bmatrix} I_{NM-1} & -\iota_{NM-1} \end{bmatrix} \otimes \iota_T' \right) \left(\iota_{NM} \otimes \begin{bmatrix} I_{T-1} \\ -\iota_{T-1}' \end{bmatrix} \right)$$
$$= \begin{bmatrix} I_{NM-1} & -\iota_{NM-1} \end{bmatrix} \iota_{NM} \otimes \iota_T' \begin{bmatrix} I_{T-1} \\ -\iota_{T-1}' \end{bmatrix} = 0.$$

Similarly, we can show the other claims in (i).
   (ii) This follows from (i) directly.

(iii) Noting that $D_I'D_I = MT\left(I_{N-1} + J_{N-1}\right)$, we have $\left(D_I'D_I\right)^{-1} = \frac{1}{MT}(I_{N-1} - \frac{1}{N}J_{N-1})$. Then

$$
\begin{aligned}
P_I &= D_I\left(D_I'D_I\right)^{-1}D_I' \\
&= \frac{1}{MT}\left(\begin{bmatrix} I_{N-1} \\ -\iota_{N-1}' \end{bmatrix} \otimes \iota_{MT}\right)\left(I_{N-1} - \frac{1}{N}J_{N-1}\right)\left(\begin{bmatrix} I_{N-1} & -\iota_{N-1} \end{bmatrix} \otimes \iota_{MT}'\right) \\
&= \left(I_N - \frac{J_N}{N}\right) \otimes \frac{J_M}{M} \otimes \frac{J_T}{T}.
\end{aligned}
$$

Similarly, other parts in (iii) follow.

(iv) Noting that $D_{IJ}'D_{IJ} = T\left(I_{NM-1} + J_{NM-1}\right)$, we have $\left(D_{IJ}'D_{IJ}\right)^{-1} = \frac{1}{T}(I_{NM-1} - \frac{1}{NM}J_{NM-1})$. Then

$$
\begin{aligned}
P_{IJ} &= D_{IJ}\left(D_{IJ}'D_{IJ}\right)^{-1}D_{IJ}' \\
&= \frac{1}{T}\left(\begin{bmatrix} I_{NM-1} \\ -\iota_{NM-1}' \end{bmatrix} \otimes \iota_T\right)\left(I_{NM-1} - \frac{1}{NM}J_{NM-1}\right)\left(\begin{bmatrix} I_{NM-1} & -\iota_{NM-1} \end{bmatrix} \otimes \iota_T'\right) \\
&= \left(I_{NM} - \frac{J_{NM}}{NM}\right) \otimes \frac{J_T}{T}.
\end{aligned}
$$

Similarly, we have $P_{JT} = \frac{J_N}{N} \otimes (I_{MT} - \frac{J_{MT}}{MT})$.

(v) Noting that $D_{IT}^{*\prime}D_{IT}^* = M\left(I_{N-1} + J_{N-1}\right) \otimes I_T$, we have $\left(D_{IT}^{*\prime}D_{IT}^*\right)^{-1} = \frac{1}{M}(I_{N-1} - \frac{1}{N}J_{N-1}) \otimes I_T$. Then

$$
\begin{aligned}
P_{IT}^* &= D_{IT}^*\left(D_{IT}^{*\prime}D_{IT}^*\right)^{-1}D_{IT}^{*\prime} \\
&= \frac{1}{M}\left(\begin{bmatrix} I_{N-1} \\ -\iota_{N-1}' \end{bmatrix} \otimes \iota_M \otimes I_T\right)\left(\left(I_{N-1} - \frac{1}{N}J_{N-1}\right) \otimes I_T\right)\left(\begin{bmatrix} I_{N-1} & -\iota_{N-1} \end{bmatrix} \otimes \iota_M' \otimes I_T\right) \\
&= \left(I_N - \frac{J_N}{N}\right) \otimes \frac{J_M}{M} \otimes I_T.
\end{aligned}
$$

Similarly, we can show the other two parts in (v). ∎

Proof of Lemma A.3. For (i), noting that $(I_N \otimes \frac{J_{MT}}{MT})A = (\bar{a}_{1..}, ..., \bar{a}_{N..})' \otimes \iota_{MT}$ and $\frac{J_{NMT}}{NMT}A = \bar{a} \cdot \iota_{NMT}$, we have

$$
\begin{aligned}
\frac{1}{NMT}A'P_IB &= \frac{1}{NMT}\left[A'(I_N \otimes \frac{J_{MT}}{MT}) - A'\frac{J_{NMT}}{NMT}\right]\left[(I_N \otimes \frac{J_{MT}}{MT})B - \frac{J_{NMT}}{NMT}B\right] \\
&= \frac{1}{NMT}A'(I_N \otimes \frac{J_{MT}}{MT})B - \frac{1}{NMT}A'\frac{J_{NMT}}{NMT}B \\
&= \frac{1}{N}\sum_{i=1}^N \bar{a}_{i..}\bar{b}_{i..}' - \bar{a}\bar{b}'.
\end{aligned}
$$

Similarly, we can show (ii) and (iii).

For (iv), noting that $(I_{NM} \otimes \frac{J_T}{T})A = (\bar{a}_{11.}, ..., \bar{a}_{1M.}, ..., \bar{a}_{N1.}, ..., \bar{a}_{NM.})' \otimes \iota_T$, we have

$$
\begin{aligned}
\frac{1}{NMT}A'P_{IJ}B &= \frac{1}{NMT}\left[A'(I_{NM} \otimes \frac{J_T}{T}) - A'\frac{J_{NMT}}{NMT}\right]\left[(I_{NM} \otimes \frac{J_T}{T})B - \frac{J_{NMT}}{NMT}B\right] \\
&= \frac{1}{NMT}A'(I_{NM} \otimes \frac{J_T}{T})B - \frac{1}{NMT}A'\frac{J_{NMT}}{NMT}B \\
&= \frac{1}{NM}\sum_{i,j} \bar{a}_{ij.}\bar{b}_{ij.}' - \bar{a}\bar{b}'.
\end{aligned}
$$

Similarly, we can prove (v).

For (vi), noting that $P_{IJ}^* = I_N \otimes (I_M - \frac{J_M}{M}) \otimes \frac{J_T}{T}$, we have

$$
\begin{aligned}
\frac{1}{NMT}A'P_{IJ}^*B &= \frac{1}{NMT}\left[A'(I_{NM} \otimes \frac{J_T}{T}) - A'(I_N \otimes \frac{J_{MT}}{MT})\right]\left[(I_{NM} \otimes \frac{J_T}{T})B - (I_N \otimes \frac{J_{MT}}{MT})B\right] \\
&= \frac{1}{NMT}A'(I_{NM} \otimes \frac{J_T}{T})B - \frac{1}{NMT}A'(I_N \otimes \frac{J_{MT}}{MT})B \\
&= \frac{1}{NM}\sum_{i,j}\overline{a}_{ij.}\overline{b}'_{ij.} - \frac{1}{N}\sum_{i=1}^{N}\overline{a}_{i..}\overline{b}'_{i..}
\end{aligned}
$$

Analogously, we can prove (vii) and (viii). ∎

Proof of Lemma A.4. (i) By Lemma A.2(ii) and Assumptions A.1(iv), A.2(i)-(iii) and A.3(i)-(iii), we have

$$
\begin{aligned}
\frac{1}{NMT}X'M_{D_2}U &= \frac{1}{NMT}X'U - \frac{1}{NMT}X'P_IU - \frac{1}{NMT}X'P_JU - \frac{1}{NMT}X'P_TU \\
&= \frac{1}{NMT}X'U - \frac{1}{N}\sum_{i=1}^{N}\overline{x}_{i..}\overline{u}_{i..} - \frac{1}{M}\sum_{j=1}^{M}\overline{x}_{.j.}\overline{u}_{.j.} - \frac{1}{T}\sum_{t=1}^{T}\overline{x}_{..t}\overline{u}_{..t} + 3\overline{xu} \\
&= O_p((NM)^{-1} + (MT)^{-1} + (NT)^{-1} + (NMT)^{-1/2}).
\end{aligned}
$$

(ii) Noting that $D_3 = D_{IJ}$, by Lemma A.2(iv) and Assumptions A.1(iv), A.2(iv) and A.3(iv) we have

$$
\begin{aligned}
\frac{1}{NMT}X'M_{D_3}U &= \frac{1}{NMT}X'U - \frac{1}{NMT}X'P_{IJ}U \\
&= \frac{1}{NMT}X'U - \frac{1}{NM}\sum_{i,j}\overline{x}_{ij.}\overline{u}_{ij.} + \overline{xu} \\
&= O_p(T^{-1} + (NMT)^{-1/2}).
\end{aligned}
$$

(iii) By Lemma A.2(ii) and Assumptions A.1(iv), A.2(iii)-(iv) and A.3(iii)-(iv) we have

$$
\begin{aligned}
\frac{1}{NMT}X'M_{D_4}U &= \frac{1}{NMT}X'U - \frac{1}{NMT}X'P_{IJ}U - \frac{1}{NMT}X'P_TU \\
&= \frac{1}{NMT}X'U - \frac{1}{NM}\sum_{i,j}\overline{x}_{ij.}\overline{u}_{ij.} - \frac{1}{T}\sum_{t}\overline{x}_{..t}\overline{u}_{..t} + 2\overline{xu} \\
&= O_p(T^{-1} + (NM)^{-1} + (NMT)^{-1/2}).
\end{aligned}
$$

(iv) Noting that $D_5 = D_{JT}$, by Lemma A.2(iv) and Assumptions A.1(iv), A.2(vi) and A.3(vi) we have

$$
\begin{aligned}
\frac{1}{NMT}X'M_{D_5}U &= \frac{1}{NMT}X'U - \frac{1}{NMT}X'P_{JT}U \\
&= \frac{1}{NMT}X'U - \frac{1}{MT}\sum_{j,t}\overline{x}_{.jt}\overline{u}_{.jt} + \overline{xu} \\
&= O_p((NMT)^{-1/2} + N^{-1}).
\end{aligned}
$$

(v) By Lemma A.2(iv) and Assumptions A.1(iv), A.2(ii) and (v)-(vi) and A.3(iii) and (v)-(vi) we have

$$
\begin{aligned}
\frac{1}{NMT}X'M_{D_6}U &= \frac{1}{NMT}X'U - \frac{1}{NMT}X'P_{IT}^*U - \frac{1}{NMT}X'P_{JT}U \\
&= \frac{1}{NMT}X'U - \frac{1}{NT}\sum_{i,t}\overline{x}_{i.t}\overline{u}_{i.t} + \frac{1}{T}\sum_{t}\overline{x}_{..t}\overline{u}_{..t} - \frac{1}{MT}\sum_{j,t}\overline{x}_{.jt}\overline{u}_{.jt} + \overline{xu} \\
&= O_p((NMT)^{-1/2} + M^{-1} + N^{-1}).
\end{aligned}
$$

9

(vi) By Lemma A.2(iv) and Assumptions A.1(iv), A.2 and A.3 we have

$$
\begin{aligned}
\frac{1}{NMT}X'M_{D_7}U &= \frac{1}{NMT}X'U - \frac{1}{NMT}X'P_{IT}^*U - \frac{1}{NMT}X'P_{JT}^*U - \frac{1}{NMT}X'P_{IJ}^*U \\
&= \frac{1}{NMT}X'U - \frac{1}{NT}\sum_{i,t}\overline{x}_{i\cdot t}\overline{u}_{i\cdot t} + \frac{1}{T}\sum_{t=1}^{T}\overline{x}_{\cdot\cdot t}\overline{u}_{\cdot\cdot t} - \frac{1}{MT}\sum_{j,t}\overline{x}_{\cdot jt}\overline{u}_{\cdot jt} \\
&\quad + \frac{1}{M}\sum_{j=1}^{M}\overline{x}_{\cdot j\cdot}\overline{u}_{\cdot j\cdot} - \frac{1}{NM}\sum_{i,j}\overline{x}_{ij\cdot}\overline{u}_{ij\cdot} + \frac{1}{N}\sum_{i=1}^{N}\overline{x}_{i\cdot\cdot}\overline{u}_{i\cdot\cdot} \\
&= O_p(N^{-1} + M^{-1} + T^{-1} + (NMT)^{-1/2}). \ \blacksquare
\end{aligned}
$$

Proof of Lemma A.5. Note that $d'_{ijt,m}(D'_mD_m)^{-1}d_{ijt,m}$ denotes the $\{(i-1)MT + (j-1)M + t\}$th diagonal element of P_{D_m} for $m = 2,...,7$. The form of P_{D_m} is given in Lemma A.2 (note that $P_{D_3} = P_{D_{IJ}}$ and $P_{D_5} = P_{D_{JT}}$), from which the results in (i)-(vi) follow immediately. $\blacksquare$

Proof of Lemma A.6. (i) Noting that $Z_m = (X, D_m)$, we can apply Lemma A.1 to obtain

$$
\begin{aligned}
&h_{ijt,m} \\
&= (x'_{ijt}, d'_{ijt,m})\begin{pmatrix} X_{D_m}^* & -X_{D_m}^*X'D_m(D'_mD_m)^{-1} \\ -(D'_mD_m)^{-1}D'_mXX_{D_m}^* & (D'_mD_m)^{-1} + (D'_mD_m)^{-1}D'_mXX_{D_m}^*X'D_m(D'_mD_m)^{-1} \end{pmatrix} \\
&\quad \times \begin{pmatrix} x_{ijt} \\ d_{ijt,m} \end{pmatrix} \\
&= d'_{ijt,m}(D'_mD_m)^{-1}d_{ijt,m} + [x_{ijt} - X'D_m(D'_mD_m)^{-1}d_{ijt,m}]'X_{D_m}^*[x_{ijt} - X'D_m(D'_mD_m)^{-1}d_{ijt,m}] \\
&= \bar{d}_m + h_{ijt,m}^*,
\end{aligned}
$$

where $X_{D_m}^* = (X'M_{D_m}X)^{-1}$.

(ii) Note that $D_m(D'_mD_m)^{-1}D'_m$ is a projection matrix with spectral norm 1 and $\bar{d}_m \equiv d'_{ijt,m}(D'_mD_m)^{-1} \times d_{ijt,m}$ is a constant which is $o(1)$ for each m by Lemma A.5

$$
\begin{aligned}
\frac{1}{NMT}\left\|X'D_m(D'_mD_m)^{-1}d_{ijt,m}\right\|^2 &= \frac{1}{NMT}d'_{ijt,m}(D'_mD_m)^{-1}D'_mX \ X'D_m(D'_mD_m)^{-1}d_{ijt,m} \\
&\leq \frac{\bar{d}_m}{NMT}\lambda_{\max}\left((D'_mD_m)^{-1/2}D'_mX \ X'D_m(D'_mD_m)^{-1/2}\right) \\
&\leq \frac{\bar{d}_m}{NMT}\text{tr}\left(XX'D_m(D'_mD_m)^{-1}D'_m\right) \\
&\leq \bar{d}_m\frac{1}{NMT}\|X'X\| = O_p\left(\bar{d}_m\right).
\end{aligned}
$$

By the Cauchy-Schwarz inequality and Assumption A.1(ii) and (v),

$$
\begin{aligned}
\max_{i,j,t}h_{ijt,m}^* &\leq 2c_{mNMT}\left[\frac{1}{NMT}\max_{i,j,t}\|x_{ijt}\|^2 + \frac{1}{NMT}\max_{i,j,t}\left\|d'_{ijt,m}(D'_mD_m)^{-1}D'_mX\right\|^2\right] \\
&= O_p((NMT)^{-1/2} + \bar{d}_m) = o_p(1),
\end{aligned}
$$

where $c_{mNMT} = [\lambda_{\min}(\frac{1}{NMT}X'M_{D_m}X)]^{-1} = O_p(1)$ by Assumption A.1(v).

(iii) Noting that $\max_{i,j,t}|c_{ijt,1} - 1| = \max_{i,j,t}\left|\frac{h_{ijt,m}}{1-h_{ijt,m}}\right| \leq \frac{\max_{i,j,t}h_{ijt,m}}{1-\max_{i,j,t}h_{ijt,m}}$, the result follows from part (ii).

(iv) This follows from the definition of $c_{ijt,1}$ and part (ii).

(v) This follows from the definition of $c_{ijt,1}$ and part (ii). $\blacksquare$

Proof of Lemma A.7. (i) For the model $y_{ijt} = x'_{ijt}\beta + d'_{ijt,m}\pi_m + u_{ijt} = z'_{ijt,m}\theta_m + u_{ijt}$, the OLS and leave-one-out OLS estimators of $\theta_m = (\beta', \pi'_m)'$ are given by $\hat{\theta}_m = (Z'_m Z_m)^{-1} Z'_m Y$ and $\hat{\theta}_{ijt,m} = (Z'_m Z_m - z_{ijt,m} z'_{ijt,m})^{-1} (Z'_m Y - z_{ijt,m} y_{ijt})$, respectively. By the updated formula for OLS estimation(e.g. Greene (2008, p.964)), we have $\hat{\theta}_{ijt,m} - \hat{\theta}_m = -\frac{1}{1-h_{ijt,m}} (Z'_m Z_m)^{-1} z_{ijt,m} e_{ijt,m}$. It follows that

$$
\begin{aligned}
y_{ijt} - \hat{y}_{ijt,m} &= y_{ijt} - z'_{ijt,m} \left(\hat{\theta}_m - \frac{1}{1-h_{ijt,m}} (Z'_m Z_m)^{-1} z_{ijt,m} e_{ijt,m} \right) \\
&= e_{ijt,m} + \frac{h_{ijt,m}}{1-h_{ijt,m}} e_{ijt,m} = \frac{e_{ijt,m}}{1-h_{ijt,m}}.
\end{aligned}
$$

(ii) When the true model is given by $Y = X\beta + D^*\pi^* + U$ but with $Z_m = (X, D_m)$ used in the regression, we have

$$
\begin{aligned}
\hat{\theta}_m &= (Z'_m Z_m)^{-1} Z'_m Y = (Z'_m Z_m)^{-1} \begin{pmatrix} X'X & X'D^* \\ D'_m X & D'_m D^* \end{pmatrix} \begin{pmatrix} \beta \\ \pi^* \end{pmatrix} + (Z'_m Z_m)^{-1} \begin{pmatrix} X'U \\ D'_m U \end{pmatrix} \\
&\equiv I + II, \text{ say.}
\end{aligned}
$$

By using the inverse formula in Lemma A.1, we can readily show that

$$
\begin{aligned}
I &= \begin{pmatrix} X^*_{D_m} & -X^*_{D_m} X'D_m (D'_m D_m)^{-1} \\ -(D'_m D_m)^{-1} D'_m X X^*_{D_m} & (D'_m D_m)^{-1} + (D'_m D_m)^{-1} D'_m X X^*_{D_m} X'D_m (D'_m D_m)^{-1} \end{pmatrix} \\
&\quad \times \begin{pmatrix} X'X\beta + X'D^*\pi^* \\ D'_m X\beta + D'_m D^*\pi^* \end{pmatrix} \\
&= \begin{pmatrix} \beta + (X'M_{D_m}X)^{-1}X'M_{D_m}D^*\pi^* \\ (D'_m D_m)^{-1}D'_m D^*\pi^* - (D'_m D_m)^{-1}D'_m X X^*_{D_m} X'M_{D_m}D^*\pi^* \end{pmatrix},
\end{aligned}
$$

and similarly

$$
II = \begin{pmatrix} (X'M_{D_m}X)^{-1}X'M_{Dp}U \\ (D'_m D_m)^{-1}D'_m U - (D'_m D_m)^{-1}D'_m X X^*_{D_m} X'M_{D_m}U \end{pmatrix},
$$

where $X^*_{D_m} = (X'M_{D_m}X)^{-1}$. It follows that

$$
\begin{aligned}
e_{ijt,m} &= y_{ijt} - z'_{ijt,m}\hat{\theta}_m = (x'_{ijt}\beta + d^{*'}_{i,j,t}\pi^* + u_{ijt}) - (x'_{ijt}, d'_{ijt,m})(I + II) \\
&= (x'_{ijt}\beta + d^{*'}_{i,j,t}\pi^* + u_{ijt}) - x'_{ijt}(\beta + X^*_{D_m}X'M_{D_m}D^*\pi^* + X^*_{D_m}X'M_{Dp}U) \\
&\quad - d'_{ijt,m}[(D'_m D_m)^{-1}D'_m D^*\pi^* - (D'_m D_m)^{-1}D'_m X X^*_{D_m} X'M_{D_m}D^*\pi^* \\
&\quad\quad + (D'_m D_m)^{-1}D'_m U - (D'_m D_m)^{-1}D'_m X X^*_{D_m} X'M_{D_m}U] \\
&= A_{ijt,m} + B_{ijt,m} + C_{ijt,m}.
\end{aligned}
$$

(iii) Noting that $A_{ijt,m}$, $B_{ijt,m}$, and $C_{ijt,m}$ are typical elements of $M_{D_m}U$, $(I - P_{M_{D_m}X})M_{D_m}D^*\pi^*$, and $-M_{D_m}X X^*_{D_m} X'M_{D_m}U$ respectively, we have

$$
\sum_{i,j,t} A^2_{ijt,m} = U'M_D U, \quad \sum_{i,j,t} B^2_{ijt,m} = \pi^{*'}D^{*'}M_{D_m}(I - P_{M_{D_m}X})M_{D_m}D^*\pi^*,
$$

$$
\sum_{i,j,t} C^2_{ijt,m} = U'M_{D_m}X X^*_{D_m} X'M_{D_m}U, \quad \sum_{i,j,t} A_{ijt,m}B_{ijt,m} = U'M_{D_m}(I - P_{M_{D_m}X})M_{D_m}D^*\pi^*,
$$

$$
\sum_{i,j,t} A_{ijt,m}C_{ijt,m} = -U'M_{D_m}X X^*_{D_m} X'M_{D_m}U, \text{ and } \sum_{i,j,t} B_{ijt,m}C_{ijt,m} = 0,
$$

where we also use the fact that $\sum_{i,j,t} u_{ijt}d'_{ijt,m} = U'D_m$, $\sum_{i,j,t} d_{ijt,m}d'_{ijt,m} = D'_m D_m$, $\sum_{i,j,t} x_{ijt,m}d'_{ijt,m} = X'D_m$, and $M_{D_m}X(I - P_{M_{D_m}X}) = 0$. $\blacksquare$

Proof of Lemma A.8. We first determine the probability order of $E_{m,1} \equiv \frac{1}{NMT} \sum_{i,j,t} c_{ijt,m}^2 C_{ijt,m}^2$ and then that of $E_{m,2} \equiv \frac{1}{NMT} \sum_{i,j,t} c_{ijt,m}^2 A_{ijt,m} C_{ijt,m}$. By Lemmas A.6(iii) and A.7(iii),

$$E_{m,1} \leq \frac{c_{2mNMT}}{NMT} \sum_{i,j,t} C_{ijt,m}^2 = c_{2mNMT} \frac{1}{NMT} U' M_{D_m} X X_{D_m}^* X' M_{D_m} U,$$

where $c_{2mNMT} = \max_{i,j,t} c_{ijt,m}^2 = 1 + o_p(1)$ by Lemma A.6(iii). One can readily show that by Assumptions A.1(iv)-(v) and Lemma A.4,

$$
\begin{aligned}
E_{1,1} &= O_p((NMT)^{-1}), \\
E_{2,1} &= O_p((NMT)^{-1} + (NM)^{-2} + (MT)^{-2} + (NT)^{-2}), \\
E_{3,1} &= O_p((NMT)^{-1} + T^{-2}), \\
E_{4,1} &= O_p((NMT)^{-1} + T^{-2} + (NM)^{-2}), \\
E_{5,1} &= O_p((NMT)^{-1} + N^{-2}), \\
E_{6,1} &= O_p((NMT)^{-1} + N^{-2} + M^{-2}), \\
E_{7,1} &= O_p((NMT)^{-1} + N^{-2} + M^{-2} + T^{-2}).
\end{aligned}
$$

For $E_{m,2} \equiv \frac{1}{NMT} \sum_{i,j,t} c_{ijt,m}^2 A_{ijt,m} C_{ijt,m}$, we make the following decomposition

$$
\begin{aligned}
E_{m,2} &= \frac{1}{NMT} \sum_{i,j,t} A_{ijt,m} C_{ijt,m} + \frac{2}{NMT} \sum_{i,j,t} h_{ijt,m} A_{ijt,m} C_{ijt,m} \\
&\quad + \frac{1}{NMT} \sum_{i,j,t} (c_{ijt,m}^2 - 1 - 2h_{ijt,m}) A_{ijt,m} C_{ijt,m} \\
&\equiv E_{m,2,1} + 2E_{m,2,2} + E_{m,2,3}, \text{ say.}
\end{aligned}
$$

By Lemma A.7(iii), $E_{m,2,1} = -\frac{1}{NMT} \sum_{i,j,t} C_{ijt,m}^2$ whose probability order is given above. By Lemma A.6(i),

$$E_{m,2,2} = \frac{\bar{d}_m}{NMT} \sum_{i,j,t} A_{ijt,m} C_{ijt,m} + \frac{1}{NMT} \sum_{i,j,t} h_{ijt,m}^* A_{ijt,m} C_{ijt,m} \equiv E_{m,2,2a} + E_{m,2,2b}, \text{ say.}$$

By the Cauchy-Schwarz inequality we have

$$|E_{m,2,2a}| \leq \bar{d}_m \left\{ \frac{1}{NMT} \sum_{i,j,t} A_{ijt,m}^2 \right\}^{1/2} \left\{ \frac{1}{NMT} \sum_{i,j,t} C_{ijt,m}^2 \right\}^{1/2} = O_p\left(\bar{d}_m^2 + \delta_{NMT,m} \right),$$

where $\delta_{NMT,m} \equiv \frac{1}{NMT} \sum_{i,j,t} C_{ijt,m}^2$ has the same probability order as $E_{m,1}$ studied above, the exact values of $\bar{d}_m$'s are given in Lemma A.5, and we use the fact that $\frac{1}{NMT} \sum_{i,j,t} A_{ijt,m}^2 = O_p(1)$ by Lemma A.10 below. For $E_{m,2,2b}$, we have

$$
\begin{aligned}
|E_{m,2,2b}| &\leq \max_{i,j,t} h_{ijt,m}^* \frac{1}{NMT} \sum_{i,j,t} |A_{ijt,m} C_{ijt,m}| \\
&= O_p\left(\bar{d}_m + (NMT)^{-1/2} \right) \left\{ \frac{1}{NMT} \sum_{i,j,t} A_{ijt,m}^2 \right\}^{1/2} \left\{ \frac{1}{NMT} \sum_{i,j,t} C_{ijt,m}^2 \right\}^{1/2} \\
&= O_p\left(\bar{d}_m^2 + (NMT)^{-1} \right) + O_p\left(\delta_{NMT,m} \right) = O_p\left(\bar{d}_m^2 + \delta_{NMT,m} \right).
\end{aligned}
$$

By Lemma A.6

$$
\begin{aligned}
|E_{m,2,3}| &\leq \frac{1}{NMT} \sum_{i,j,t} |(c_{ijt,m}^2 - 1 - 2h_{ijt,m}) A_{ijt,m} C_{ijt,m}| \\
&\leq [3 + o_p(1)] \max_{i,j,t} h_{ijt,m}^2 \frac{1}{NMT} \sum_{i,j,t} |A_{ijt,m} C_{ijt,m}| \\
&\leq [3 + o_p(1)] O_p\left(\bar{d}_m^2 + (NMT)^{-1}\right) \frac{1}{NMT} \sum_{i,j,t} |A_{ijt,m} C_{ijt,m}| \\
&= o(E_{m,2,2}) = o_p\left(\bar{d}_m^2 + \delta_{NMT,m}\right).
\end{aligned}
$$

Summarizing the above results yields the claims in the lemma. ∎

Proof of Lemma A.9. The key observation is that when Model m is just- or over-fitted, $M_{D_m} D^* = 0$ and $d'_{ijt,m}(D'_m D_m)^{-1} D'_m D^* \pi^* = d'_{ijt,m}(D'_m D_m)^{-1} D'_m D_m \pi_m = d'_{ijt,m} \pi_m = d^{*\prime}_{ijt,m} \pi^*$ where the coefficients in π_m corresponding to the redundant dummies in Model m are zero. As a result, $B_{ijt,m} = 0$ whenever Model m is just- or over-fitted for $m \in \{2, 3, ..., 7\}$. ∎

Proof of Lemma A.10. (i) For H_1, we make the following decomposition:

$$
H_1 = \frac{1}{NMT} \sum_{i,j,t} A_{ijt,1}^2 + \frac{1}{NMT} \sum_{i,j,t} (c_{ijt,1}^2 - 1) A_{ijt,1}^2 \equiv H_{1,1} + H_{1,2}, \text{ say.}
$$

Note that $H_{1,1} = \frac{1}{NMT} \sum_{i,j,t} u_{ijt}^2$. For $H_{1,2}$, we have

$$
\begin{aligned}
H_{1,2} &= \frac{1}{NMT} \sum_{i,j,t} \left(c_{ijt,1}^2 - 1\right) u_{ijt}^2 = \frac{1}{NMT} \sum_{i,j,t} \frac{2h_{ijt,1} - h_{ijt,1}^2}{(1 - h_{ijt,1})^2} u_{ijt}^2 \\
&= \frac{2}{NMT} \sum_{i,j,t} h_{ijt,1} u_{ijt}^2 + \frac{1}{NMT} \sum_{i,j,t} \frac{3h_{ijt,1} - 2h_{ijt,1}^2}{(1 - h_{ijt,1})^2} h_{ijt,1} u_{ijt}^2 \equiv H_{1,2,1} + H_{1,2,2}, \text{ say.}
\end{aligned}
$$

For $H_{1,2,1}$, we have that by Assumption A.1(iii) and (v) and the Markov inequality,

$$
\begin{aligned}
H_{1,2,1} &= \frac{2}{NMT} \sum_{i,j,t} h_{ijt,1} u_{ijt}^2 = \frac{2}{NMT} \sum_{i,j,t} x'_{ijt}(X'X)^{-1} x_{ijt} u_{ijt}^2 \\
&\leq \left\| (X'X)^{-1} \right\| \frac{2}{NMT} \sum_{i,j,t} \|x_{ijt}\|^2 u_{ijt}^2 = O_p((NMT)^{-1}).
\end{aligned}
$$

This, in conjunction with Lemma A.6(iv), implies that

$$
H_{1,2,2} \leq \max_{i,j,t} \left| \frac{3h_{ijt,1} - 2h_{ijt,1}^2}{(1 - h_{ijt,1})^2} \right| \frac{1}{NMT} \sum_{i,j,t} h_{ijt,1} u_{ijt}^2 = o_p(1) O_p((NMT)^{-1}) = o_p((NMT)^{-1}).
$$

Thus we have shown that $H_1 = \frac{1}{NMT} \sum_{i,j,t} u_{ijt}^2 + O_p((NMT)^{-1})$.

(ii) For H_2, we make the following decomposition

$$
\begin{aligned}
H_2 &= \frac{1}{NMT}\sum_{i,j,t}(1+2h_{ijt,2})A_{ijt,2}^2 + \frac{1}{NMT}\sum_{i,j,t}\left(c_{ijt,2}^2 - 1 - 2h_{ijt,2}\right)A_{ijt,2}^2 \\
&= (1 + 2\,(NM)^{-1} + 2\,(MT)^{-1} + 2\,(NT)^{-1})\frac{1}{NMT}\sum_{i,j,t}A_{ijt,2}^2 \\
&\quad + \frac{2}{NMT}\sum_{i,j,t}\left[h_{ijt,2} - (NM)^{-1} - (MT)^{-1} - (NT)^{-1}\right]A_{ijt,2}^2 \\
&\quad + \frac{1}{NMT}\sum_{i,j,t}\left(c_{ijt,2}^2 - 1 - 2h_{ijt,2}\right)A_{ijt,2}^2 \\
&\equiv H_{2,1} + H_{2,2} + H_{2,3}, \ \text{say}.
\end{aligned}
$$

By Lemmas A.2(ii) and A.3(i)-(iii) and Assumptions A.1(iv) and A.2(i)-(iii),

$$
\begin{aligned}
H_{2,1} &= (1 + \frac{2}{NM} + \frac{2}{MT} + \frac{2}{NT})\frac{1}{NMT}U'M_{D_2}U \\
&= (1 + \frac{2}{NM} + \frac{2}{MT} + \frac{2}{NT})\frac{1}{NMT}U'\left(I_{NMT} - P_I - P_J - P_T\right)U \\
&= (1 + \frac{2}{NM} + \frac{2}{MT} + \frac{2}{NT})\left\{\frac{1}{NMT}\sum_{i,j,t}u_{ijt}^2 - \frac{1}{N}\sum_{i=1}^{N}\overline{u}_{i\cdot\cdot}^2 - \frac{1}{M}\sum_{j=1}^{M}\overline{u}_{\cdot j\cdot}^2 - \frac{1}{T}\sum_{t=1}^{T}\overline{u}_{\cdot\cdot t}^2 + 3\overline{u}^2\right\} \\
&= (1 + \frac{2}{NM} + \frac{2}{MT} + \frac{2}{NT})\frac{1}{NMT}\sum_{i,j,t}u_{ijt}^2 - \frac{1}{N}\sum_{i=1}^{N}\overline{u}_{i\cdot\cdot}^2 - \frac{1}{M}\sum_{j=1}^{M}\overline{u}_{\cdot j\cdot}^2 - \frac{1}{T}\sum_{t=1}^{T}\overline{u}_{\cdot\cdot t}^2 \\
&\quad + O_p((NM)^{-2} + (MT)^{-2} + (NT)^{-2} + (NMT)^{-1}).
\end{aligned}
$$

For $H_{2,2}$, we apply the results of Lemmas A.5(i) and A.6(i) and Assumption A.1(iii) and (v),

$$
\begin{aligned}
H_{2,2} &= \frac{2}{NMT}\sum_{i,j,t}\left[h_{ijt,2} - (NM)^{-1} - (MT)^{-1} - (NT)^{-1}\right]A_{ijt,2}^2 \\
&= \frac{2}{NMT}\sum_{i,j,t}h_{ijt,2}^* A_{ijt,2}^2 + O_p((NMT)^{-1}) \leqslant 2\left\|(X'M_{D_2}X)^{-1}\right\|_{\mathrm{sp}}\overline{H}_{2,2} + O_p((NMT)^{-1}),
\end{aligned}
$$

where $\overline{H}_{2,2} = \frac{1}{NMT}\sum_{i,j,t}\left\|x_{ijt} - X'D_2\left(D_2'D_2\right)^{-1}d_{ijt,2}\right\|^2 (u_{ijt} - U'D_2\left(D_2'D_2\right)^{-1}d_{ijt,2})^2$. Noting that

$$
X'D_2\left(D_2'D_2\right)^{-1}d_{ijt,2} = \overline{x}_{i\cdot\cdot} + \overline{x}_{\cdot j\cdot} + \overline{x}_{\cdot\cdot t} - 3\overline{x} \quad \text{and} \quad U'D_2\left(D_2'D_2\right)^{-1}d_{ijt,2} = \overline{u}_{i\cdot\cdot} + \overline{u}_{\cdot j\cdot} + \overline{u}_{\cdot\cdot t} - 3\overline{u},
$$

we can readily apply the Cauchy-Schwarz inequality and Assumption A.1(iii)-(iv) and show that $\overline{H}_{2,2} = O_p(1)$. It follows that $H_{2,2} = O_p((NMT)^{-1})$. For $H_{2,3}$, we make the following decomposition

$$
\begin{aligned}
H_{2,3} &= \frac{1}{NMT}\sum_{i,j,t}\left(c_{ijt,2}^2 - 1 - 2h_{ijt,2}\right)A_{ijt,2}^2 = \frac{1}{NMT}\sum_{i,j,t}\frac{3 - 2h_{ijt,2}}{(1 - h_{ijt,2})^2}h_{ijt,2}^2 A_{ijt,2}^2 \\
&= \frac{3}{NMT}\sum_{i,j,t}h_{ijt,2}^2 A_{ijt,2}^2 + \frac{1}{NMT}\sum_{i,j,t}\left[\frac{3 - 2h_{ijt,2}}{(1 - h_{ijt,2})^2} - 3\right]h_{ijt,2}^2 A_{ijt,2}^2 \\
&\equiv H_{2,3,1} + H_{2,3,2}, \ \text{say}.
\end{aligned}
$$

For $H_{2,3,1}$, we have

$$
H_{2,3,1} \leq \max_{i,j,t}h_{ijt,2}^2 \frac{3}{NMT}\sum_{i,j,t}A_{ijt,2}^2 = O_p(\overline{d}_2^2).
$$

14

By Lemma A.6(v) and the dominated convergence theorem (DCT), $H_{2,3,2} = o_p(\bar{d}_2^2)$. Combining the above results yields the conclusion in (ii).

(iii) As in the proof of (ii), we make the following decomposition

$$
\begin{aligned}
H_3 &= \frac{1}{NMT} \sum_{i,j,t}(1+2h_{ijt,3})A_{ijt,3}^2 + \frac{1}{NMT}\sum_{i,j,t}(c_{ijt,3}^2 - 1 - 2h_{ijt,3})A_{ijt,3}^2 \\
&= (1+2T^{-1})\frac{1}{NMT}\sum_{i,j,t}A_{ijt,3}^2 + \frac{2}{NMT}\sum_{i,j,t}\left(h_{ijt,3} - T^{-1}\right)A_{ijt,3}^2 \\
&\quad + \frac{1}{NMT}\sum_{i,j,t}(c_{ijt,3}^2 - 1 - 2h_{ijt,3})A_{ijt,3}^2 \\
&\equiv H_{3,1} + H_{3,2} + H_{3,3}, \text{ say.}
\end{aligned}
$$

By Lemma A.3(iv) and Assumptions A.1(iv) and A.2(iv),

$$
\begin{aligned}
H_{3,1} &= (1+2T^{-1})\frac{1}{NMT}U'M_{D_3}U = (1+2T^{-1})\frac{1}{NMT}U'(I - P_{IJ})U \\
&= (1+2T^{-1})\left(\frac{1}{NMT}\sum_{i,j,t}u_{ijt}^2 - \frac{1}{NM}\sum_{i,j}\bar{u}_{ij\cdot}^2 + \bar{u}^2\right) \\
&= (1+2T^{-1})\frac{1}{NMT}\sum_{i,j,t}u_{ijt}^2 - \frac{1}{NM}\sum_{i,j}\bar{u}_{ij\cdot}^2 + O_p(T^{-2} + (NMT)^{-1}).
\end{aligned}
$$

By Lemmas A.5(ii) and A.6(i),

$$
\begin{aligned}
H_{3,2} &= \frac{2}{NMT}\sum_{i,j,t}\left(h_{ijt,3} - T^{-1}\right)A_{ijt,3}^2 = \frac{2}{NMT}\sum_{i,j,t}h_{ijt,3}^*A_{ijt,3}^2 + O_p((NMT)^{-1}) \\
&\leqslant 2\left\|(X'M_{D_3}X)^{-1}\right\|_{\text{sp}}\overline{H}_{3,2} + O_p((NMT)^{-1}),
\end{aligned}
$$

where $\overline{H}_{3,2} = \frac{1}{NMT}\sum_{i,j,t}\left\|x_{ijt} - X'D_3\left(D_3'D_3\right)^{-1}d_{ijt,3}\right\|^2\left(u_{ijt} - U'D_3\left(D_3'D_3\right)^{-1}d_{ijt,3}\right)^2$. Noting that

$$
X'D_3\left(D_3'D_3\right)^{-1}d_{ijt,3} = \bar{x}_{ij\cdot} - \bar{x} \text{ and } U'D_3\left(D_3'D_3\right)^{-1}d_{ijt,3} = \bar{u}_{ij\cdot} - \bar{u},
$$

we can readily show that $\overline{H}_{3,2} = O_p(1)$. Then $H_{3,2} = O_p((NMT)^{-1})$. Following the analysis of $H_{2,3}$, we can readily show that $H_{3,3} = O_p(T^{-2} + (NMT)^{-1})$. It follows that $H_3 = (1+\frac{2}{T})\frac{1}{NMT}\sum_{i,j,t}u_{ijt}^2 - \frac{1}{NM}\sum_{i,j}\bar{u}_{ij\cdot}^2 + O_p(T^{-2} + (NMT)^{-1})$.

(iv) As in the proof of (ii), we make the following decomposition

$$
\begin{aligned}
H_4 &= \frac{1}{NMT}\sum_{i,j,t}(1+2h_{ijt,4})A_{ijt,4}^2 + \frac{1}{NMT}\sum_{i,j,t}\left(c_{ijt,4}^2 - 1 - 2h_{ijt,4}\right)A_{ijt,4}^2 \\
&= (1 + 2T^{-1} + 2(NM)^{-1})\frac{1}{NMT}\sum_{i,j,t}A_{ijt,4}^2 + \frac{2}{NMT}\sum_{i,j,t}\left(h_{ijt,4} - T^{-1} - (NM)^{-1}\right)A_{ijt,4}^2 \\
&\quad + \frac{1}{NMT}\sum_{i,j,t}\left(c_{ijt,4}^2 - 1 - 2h_{ijt,4}\right)A_{ijt,4}^2 \\
&\equiv H_{4,1} + H_{4,2} + H_{4,3}, \text{ say.}
\end{aligned}
$$

15

By Lemmas A.2(ii) and A.3(iii)-(iv) and Assumptions A.1(iv) and A.2(iii)-(iv),

$$
\begin{aligned}
H_{4,1} &= (1 + 2T^{-1} + 2(NM)^{-1})\frac{1}{NMT}U'M_{D_4}U \\
&= (1 + 2T^{-1} + 2(NM)^{-1})\frac{1}{NMT}U'(I - P_{IJ} - P_T)U \\
&= (1 + 2T^{-1} + 2(NM)^{-1})\left(\frac{1}{NMT}\sum_{i,j,t}u_{ijt}^2 - \frac{1}{NM}\sum_{i,j}\bar{u}_{ij\cdot}^2 - \frac{1}{T}\sum_t \bar{u}_{\cdot\cdot t}^2 + 2\bar{u}^2\right) \\
&= (1 + 2T^{-1} + 2(NM)^{-1})\frac{1}{NMT}\sum_{i,j,t}u_{ijt}^2 - \frac{1}{NM}\sum_{i,j}\bar{u}_{ij\cdot}^2 - \frac{1}{T}\sum_t \bar{u}_{\cdot\cdot t}^2 \\
&\quad + O_p((NM)^{-2} + T^{-2} + (NMT)^{-1}).
\end{aligned}
$$

Using arguments as used in the analyses of $H_{2,2}$ and $H_{2,3}$, we can readily show that $H_{4,2} = O_p((NMT)^{-1})$ and $H_{4,3} = O_p((NM)^{-2} + T^{-2} + (NMT)^{-1})$. Then (iv) follows.

(v) As in the proof of (ii), we make the following decomposition

$$
\begin{aligned}
H_5 &= \frac{1}{NMT}\sum_{i,j,t}(1 + 2h_{ijt,5})A_{ijt,5}^2 + \frac{1}{NMT}\sum_{i,j,t}\left(c_{ijt,5}^2 - 1 - 2h_{ijt,5}\right)A_{ijt,5}^2 \\
&= (1 + 2N^{-1})\frac{1}{NMT}\sum_{i,j,t}A_{ijt,5}^2 + \frac{2}{NMT}\sum_{i,j,t}\left(h_{ijt,5} - N^{-1}\right)A_{ijt,5}^2 \\
&\quad + \frac{1}{NMT}\sum_{i,j,t}\left(c_{ijt,5}^2 - 1 - 2h_{ijt,5}\right)A_{ijt,5}^2 \\
&\equiv H_{5,1} + H_{5,2} + H_{5,3}, \text{ say.}
\end{aligned}
$$

By Lemma A.3(v) and Assumptions A.1(iv) and A.2(vi),

$$
\begin{aligned}
H_{5,1} &= (1 + 2N^{-1})\frac{1}{NMT}U'M_{D_5}U = (1 + 2N^{-1})\frac{1}{NMT}U'(I - P_{JT})U \\
&= (1 + 2N^{-1})\left(\frac{1}{NMT}\sum_{i,j,t}u_{ijt}^2 - \frac{1}{MT}\sum_{j,t}\bar{u}_{\cdot jt}^2 + \bar{u}^2\right) \\
&= (1 + 2N^{-1})\frac{1}{NMT}\sum_{i,j,t}u_{ijt}^2 - \frac{1}{MT}\sum_{j,t}\bar{u}_{\cdot jt}^2 + O_p(N^{-2} + (NMT)^{-1}).
\end{aligned}
$$

Using arguments as used in the analyses of $H_{2,2}$ and $H_{2,3}$, we can readily show that $H_{5,2} = O_p((NMT)^{-1})$ and $H_{5,3} = O_p(N^{-2} + (NMT)^{-1})$. Then (v) follows.

(vi) As in the proof of (ii), we make the following decomposition

$$
\begin{aligned}
H_6 &= (1 + 2N^{-1} + 2M^{-1})\frac{1}{NMT}\sum_{i,j,t}A_{ijt,6}^2 + \frac{2}{NMT}\sum_{i,j,t}\left(h_{ijt,6} - N^{-1} - M^{-1}\right)A_{ijt,6}^2 \\
&\quad + \frac{1}{NMT}\sum_{i,j,t}\left(c_{ijt,6}^2 - 1 - 2h_{ijt,6}\right)A_{ijt,6}^2 \\
&\equiv H_{6,1} + H_{6,2} + H_{6,3}, \text{ say.}
\end{aligned}
$$

By Lemmas A.2(ii) and A.3(v) and (vii), and Assumptions A.1(iv) and A.2(iii) and (v)-(vi),

$$
\begin{aligned}
H_{6,1} &= (1 + 2N^{-1} + 2M)^{-1}) \frac{1}{NMT} U' M_{D_6} U \\
&= (1 + 2N^{-1} + 2M)^{-1}) \frac{1}{NMT} U' (I - P_{IT}^* - P_{JT}) U \\
&= (1 + 2N^{-1} + 2M)^{-1}) \left(\frac{1}{NMT} \sum_{i,j,t} u_{ijt}^2 - \frac{1}{MT} \sum_{j,t} \overline{u}_{\cdot jt}^2 - \frac{1}{NT} \sum_{i,t} \overline{u}_{i \cdot t}^2 + \frac{1}{T} \sum_t \overline{u}_{\cdot \cdot t}^2 + \overline{u}^2 \right) \\
&= (1 + 2N^{-1} + 2M)^{-1}) \frac{1}{NMT} \sum_{i,j,t} u_{ijt}^2 - \frac{1}{MT} \sum_{j,t} \overline{u}_{\cdot jt}^2 - \frac{1}{NT} \sum_{i,t} \overline{u}_{i \cdot t}^2 \\
&\quad + O_p(N^{-2} + M^{-2} + (NMT)^{-1}).
\end{aligned}
$$

Using arguments as used in the analyses of $H_{2,2}$ and $H_{2,3}$, we can readily show that $H_{6,2} = O_p((NMT)^{-1})$ and $H_{6,3} = O_p(N^{-2} + M^{-2} + (NMT)^{-1})$. Then (iv) follows.

(vii) As in the proof of (ii), we make the following decomposition

$$
\begin{aligned}
H_7 &= (1 + 2N^{-1} + 2M^{-1} + 2T^{-1}) \frac{1}{NMT} \sum_{i,j,t} A_{ijt,7}^2 + \frac{2}{NMT} \sum_{i,j,t} \left(h_{ijt,7} - N^{-1} - M^{-1} - T^{-1} \right) A_{ijt,7}^2 \\
&\quad + \frac{1}{NMT} \sum_{i,j,t} \left(c_{ijt,7}^2 - 1 - 2h_{ijt,7} \right) A_{ijt,7}^2 \\
&\equiv H_{7,1} + H_{7,2} + H_{7,3}, \text{ say.}
\end{aligned}
$$

By Lemmas A.2(ii) and A.3(vi)-(vii) and Assumptions A.1(iv) and A.2(i)-(vi)

$$
\begin{aligned}
&H_{7,1} \\
&= (1 + 2N^{-1} + 2M^{-1} + 2T^{-1}) \frac{1}{NMT} U' M_{D_7} U \\
&= (1 + 2N^{-1} + 2M^{-1} + 2T^{-1}) \frac{1}{NMT} U' (I - P_{IT}^* - P_{IJ}^* - P_{JT}^*) U \\
&= (1 + 2N^{-1} + 2M^{-1} + 2T^{-1}) \\
&\quad \times \left(\frac{1}{NMT} \sum_{i,j,t} u_{ijt}^2 - \frac{1}{NM} \sum_{i,j} \overline{u}_{ij \cdot}^2 - \frac{1}{NT} \sum_{i,t} \overline{u}_{i \cdot t}^2 - \frac{1}{MT} \sum_{j,t} \overline{u}_{\cdot jt}^2 + \frac{1}{N} \sum_i \overline{u}_{i \cdot \cdot}^2 + \frac{1}{T} \sum_t \overline{u}_{\cdot \cdot t}^2 + \frac{1}{M} \sum_j \overline{u}_{\cdot j \cdot}^2 \right) \\
&= (1 + 2N^{-1} + 2M^{-1} + 2T^{-1}) \frac{1}{NMT} \sum_{i,j,t} u_{ijt}^2 - \frac{1}{NM} \sum_{i,j} \overline{u}_{ij \cdot}^2 - \frac{1}{NT} \sum_{i,t} \overline{u}_{i \cdot t}^2 - \frac{1}{MT} \sum_{j,t} \overline{u}_{\cdot jt}^2 \\
&\quad + O_p(N^{-2} + M^{-2} + T^{-2} + (NMT)^{-1}).
\end{aligned}
$$

Using arguments as used in the analyses of $H_{2,2}$ and $H_{2,3}$, we can readily show that $H_{7,2} = O_p((NMT)^{-1})$ and $H_{7,3} = O_p(N^{-2} + M^{-2} + T^{-2} + +(NMT)^{-1})$. Then (vii) follows. ∎

Proof of Lemma A.11. (i) We write $G_m = \frac{2}{NMT} \sum_{i,j,t} c_{ijt,m}^2 A_{ijt,m} B_{ijt,m} + \frac{2}{NMT} \sum_{i,j,t} c_{ijt,m}^2 B_{ijt,m} \times C_{ijt,m} \equiv 2G_{m,1} + 2G_{m,2}$, say. For $G_{m,1}$, we make futher decomposition

$$
G_{m,1} = \frac{1}{NMT} \sum_{i,j,t} A_{ijt,m} B_{ijt,m} + \frac{1}{NMT} \sum_{i,j,t} (c_{ijt,m}^2 - 1) A_{ijt,m} B_{ijt,m} \equiv G_{m,1,1} + G_{m,1,2}, \text{ say.}
$$

Noting that $M_{D_m}(I - P_{M_{D_m} X}) M_{D_m} = I - P_{M_{D_m} X} - P_{D_m}$, we can readily show that by Lemma A.7(iii) and

Assumption A.4

$$G_{m,1,1} = \frac{1}{NMT}\sum_{i,j,t}A_{ijt,m}B_{ijt,m} = \frac{1}{NMT}U'M_{D_m}(I - P_{M_{D_m}X})M_{D_m}D^*\pi^*$$

$$= \frac{1}{NMT}U'\left(I - P_{M_{D_m}X} - P_{D_m}\right)D^*\pi^*$$

$$= \frac{1}{NMT}\sum_{i,j,t}[d^*_{ijt}\pi^* - x'_{ijt,m}(Z'_mZ_m)^{-1}Z'_mD^*\pi^*]u_{ijt} = o_p(1).$$

For $G_{m,1,2}$, we have by Lemma A.6(iv)

$$|G_{m,1,2}| \leq \frac{2[1 + o_p(1)]}{NMT}\sum_{i,j,t}h_{ijt,m}|A_{ijt,m}B_{ijt,m}|$$

$$\leq 2[1 + o_p(1)]\max_{i,j,t}h_{ijt,m}\left\{\frac{1}{NMT}\sum_{i,j,t}A^2_{ijt,m}\right\}^{1/2}\left\{\frac{1}{NMT}\sum_{i,j,t}B^2_{ijt,m}\right\}^{1/2}$$

$$\leq O_p(\bar{d}_m + (NMT)^{-1/2})O_p(1)O_p(1) = o_p(1),$$

where we use the fact that $\frac{1}{NMT}\sum_{i,j,t}A^2_{ijt,m} = O_p(1)$ and $\frac{1}{NMT}\sum_{i,j,t}B^2_{ijt,m} = O_p(1)$. Thus $G_{m,1} = o_p(1)$.

For $G_{m,2}$, noting that $\sum_{i,j,t}B_{ijt,m}C_{ijt,m} = 0$ by Lemma A.7(iii), we have $G_{m,2} = \frac{1}{NMT}\sum_{i,j,t}(c^2_{ijt,m} - 1)B_{ijt,m}C_{ijt,m}$. Using arguments used in the analysis of $G_{m,1,2}$, we can show that $G_{m,2} = o_p(1)$. This completes the proof of the lemma.

(ii) Note that $K_{m^*,m} = \frac{1}{NMT}\sum_{i,j,t}B^2_{ijt,m} + \frac{1}{NMT}\sum_{i,j,t}(c^2_{ijt,m} - 1)B^2_{ijt,m} \equiv K_{m^*,m,1} + K_{m^*,m,2}$, say. For $K_{m^*,m,1}$, we have by Assumption A.4

$$K_{m^*,m,1} = \frac{1}{NMT}\sum_{i,j,t}B^2_{ijt,m} = \pi^{*'}D^{*'}M_{D_m}(I - P_{M_{D_m}X})M_{D_m}D^*\pi^*$$

$$= \pi^{*'}D^{*'}(I - P_{M_{D_m}X} - P_{D_m})D^*\pi^* \xrightarrow{m} \varphi_{m^*,m}.$$

By Lemma A.6(iii) and the DCT, $K_{m^*,m,2} = o_p(1)$. This completes the proof of the lemma. ∎

Proof of Lemma A.12. (i) Noting that $P_{D_4} - P_{D_2} = P_{IJ} - P_I - P_J$ by Lemma A.2(ii), we have by Lemma A.3(i), (ii), and (iv) and Assumptions A.1-A.3

$$\frac{1}{NMT}A'(M_{D_2} - M_{D_4})B = \frac{1}{NMT}A'(P_{IJ} - P_I - P_J)B$$

$$= \frac{1}{NM}\sum_{i,j}\bar{a}_{ij\cdot}\bar{b}'_{ij\cdot} - \frac{1}{N}\sum_{i=1}\bar{a}_{i\cdot\cdot}\bar{b}'_{i\cdot\cdot} - \frac{1}{M}\sum_{j=1}\bar{a}_{\cdot j\cdot}\bar{b}'_{\cdot j\cdot} + \bar{a}\bar{b}'$$

$$= \frac{1}{NM}\sum_{i,j}(\bar{a}_{ij\cdot} - \bar{a})(\bar{b}_{ij\cdot} - \bar{b})' - \frac{1}{N}\sum_{i=1}(\bar{a}_{i\cdot\cdot} - \bar{a})(\bar{b}_{i\cdot\cdot} - \bar{b})'$$

$$- \frac{1}{M}\sum_{j=1}(\bar{a}_{\cdot j\cdot} - \bar{a})(\bar{b}_{\cdot j\cdot} - \bar{b})'$$

$$= \frac{1}{NM}\sum_{i,j}(\bar{a}_{ij\cdot} - \bar{a})(\bar{b}_{ij\cdot} - \bar{b})' + O_p((MT)^{-1} + (NT)^{-1}) = O_p(T^{-1}).$$

(ii) Noting that $P_{D_4} - P_{D_3} = P_T$ by Lemma A.2(ii) and the fact $D_3 = D_{IJ}$, we have by Lemma A.3(iii)

$$\frac{1}{NMT}A'(M_{D_3} - M_{D_4})B = \frac{1}{NMT}A'P_TB = \frac{1}{T}\sum_t(\bar{a}_{\cdot\cdot t} - \bar{a})(\bar{b}_{\cdot\cdot t} - \bar{b}_\cdot)' = O_p((NM)^{-1}).$$

(iii) Noting that $P_{D_7} - P_{D_3} = P_{IT}^* + P_{IJ}^* + P_{JT}^* - P_{IJ}$ by Lemma A.2(ii),

$$
\begin{aligned}
\frac{1}{NMT} A'\left(M_{D_3} - M_{D_7}\right) B &= \frac{1}{NMT} A'\left(P_{IT}^* + P_{IJ}^* + P_{JT}^* - P_{IJ}\right) B \\
&= \frac{1}{NT}\sum_{i,t}\overline{a}_{i\cdot t}\overline{b}_{i\cdot t}' + \frac{1}{MT}\sum_{j,t}\overline{a}_{\cdot jt}\overline{b}_{\cdot jt}' - \frac{1}{T}\sum_{t=1}\overline{a}_{\cdot\cdot t}\overline{b}_{\cdot\cdot t}' \\
&\quad - \frac{1}{N}\sum_{i=1}\overline{a}_{i\cdot\cdot}\overline{b}_{i\cdot\cdot}' - \frac{1}{M}\sum_{j=1}\overline{a}_{\cdot j\cdot}\overline{b}_{\cdot j\cdot}' + \overline{a}\overline{b}' \\
&= \frac{1}{NT}\sum_{i,t}(\overline{a}_{i\cdot t} - \overline{a})(\overline{b}_{i\cdot t} - \overline{b})' + \frac{1}{MT}\sum_{j,t}\left(\overline{a}_{\cdot jt} - \overline{a}\right)\left(\overline{b}_{\cdot jt} - \overline{b}\right)' \\
&\quad + O_p((NM)^{-1} + (MT)^{-1} + (NT)^{-1}) \\
&= O_p(M^{-1} + N^{-1}).
\end{aligned}
$$

(iv) Noting that $P_{D_7} - P_{D_4} = P_{IT}^* + P_{IJ}^* + P_{JT}^* - P_{IJ} - P_T$ by Lemma A.2(ii), we have by (iii) and Lemma A.3(iii)

$$
\begin{aligned}
\frac{1}{NMT} A'\left(M_{D_4} - M_{D_7}\right) B &= \frac{1}{NMT} A'\left(P_{IT}^* + P_{IJ}^* + P_{JT}^* - P_{IJ} - P_T\right) B \\
&= \frac{1}{NT}\sum_{i,t}(\overline{a}_{i\cdot t} - \overline{a})(\overline{b}_{i\cdot t} - \overline{b})' + \frac{1}{MT}\sum_{j,t}\left(\overline{a}_{\cdot jt} - \overline{a}\right)\left(\overline{b}_{\cdot jt} - \overline{b}\right)' \\
&\quad - \frac{2}{T}\sum_{t=1}^{T}\left(\overline{a}_{\cdot\cdot t} - \overline{a}\right)\left(\overline{b}_{\cdot\cdot t} - \overline{b}\right)' - \frac{1}{N}\sum_{i=1}^{N}\left(\overline{a}_{i\cdot\cdot} - \overline{a}\right)\left(\overline{b}_{i\cdot\cdot} - \overline{b}\right)' \\
&\quad - \frac{1}{M}\sum_{j=1}^{M}(\overline{a}_{\cdot j\cdot} - \overline{a})(\overline{b}_{\cdot j\cdot} - \overline{b})' \\
&= \frac{1}{NT}\sum_{i,t}(\overline{a}_{i\cdot t} - \overline{a})(\overline{b}_{i\cdot t} - \overline{b})' + \frac{1}{MT}\sum_{j,t}\left(\overline{a}_{\cdot jt} - \overline{a}\right)\left(\overline{b}_{\cdot jt} - \overline{b}\right)' \\
&\quad + O_p((NM)^{-1} + (MT)^{-1} + (NT)^{-1}) = O_p(M^{-1} + N^{-1}).
\end{aligned}
$$

(v) Noting that $P_{D_6} - P_{D_5} = P_{IT}^*$ by Lemma A.2(ii) and the fact that $D_5 = D_{JT}$, we have Lemma A.3(vii)

$$
\begin{aligned}
\frac{1}{NMT} A'\left(M_{D_5} - M_{D_6}\right) B &= \frac{1}{NMT} A' P_{IT}^* B = \frac{1}{NT}\sum_{i,t}(\overline{a}_{i\cdot t} - \overline{a})(\overline{b}_{i\cdot t} - \overline{b})' - \frac{1}{T}\sum_{t=1}^{T}\left(\overline{a}_{\cdot\cdot t} - \overline{a}\right)\left(\overline{b}_{\cdot\cdot t} - \overline{b}\right)' \\
&= \frac{1}{NT}\sum_{i,t}(\overline{a}_{i\cdot t} - \overline{a})(\overline{b}_{i\cdot t} - \overline{b})' + O_p((NM)^{-1}) = O_p(M^{-1}).
\end{aligned}
$$

(vi) Noting that $P_{D_7} - P_{D_5} = P_{IT}^* + P_{IJ}^* + P_{JT}^* - P_{JT}$ by Lemma A.2(ii) and the fact that $D_5 = D_{JT}$,

we have Lemma A.3(v)-(viii)

$$
\begin{aligned}
\frac{1}{NMT}A'\left(M_{D_5}-M_{D_7}\right)B &= \frac{1}{NMT}A'\left(P_{IT}^*+P_{IJ}^*+P_{JT}^*-P_{JT}\right)B \\
&= \frac{1}{NT}\sum_{i,t}(\overline{a}_{i\cdot t}-\overline{a})(\overline{b}_{i\cdot t}-\overline{b})' + \frac{1}{NM}\sum_{i,j}(\overline{a}_{ij\cdot}-\overline{a})(\overline{b}_{ij\cdot}-\overline{b})' \\
&\quad -\frac{1}{T}\sum_{t=1}^{T}(\overline{a}_{\cdot\cdot t}-\overline{a})(\overline{b}_{\cdot\cdot t}-\overline{b})' - \frac{1}{N}\sum_{i=1}^{N}(\overline{a}_{i\cdot\cdot}-\overline{a})(\overline{b}_{i\cdot\cdot}-\overline{b})' \\
&\quad -\frac{1}{M}\sum_{j=1}^{M}(\overline{a}_{\cdot j\cdot}-\overline{a})(\overline{b}_{\cdot j\cdot}-\overline{b})' \\
&= \frac{1}{NT}\sum_{i,t}(\overline{a}_{i\cdot t}-\overline{a})(\overline{b}_{i\cdot t}-\overline{b})' + \frac{1}{NM}\sum_{i,j}(\overline{a}_{ij\cdot}-\overline{a})(\overline{b}_{ij\cdot}-\overline{b})' \\
&\quad +O_p((NM)^{-1}+(MT)^{-1}+(NT)^{-1})=O_p(M^{-1}+T^{-1}).
\end{aligned}
$$

(vii) Noting that $P_{D_7}-P_{D_6}=P_{IJ}^*+P_{JT}^*-P_{JT}$ by Lemma A.2(ii), we have Lemma A.3(v), (vi) and (viii),

$$
\begin{aligned}
\frac{1}{NMT}A'\left(M_{D_6}-M_{D_7}\right)B &= \frac{1}{NMT}A'\left(P_{IJ}^*+P_{JT}^*-P_{JT}\right)B \\
&= \frac{1}{NM}\sum_{i,j}(\overline{a}_{ij\cdot}-\overline{a})(\overline{b}_{ij\cdot}-\overline{b})' - \frac{1}{N}\sum_{i=1}^{N}(\overline{a}_{i\cdot\cdot}-\overline{a})(\overline{b}_{i\cdot\cdot}-\overline{b})' \\
&\quad -\frac{1}{M}\sum_{j=1}^{M}(\overline{a}_{\cdot j\cdot}-\overline{a})(\overline{b}_{\cdot j\cdot}-\overline{b})' \\
&= \frac{1}{NM}\sum_{i,j}(\overline{a}_{ij\cdot}-\overline{a})(\overline{b}_{ij\cdot}-\overline{b})' + O_p((MT)^{-1}+(NT)^{-1})=O_p(T^{-1}).\ \blacksquare
\end{aligned}
$$

Proof of Lemma A.13. We have $\overline{D}_{ijt,m}=D_m(D_m'D_m)^{-1}d_{ijt,m}$ which is the $((i-1)MT+(j-1)T+t)$th column of P_m. Then, for $A=\{a_{ijt}\}$, we have:

$$
\begin{aligned}
\overline{D}'_{ijt,2}A &= \overline{a}_{i\cdot\cdot}+\overline{a}_{\cdot j\cdot}+\overline{a}_{\cdot\cdot t}-3\overline{a}, \\
\overline{D}'_{ijt,3}A &= \overline{a}_{ij\cdot}-\overline{a}, \\
\overline{D}'_{ijt,4}A &= \overline{a}_{ij\cdot}+\overline{a}_{\cdot\cdot t}-2\overline{a}, \\
\overline{D}'_{ijt,5}A &= \overline{a}_{\cdot jt}-\overline{a}, \\
\overline{D}'_{ijt,6}A &= \overline{a}_{i\cdot t}+\overline{a}_{\cdot jt}-\overline{a}_{\cdot\cdot t}-\overline{a}, \\
\overline{D}'_{ijt,7}A &= \overline{a}_{i\cdot t}+\overline{a}_{\cdot jt}+\overline{a}_{ij\cdot}-\overline{a}_{i\cdot\cdot}-\overline{a}_{\cdot j\cdot}-\overline{a}_{\cdot\cdot t}.
\end{aligned}
$$

Below we focus on the proof of (i) as the proofs of the other parts in the lemma are analogous.

(i) If Model 2 is the true model, we have $e_{ijt,4}=A_{ijt,4}+C_{ijt,4}$ and $e_{ijt,2}=A_{ijt,2}+C_{ijt,2}$. It follows that

$$
\begin{aligned}
e_{ijt,4}^2-e_{ijt,2}^2 &= (A_{ijt,4}^2-A_{ijt,2}^2)+(C_{ijt,4}^2-C_{ijt,2}^2)+2(A_{ijt,4}-A_{ijt,2})C_{ijt,4}+2A_{ijt,2}(C_{ijt,4}-C_{ijt,2}) \\
&\equiv \sum_{l=1}^{4}e_{ijt,42}(l),\ \text{say}.
\end{aligned}
$$

By the triangle inequality, we have $\frac{1}{NMT}\sum_{i,j,t}h_{ijt,4}^*\left|e_{ijt,4}^2-e_{ijt,2}^2\right| \le \sum_{l=1}^{4}\frac{1}{NMT}\sum_{i,j,t}h_{ijt,4}^*\left|e_{ijt,42}(l)\right| \equiv \sum_{l=1}^{4}E_{2,4}(l)$, say.

First, we study $E_{2,4}(1)$. Noting that

$$
\begin{aligned}
A_{ijt,4}^2 - A_{ijt,2}^2 &= (A_{ijt,4} - A_{ijt,2})(A_{ijt,4} + A_{ijt,2}) \\
&= -(\bar{D}_{ijt,4} - \bar{D}_{ijt,2})'U\left[2u_{ijt} - (\bar{D}_{ijt,4} + \bar{D}_{ijt,2})'U\right] \\
&= -2(\bar{D}_{ijt,4} - \bar{D}_{ijt,2})'Uu_{ijt} + (\bar{D}_{ijt,4} - \bar{D}_{ijt,2})'UU'(\bar{D}_{ijt,4} + \bar{D}_{ijt,2}),
\end{aligned}
$$

we have

$$
\begin{aligned}
E_{2,4}(1) &= \frac{1}{NMT}\sum_{i,j,t} h_{ijt,4}^* \left|A_{ijt,4}^2 - A_{ijt,2}^2\right| \\
&\leq \frac{2}{NMT}\sum_{i,j,t} h_{ijt,4}^* \left|(\bar{D}_{ijt,4} - \bar{D}_{ijt,2})'Uu_{ijt}\right| \\
&\quad + \frac{1}{NMT}\sum_{i,j,t} h_{ijt,4}^* \left|(\bar{D}_{ijt,4} - \bar{D}_{ijt,2})'UU'(\bar{D}_{ijt,4} + \bar{D}_{ijt,2})\right| \equiv 2E_{2,4}(1,1) + E_{2,4}(1,2).
\end{aligned}
$$

By the Cauchy-Schwarz inequality,

$$
E_{2,4}(1,1) \leq \left\{\frac{1}{NMT}\sum_{i,j,t} h_{ijt,4}^*\right\}^{1/2} \left\{\frac{1}{NMT}\sum_{i,j,t} h_{ijt,4}^* \left\|(\bar{D}_{ijt,4} - \bar{D}_{ijt,2})'Uu_{ijt}\right\|^2\right\}^{1/2}.
$$

Noting that $x_{ijt} - X'\bar{D}_{ijt,4} = x_{ijt} - X'D_4(D_4'D_4)^{-1}d_{ijt,4}$ denotes the residual in the OLS regression of x_{ijt} on $d_{ijt,4}$, we have

$$
\begin{aligned}
\frac{1}{NMT}\sum_{i,j,t} h_{ijt,4}^* &= \frac{1}{NMT}\sum_{i,j,t} (x_{ijt} - X'\bar{D}_{ijt,4})' X_{D_4}^* (x_{ijt} - X'\bar{D}_{ijt,4}) \\
&\leq \|X_{D_4}^*\|_{\mathrm{sp}} \frac{1}{NMT}\sum_{i,j,t} (x_{ijt} - X'\bar{D}_{ijt,4})'(x_{ijt} - X'\bar{D}_{ijt,4}) \\
&\leq \|X_{D_4}^*\|_{\mathrm{sp}} \frac{1}{NMT}\sum_{i,j,t} \|x_{ijt}\|^2 = O_p((NMT)^{-1}).
\end{aligned}
$$

Noting that $h_{ijt,4}^* \leq 2x_{ijt}'X_{D_4}^* x_{ijt} + 2\bar{D}_{ijt,4}'XX_{D_4}^*X'\bar{D}_{ijt,4}$, we have

$$
\begin{aligned}
&\frac{1}{NMT}\sum_{i,j,t} h_{ijt,4}^* \left\|(\bar{D}_{ijt,4} - \bar{D}_{ijt,2})'Uu_{ijt}\right\|^2 \\
&\leq \frac{2}{NMT}\sum_{i,j,t} x_{ijt}'X_{D_4}^* x_{ijt} \left\|(\bar{D}_{ijt,4} - \bar{D}_{ijt,2})'Uu_{ijt}\right\|^2 \\
&\quad + \frac{2}{NMT}\sum_{i,j,t} \bar{D}_{ijt,4}'XX_{D_4}^*X'\bar{D}_{ijt,4} \left\|(\bar{D}_{ijt,4} - \bar{D}_{ijt,2})'Uu_{ijt}\right\|^2 \equiv I + II, \text{ say,}
\end{aligned}
$$

where

$$
\begin{aligned}
I &\leq \|X_{D_4}^*\|_{\mathrm{sp}} \frac{2}{NMT}\sum_{i,j,t} \|x_{ijt}\|^2 u_{ijt}^2 \left[(\bar{D}_{ijt,4} - \bar{D}_{ijt,2})'U\right]^2 \\
&= \|X_{D_4}^*\|_{\mathrm{sp}} \frac{2}{NMT}\sum_{i,j,t} \|x_{ijt}\|^2 u_{ijt}^2 (\bar{u}_{ij\cdot} - \bar{u}_{i\cdot\cdot} - \bar{u}_{\cdot j\cdot} + \bar{u})^2 \\
&\leq \|X_{D_4}^*\|_{\mathrm{sp}} \frac{8}{NMT}\sum_{i,j,t} \|x_{ijt}\|^2 (\bar{u}_{ij\cdot}^2 + \bar{u}_{i\cdot\cdot}^2 + \bar{u}_{\cdot j\cdot}^2) + O_p((NMT)^{-2}) \\
&= O_p((NMT)^{-1}),
\end{aligned}
$$

21

since we can show $\frac{1}{NMT}\sum_{i,j,t}\|x_{ijt}\|^2 u_{ijt}^2(\overline{u}_{ij.}^2+\overline{u}_{i..}^2+\overline{u}_{.j.}^2)=O_p(1)$ by Assumption A.1(iii) and the Cauchy-Schwarz and Jensen inequalities. Similarly

$$
\begin{aligned}
II &\leq \left\|X_{D_4}^*\right\|_{\mathrm{sp}}\frac{1}{NMT}\sum_{i,j,t}u_{ijt}^2\left\|\bar{D}_{ijt,4}'X\right\|^2\left[\left(\bar{D}_{ijt,4}-\bar{D}_{ijt,2}\right)'U\right]^2 \\
&= \left\|X_{D_4}^*\right\|_{\mathrm{sp}}\frac{1}{NMT}\sum_{i,j,t}u_{ijt}^2\left\|\overline{x}_{ij.}+\overline{x}_{..t}-2\overline{x}\right\|^2\left(\overline{u}_{ij.}-\overline{u}_{i..}-\overline{u}_{.j.}+\overline{u}\right)^2 \\
&= O_p((NMT)^{-1}),
\end{aligned}
$$

where we use the fact $\|\overline{x}_{ij.}+\overline{x}_{..t}-2\overline{x}\|^2\leq 3(\|\overline{x}_{ij.}\|^2+\|\overline{x}_{..t}\|^2+\|2\overline{x}\|^2)$, $(\overline{u}_{ij.}-\overline{u}_{i..}-\overline{u}_{.j.}+\overline{u})^2\leq 4(\overline{u}_{ij.}^2+\overline{u}_{i..}^2+\overline{u}_{.j.}^2+\overline{u}^2)$, and that $\frac{1}{NMT}\sum_{i,j,t}u_{ijt}^2\|x_{ijt}^\dagger\|^2 u_{ijt}^{\dagger 2}=O_p(1)$ for $x_{ijt}^\dagger=\overline{x}_{ij.}$, $\overline{x}_{..t}$ and $\overline{x}$ and $u_{ijt}^\dagger=\overline{u}_{ij.}^2$, $\overline{u}_{i..}^2$, $\overline{u}_{.j.}^2$ and $\overline{u}^2$ by Assumption A.1(iii). Consequently, we have $E_{2,4}(1,1)=O_p((NMT)^{-1})=o_p\left(T^{-1}\right)$.

For $E_{2,4}(1,2)$ we have

$$
E_{2,4}(1,2)\leq\left\{\frac{1}{NMT}\sum_{i,j,t}h_{ijt,4}^*\right\}^{1/2}\left\{\frac{1}{NMT}\sum_{i,j,t}h_{ijt,4}^*\left[\left(\bar{D}_{ijt,4}-\bar{D}_{ijt,2}\right)'UU'\left(\bar{D}_{ijt,4}+\bar{D}_{ijt,2}\right)\right]^2\right\}^{1/2},
$$

where

$$
\begin{aligned}
&\frac{1}{NMT}\sum_{i,j,t}h_{ijt,4}^*\left[\left(\bar{D}_{ijt,4}-\bar{D}_{ijt,2}\right)'UU'\left(\bar{D}_{ijt,4}+\bar{D}_{ijt,2}\right)\right]^2 \\
\leq\quad &\frac{2}{NMT}\sum_{i,j,t}x_{ijt}'X_{D_4}^*x_{ijt}\left[\left(\bar{D}_{ijt,4}-\bar{D}_{ijt,2}\right)'UU'\left(\bar{D}_{ijt,4}+\bar{D}_{ijt,2}\right)\right]^2 \\
&+\frac{2}{NMT}\sum_{i,j,t}\bar{D}_{ijt,4}'XX_{D_4}^*X'\bar{D}_{ijt,4}\left[\left(\bar{D}_{ijt,4}-\bar{D}_{ijt,2}\right)'UU'\left(\bar{D}_{ijt,4}+\bar{D}_{ijt,2}\right)\right]^2\equiv III+IV.
\end{aligned}
$$

Note that

$$
\begin{aligned}
III &\leq \left\|X_{D_4}^*\right\|_{\mathrm{sp}}\frac{2}{NMT}\sum_{i,j,t}\|x_{ijt}\|^2\left[\left(\bar{D}_{ijt,4}-\bar{D}_{ijt,2}\right)'UU'\left(\bar{D}_{ijt,4}+\bar{D}_{ijt,2}\right)\right]^2 \\
&= \left\|X_{D_4}^*\right\|_{\mathrm{sp}}\frac{1}{NMT}\sum_{i,j,t}\|x_{ijt}\|^2\left[\left(\overline{u}_{ij.}-\overline{u}_{i..}-\overline{u}_{.j.}+\overline{u}\right)\left(\overline{u}_{ij.}+\overline{u}_{i..}+\overline{u}_{.j.}+2\overline{u}_{..t}-5\overline{u}\right)\right]^2 \\
&= O_p((NMT)^{-1}),
\end{aligned}
$$

since we can show $\frac{1}{NMT}\sum_{i,j,t}\|x_{ijt}\|^2\left(\overline{u}_{ij.}^4+\overline{u}_{i..}^4+\overline{u}_{.j.}^4+\overline{u}_{..t}^4\right)=O_p(1)$ by Assumption A.1(iii). Similarly, we can show that $IV=O_p(\bar{d}_4)O_p((NMT)^{-1})$. It follows that $E_{2,4}(1,2)=O_p((NMT)^{-1/2})\times O_p(\bar{d}_4^{1/2}(NMT)^{-1/2})=o_p\left(T^{-1}\right)$. In sum, we have shown that $E_{2,4}(1)=o_p\left(T^{-1}\right)$.

Next, we study $E_{2,4}(2)$. Noting that

$$
\begin{aligned}
C_{ijt,4}-C_{ijt,2} &= \left(\bar{D}_{ijt,4}'X-x_{ijt}'\right)X_{D_4}^*X'M_{D_4}U-\left(\bar{D}_{ijt,2}'X-x_{ijt}'\right)X_{D_2}^*X'M_{D_2}U \\
&= \left(\bar{D}_{ijt,4}-\bar{D}_{ijt,2}\right)'XX_{D_4}^*X'M_{D_4}U+\left(\bar{D}_{ijt,2}'X-x_{ijt}'\right)(X_{D_4}^*-X_{D_2}^*)X'M_{D_4}U \\
&\quad+\left(\bar{D}_{ijt,2}'X-x_{ijt}'\right)X_{D_2}^*X'\left(M_{D_4}-M_{D_2}\right)U,
\end{aligned}
$$

we have

$$
\begin{aligned}
E_{2,4}\left(2\right) & = \frac{1}{NMT}\sum_{i,j,t}h^*_{ijt,4}\left|\left(C_{ijt,4}-C_{ijt,2}\right)\left(C_{ijt,4}+C_{ijt,2}\right)\right| \\
& \leq \frac{1}{NMT}\sum_{i,j,t}h^*_{ijt,4}\left|(\bar{D}_{ijt,4}-\bar{D}_{ijt,2})'XX^*_{D_4}X'M_{D_4}U\left(C_{ijt,4}+C_{ijt,2}\right)\right| \\
& \quad +\frac{1}{NMT}\sum_{i,j,t}h^*_{ijt,4}\left|(\bar{D}'_{ijt,2}X-x'_{ijt})(X^*_{D_4}-X^*_{D_2})X'M_{D_4}U\left(C_{ijt,4}+C_{ijt,2}\right)\right| \\
& \quad +\frac{1}{NMT}\sum_{i,j,t}h^*_{ijt,4}\left|(\bar{D}'_{ijt,2}X-x'_{ijt})X^*_{D_2}X'\left(M_{D_4}-M_{D_2}\right)U\left(C_{ijt,4}+C_{ijt,2}\right)\right| \\
& \equiv E_{2,4}\left(2,1\right)+E_{2,4}\left(2,2\right)+E_{2,4}\left(2,3\right),\text{ say.}
\end{aligned}
$$

It suffices to consider the probability bound for each term. By the Cauchy-Schwarz inequality

$$
\begin{aligned}
E_{2,4}\left(2,1\right) & \leq \left\{\frac{1}{NMT}\sum_{i,j,t}h^*_{ijt,4}(\bar{D}_{ijt,4}-\bar{D}_{ijt,2})'XX^*_{D_4}X'(\bar{D}_{ijt,4}-\bar{D}_{ijt,2})\right\}^{1/2} \\
& \quad \times \left\{\frac{1}{NMT}\sum_{i,j,t}h^*_{ijt,4}\left(C_{ijt,4}+C_{ijt,2}\right)^2U'M_{D_4}XX^*_{D_4}X'M_{D_4}U\right\}^{1/2}.
\end{aligned}
$$

Noting that $\overline{D}'_{ijt,4}A=\bar{a}_{ij\cdot}+\bar{a}_{\cdot\cdot t}-2\bar{a}$ and $\overline{D}'_{ijt,2}A=\bar{a}_{i\cdot\cdot}+\bar{a}_{\cdot j\cdot}+\bar{a}_{\cdot\cdot t}-3\bar{a}$, respectively, and using $h^*_{ijt,4}\leq 2\left\|X^*_{D_4}\right\|_{\mathrm{sp}}(\|x_{ijt}\|^2+\|\bar{D}'_{ijt,4}X\|^2)$, we have

$$
\begin{aligned}
& \frac{1}{NMT}\sum_{i,j,t}h^*_{ijt,4}(\bar{D}_{ijt,4}-\bar{D}_{ijt,2})'XX^*_{D_4}X'(\bar{D}_{ijt,4}-\bar{D}_{ijt,2}) \\
& \leq \left\|X^*_{D_4}\right\|^2_{\mathrm{sp}}\frac{2}{NMT}\sum_{i,j,t}\left\{\|x_{ijt}\|^2+\|\bar{D}'_{ijt,4}X\|^2\right\}(\bar{D}_{ijt,4}-\bar{D}_{ijt,2})'XX'(\bar{D}_{ijt,4}-\bar{D}_{ijt,2}) \\
& \leq \left\|X^*_{D_4}\right\|^2_{\mathrm{sp}}\frac{2}{NMT}\sum_{i,j,t}\left\{\|x_{ijt}\|^2+\|\overline{x}_{ij\cdot}+\overline{x}_{\cdot\cdot t}-2\overline{x}\|^2\right\}\|\overline{x}_{ij\cdot}-\overline{x}_{i\cdot\cdot}-\overline{x}_{\cdot j\cdot}+\overline{x}\|^2 \\
& = O_p((NMT)^{-2}),
\end{aligned}
$$

and

$$
\begin{aligned}
& \frac{1}{NMT}\sum_{i,j,t}h^*_{ijt,4}\left(C_{ijt,4}+C_{ijt,2}\right)^2U'M_{D_4}XX^*_{D_4}X'M_{D_4}U \\
& \leq \left\|X^*_{D_4}\right\|^2_{\mathrm{sp}}\frac{2}{NMT}\sum_{i,j,t}\left\{\|x_{ijt}\|^2+\|\bar{D}'_{ijt,4}X\|^2\right\}\left(C_{ijt,4}+C_{ijt,2}\right)^2U'M_{D_4}XX^*_{D_4}X'M_{D_4}U \\
& \leq \left\|X^*_{D_4}\right\|^2_{\mathrm{sp}}\|X'M_{D_4}U\|^2\frac{2}{NMT}\sum_{i,j,t}\left\{\|x_{ijt}\|^2+\|\overline{x}_{ij\cdot}+\overline{x}_{\cdot\cdot t}-2\overline{x}\|^2\right\}\left(C_{ijt,4}+C_{ijt,2}\right)^2 \\
& = O_p\left(T^{-2}+(NM)^{-2}+(NMT)^{-1}\right)O_p\left(1\right),
\end{aligned}
$$

where we use the fact $\left\|X^*_{D_4}\right\|_{\mathrm{sp}}=O_p((NMT)^{-1})$, $\frac{1}{NMT}\|X'M_{D_4}U\|=O_p(T^{-1}+(NM)^{-1}+(NMT)^{-1/2})$ by Lemma A.4(iii), $\frac{1}{NMT}\sum_{i,j,t}\|x_{ijt}\|^2\left(C_{ijt,4}+C_{ijt,2}\right)^2=O_p\left(1\right)$, and we can readily show $\frac{1}{NMT}\sum_{i,j,t}$

$||\overline{x}_{ij\cdot} + \overline{x}_{\cdot\cdot t} - 2\overline{x}||^2 \left(C_{ijt,4} + C_{ijt,2}\right)^2 = O_p(1).$ Then $E_{2,4}\left(2,1\right) = o_p\left(T^{-1}\right).$ In addition, we have

$$
\begin{aligned}
E_{2,4}\left(2,2\right) &= \frac{1}{NMT}\sum_{i,j,t} h_{ijt,4}^* \left| \left(x_{ijt} - X'\bar{D}_{ijt,2}\right)' \left(X'M_{D_4}X\right)^{-1} \left[X'\left(M_{D_4} - M_{D_2}\right)X\right]\left(X'M_{D_4}X\right)^{-1}\right. \\
&\quad \left. \times X'M_{D_4}U\left(C_{ijt,4} + C_{ijt,2}\right)\right| \\
&\leq \left\|\left(\frac{1}{NMT}X'M_{D_4}X\right)^{-1}\right\|_{\mathrm{sp}}^2 \left\|\frac{1}{NMT}X'\left(M_{D_4} - M_{D_2}\right)X\right\| \left\|\frac{1}{NMT}X'M_{D_4}U\right\| \\
&\quad \times \frac{1}{NMT}\sum_{i,j,t} h_{ijt,4}^* \left\|\left(x_{ijt} - X'\bar{D}_{ijt,2}\right)\left(C_{ijt,4} + C_{ijt,2}\right)\right\| \\
&= O_p\left(1\right) O_p\left(T^{-1}\right) O_p(T^{-1} + (NM)^{-1} + (NMT)^{-1/2}) o_p\left(1\right) = o_p\left(T^{-1}\right),
\end{aligned}
$$

and

$$
\begin{aligned}
E_{2,4}\left(2,3\right) &= \frac{1}{NMT}\sum_{i,j,t} h_{ijt,4}^* \left|\left(x_{ijt} - X'\bar{D}_{ijt,2}\right)' X_{D_2}^* X'\left(M_{D_4} - M_{D_2}\right)U\left(C_{ijt,4} + C_{ijt,2}\right)\right| \\
&\leq \left\|\left(\frac{1}{NMT}X'M_{D_4}X\right)^{-1}\right\|_{\mathrm{sp}} \left\|\frac{1}{NMT}X'\left(M_{D_4} - M_{D_2}\right)U\right\| \\
&\quad \times \frac{1}{NMT}\sum_{i,j,t} h_{ijt,4}^* \left\|\left(x_{ijt} - X'\bar{D}_{ijt,2}\right)'\left(C_{ijt,4} + C_{ijt,2}\right)\right\| \\
&= O_p\left(1\right) O_p\left(T^{-1}\right) o_p\left(1\right) = o_p\left(T^{-1}\right),
\end{aligned}
$$

as we can readily show that $\frac{1}{NMT}\sum_{i,j,t} h_{ijt,4}^* \left\|\left(x_{ijt} - X'\bar{D}_{ijt,2}\right)\left(C_{ijt,4} + C_{ijt,2}\right)\right\| = o_p\left(1\right).$ Consequently, $E_{2,4}\left(2\right) = o_p\left(T^{-1}\right).$

Next, we study $E_{2,4}\left(3\right).$ Noting that $(A_{ijt,4} - A_{ijt,2})C_{ijt,4} = -\left(\bar{D}_{ijt,4} - \bar{D}_{ijt,2}\right)'UC_{ijt,4},$

$$
\begin{aligned}
E_{2,4}\left(3\right) &= \frac{4}{NMT}\sum_{i,j,t} h_{ijt,4}^* \left|\left(\bar{D}_{ijt,4} - \bar{D}_{ijt,2}\right)'UC_{ijt,4}\right| \\
&\leq 4\left\{\frac{1}{NMT}\sum_{i,j,t} h_{ijt,4}^*\right\}^{1/2} \left\{\frac{1}{NMT}\sum_{i,j,t} h_{ijt,4}^* C_{ijt,4}^2 \left[\left(\bar{D}_{ijt,4} - \bar{D}_{ijt,2}\right)'U\right]^2\right\}^{1/2} \\
&= O_p((NMT)^{-1/2}) O_p(T^{-1/2} + (NMT)^{-1/2}) = o_p\left(T^{-1}\right),
\end{aligned}
$$

as we can show that

$$
\begin{aligned}
&\frac{1}{NMT}\sum_{i,j,t} h_{ijt,4}^* C_{ijt,4}^2 \left[\left(\bar{D}_{ijt,4} - \bar{D}_{ijt,2}\right)'U\right]^2 \\
&\leq \frac{2}{NMT}\sum_{i,j,t} x_{ijt}' X_{D_4}^* x_{ijt} C_{ijt,4}^2 \left[\left(\bar{D}_{ijt,4} - \bar{D}_{ijt,2}\right)'U\right]^2 \\
&\quad + \frac{2}{NMT}\sum_{i,j,t} \bar{D}_{ijt,4}' X X_{D_4}^* X'\bar{D}_{ijt,4} C_{ijt,4}^2 \left[\left(\bar{D}_{ijt,4} - \bar{D}_{ijt,2}\right)'U\right]^2 \\
&\leq \left\|X_{D_4}^*\right\|_{\mathrm{sp}} \frac{2}{NMT}\sum_{i,j,t} \left\|x_{ijt}\right\|^2 C_{ijt,4}^2 \left[\left(\bar{D}_{ijt,4} - \bar{D}_{ijt,2}\right)'U\right]^2 \\
&\quad + \left\|X_{D_4}^*\right\|_{\mathrm{sp}} \frac{2}{NMT}\sum_{i,j,t} \left\|\bar{D}_{ijt,4}'X\right\|^2 C_{ijt,4}^2 \left[\left(\bar{D}_{ijt,4} - \bar{D}_{ijt,2}\right)'U\right]^2 \\
&= O_p\left((NMT)^{-1}\right).
\end{aligned}
$$

24

Analogously, we can show that $E_{2,4}(4) = o_p(T^{-1})$. Consequently, we have $E_{2,4} = o_p(T^{-1})$.

(ii) The proof is analogous to that of (i). The main difference is that we now heavily rely on the fact that $\overline{D}'_{ijt,3}A = \overline{a}_{ij\cdot} - \overline{a}$ and $\overline{D}'_{ijt,4}A = \overline{a}_{ij\cdot} + \overline{a}_{\cdot\cdot t} - 2\overline{a}$.

(iii) The proof is analogous to that of (i). The main difference is that we now heavily rely on the fact that $\overline{D}'_{ijt,7}A = \overline{a}_{i\cdot t} + \overline{a}_{\cdot jt} + \overline{a}_{ij\cdot} - \overline{a}_{i\cdot\cdot} - \overline{a}_{\cdot j\cdot} - \overline{a}_{\cdot\cdot t}$ and $\overline{D}'_{ijt,3}A = \overline{a}_{ij\cdot} - \overline{a}$.

(iv) The proof is analogous to that of (i). The main difference is that we now heavily rely on the fact that $\overline{D}'_{ijt,7}A = \overline{a}_{i\cdot t} + \overline{a}_{\cdot jt} + \overline{a}_{ij\cdot} - \overline{a}_{i\cdot\cdot} - \overline{a}_{\cdot j\cdot} - \overline{a}_{\cdot\cdot t}$ and $\overline{D}'_{ijt,4}A = \overline{a}_{ij\cdot} + \overline{a}_{\cdot\cdot t} - 2\overline{a}$.

(v) The proof is analogous to that of (i). The main difference is that we now heavily rely on the fact that $\overline{D}'_{ijt,6}A = \overline{a}_{i\cdot t} + \overline{a}_{\cdot jt} - \overline{a}_{\cdot\cdot t} - \overline{a}$ and $\overline{D}'_{ijt,5}A = \overline{a}_{\cdot jt} - \overline{a}$.

(vi) The proof is analogous to that of (i). The main difference is that we now heavily rely on the fact that $\overline{D}'_{ijt,7}A = \overline{a}_{i\cdot t} + \overline{a}_{\cdot jt} + \overline{a}_{ij\cdot} - \overline{a}_{i\cdot\cdot} - \overline{a}_{\cdot j\cdot} - \overline{a}_{\cdot\cdot t}$ and $\overline{D}'_{ijt,5}A = \overline{a}_{\cdot jt} - \overline{a}$.

(vii) The proof is analogous to that of (i). The main difference is that we now heavily rely on the fact that $\overline{D}'_{ijt,7}A = \overline{a}_{i\cdot t} + \overline{a}_{\cdot jt} + \overline{a}_{ij\cdot} - \overline{a}_{i\cdot\cdot} - \overline{a}_{\cdot j\cdot} - \overline{a}_{\cdot\cdot t}$ and $\overline{D}'_{ijt,6}A = \overline{a}_{i\cdot t} + \overline{a}_{\cdot jt} - \overline{a}_{\cdot\cdot t} - \overline{a}$. ∎

Proof of Lemma A.14. (i) Recall that $\bar{D}_{ijt,m} = D_m(D'_m D_m)^{-1}d_{ijt,m}$ and $X^*_{D_m} = (X'M_{D_m}X)^{-1}$ for $m = 2, ..., 7$. Noting that

$$
\begin{aligned}
h^*_{ijt,4} - h^*_{ijt,2} &= \left(x_{ijt} - X'\bar{D}_{ijt,4}\right)' X^*_{D_4}\left(x_{ijt} - X'\bar{D}_{ijt,4}\right) - \left(x_{ijt} - X'\bar{D}_{ijt,2}\right)' X^*_{D_2}\left(x_{ijt} - X'\bar{D}_{ijt,2}\right) \\
&= x'_{ijt}(X^*_{D_4} - X^*_{D_2})x_{ijt} + (\bar{D}_{ijt,4} - \bar{D}_{ijt,2})'XX^*_{D_4}X'(\bar{D}_{ijt,4} - \bar{D}_{ijt,2}) \\
&\quad + 2\bar{D}'_{ijt,2}XX^*_{D_4}X'(\bar{D}_{ijt,4} - \bar{D}_{ijt,2}) + 2\bar{D}'_{ijt,2}X(X^*_{D_4} - X^*_{D_2})X'\bar{D}_{ijt,2} \\
&\quad + 2x'_{ijt}(X^*_{D_4} - X^*_{D_2})X'\bar{D}_{ijt,4} + 2x'_{ijt}X^*_{D_2}X'\left(\bar{D}_{ijt,4} - \bar{D}_{ijt,2}\right) \\
&\equiv \sum_{l=1}^{6} h_{ijt,24}(l),
\end{aligned}
$$

we have $L_{2,4} \leq \sum_{l=1}^{6}\frac{1}{NMT}\sum_{i,j,t}|h_{ijt,24}(l)|e^2_{ijt,2} \equiv \sum_{l=1}^{6}L_{2,4}(l)$, say. It suffices to show that $L_{2,4}(l) = o_p(T^{-1})$ for $l = 1, 2, ..., 6$.

For $L_{2,4}(1)$, we have

$$
L_{2,4}(1) \leq \left\|X^*_{D_4} - X^*_{D_2}\right\|_{\mathrm{sp}}\frac{1}{NMT}\sum_{i,j,t}\|x_{ijt}\|^2 e^2_{ijt,2} = O_p\left((NMT)^{-1}\right),
$$

as we can readily show that $\frac{1}{NMT}\sum_{i,j,t}\|x_{ijt}\|^2 e^2_{ijt,2} = \frac{1}{NMT}\sum_{i,j,t}\|x_{ijt}\|^2 u^2_{ijt,2} + o_p(1) = O_p(1)$. For $L_{2,4}(2)$, we have

$$
L_{2,4}(2) \leq \left\|X^*_{D_4}\right\|_{\mathrm{sp}}\frac{2}{NMT}\sum_{i,j,t}\left\|(\bar{D}_{ijt,4} - \bar{D}_{ijt,2})'X\right\|^2 e^2_{ijt,2} = O_p((NMT)^{-1})O_p(1) = o_p(T^{-1}).
$$

as we can show that $\frac{1}{NMT}\sum_{i,j,t}\left\|(\bar{D}_{ijt,4} - \bar{D}_{ijt,2})'X\right\|^2 e^2_{ijt,2} = \frac{1}{NMT}\sum_{i,j,t}\|\overline{x}_{ij\cdot} - \overline{x}_{i\cdot\cdot} - \overline{x}_{\cdot j\cdot} + \overline{x}\|^2 e^2_{ijt,2} = O_p(1)$. For $L_{2,4}(3)$, we have by the Cauchy-Schwarz inequality

$$
L_{2,4}(3) \leq 2\left\{\frac{1}{NMT}\sum_{i,j,t}\bar{D}'_{ijt,2}XX^*_{D_4}X'\bar{D}_{ijt,2}e^2_{ijt,2}\right\}^{1/2}\{L_{2,4}(2)\}^{1/2}.
$$

We have

$$
\begin{aligned}
\frac{1}{NMT}\sum_{i,j,t}\bar{D}'_{ijt,2}XX^*_{D_4}X'\bar{D}_{ijt,2}e^2_{ijt,2} &\leq \left\|X^*_{D_4}\right\|_{\mathrm{sp}}\frac{1}{NMT}\sum_{i,j,t}\|\overline{x}_{i\cdot\cdot} + \overline{x}_{\cdot j\cdot} + \overline{x}_{\cdot\cdot t} - 3\overline{x}\|^2 e^2_{ijt,2} \\
&= O_p\left((NMT)^{-1}\right).
\end{aligned}
$$

25

It follows that $L_{2,4}(3) = O_p\left((NMT)^{-1/2}\right)\{L_{2,4}(2)\}^{1/2} = o_p\left(T^{-1}\right).$

Noting that

$$
\begin{aligned}
\left\|X(X_{D_4}^* - X_{D_2}^*)X'\right\|_{\mathrm{sp}} &= \left\|X[(X'M_{D_4}X)^{-1} - (X'M_{D_2}X)^{-1}]X'\right\|_{\mathrm{sp}} \\
&= \left\|X\left(X'M_{D_4}X\right)^{-1}[X'(M_{D_4} - M_{D_2})X]\left(X'M_{D_2}X\right)^{-1}X'\right\|_{\mathrm{sp}} \\
&\leq \left\|\left(\frac{1}{NMT}X'M_{D_4}X\right)^{-1}\right\|_{\mathrm{sp}}\left\|\left(\frac{1}{NMT}X'M_{D_2}X\right)^{-1}\right\|_{\mathrm{sp}}\frac{1}{NMT}\|X\|_{\mathrm{sp}}^2 \\
&\quad \times \left\|\frac{1}{NMT}X'(M_{D_4} - M_{D_2})X\right\| \\
&= O_p(1)O_p(1)O_p(1)O_p\left(T^{-1}\right) = O_p\left(T^{-1}\right)
\end{aligned}
$$

and similarly, $\left\|(X_{D_4}^* - X_{D_2}^*)X'\right\|_{\mathrm{sp}} = (NMT)^{-1/2}O_p\left(T^{-1}\right),$ we have

$$
\begin{aligned}
L_{2,4}(4) &= \frac{2}{NMT}\sum_{i,j,t}\left|\bar{D}_{ijt,2}'X(X_{D_4}^* - X_{D_2}^*)X'\bar{D}_{ijt,2}\right|e_{ijt,2}^2 \\
&\leq \max_{i,j,t}\left\|\bar{D}_{ijt,2}\right\|^2\left\|X(X_{D_4}^* - X_{D_2}^*)X'\right\|_{\mathrm{sp}}\frac{2}{NMT}\sum_{i,j,t}e_{ijt,2}^2 \leq \bar{d}_2 O_p\left(T^{-1}\right)O_p(1) = o_p\left(T^{-1}\right),
\end{aligned}
$$

and

$$
\begin{aligned}
L_{2,4}(5) &= \frac{2}{NMT}\sum_{i,j,t}\left|x_{ijt}'(X_{D_4}^* - X_{D_2}^*)X'\bar{D}_{ijt,4}\right|e_{ijt,2}^2 \\
&\leq \max_{i,j,t}\left\|\bar{D}_{ijt,2}\right\|\left\|(X_{D_4}^* - X_{D_2}^*)X\right\|_{\mathrm{sp}}\frac{2}{NMT}\sum_{i,j,t}\|x_{ijt}\|e_{ijt,2}^2 \\
&= \bar{d}_2^{1/2}O_p((NMT)^{-1/2})O_p(T^{-1})O_p(1) = o_p\left(T^{-1}\right),
\end{aligned}
$$

where we use the fact that $\left\|\bar{D}_{ijt,2}\right\|^2 = \bar{d}_2.$

For $L_{2,4}(6),$ we have

$$
\begin{aligned}
L_{2,4}(6) &= \frac{2}{NMT}\sum_{i,j,t}\left|x_{ijt}'X_{D_2}^*X'\left(\bar{D}_{ijt,4} - \bar{D}_{ijt,2}\right)\right|e_{ijt,2}^2 \\
&\leq 2\left\{\frac{1}{NMT}\sum_{i,j,t}x_{ijt}'X_{D_2}^*x_{ijt}e_{ijt,2}^2\right\}^{1/2}\{L_{2,4}(2)\}^{1/2} \\
&\leq \left\|X_{D_2}^*\right\|_{\mathrm{sp}}^{1/2}\left\{\frac{1}{NMT}\sum_{i,j,t}\|x_{ijt}\|^2e_{ijt,2}^2\right\}^{1/2}\{L_{2,4}(2)\}^{1/2} \\
&= O_p\left((NMT)^{-1/2}\right)O_p(1)o_p\left(T^{-1/2}\right) = o_p\left(T^{-1}\right).
\end{aligned}
$$

This completes the proof of (i).

(ii)-(vii) The proofs are completely analogous to that of (i). the main difference is that we need to use the probability order of $\overline{D}_{ijt,2}'A = \bar{a}_{i..} + \bar{a}_{.j.} + \bar{a}_{..t} - 3\bar{a},$ $\overline{D}_{ijt,3}'A = \bar{a}_{ij.} - \bar{a},$ $\overline{D}_{ijt,4}'A = \bar{a}_{ij.} + \bar{a}_{..t} - 2\bar{a},$ $\overline{D}_{ijt,5}'A = \bar{a}_{.jt} - \bar{a},$ $\overline{D}_{ijt,6}'A = \bar{a}_{i.t} + \bar{a}_{.jt} - \bar{a}_{..t} - \bar{a}$ and $\overline{D}_{ijt,7}'A = \bar{a}_{i.t} + \bar{a}_{.jt} + \bar{a}_{ij.} - \bar{a}_{i..} - \bar{a}_{.j.} - \bar{a}_{..t}$ in order. ∎

Proof of Lemma B.1. (i) When Model 7 is used, the residual vector is given by $\hat{\mathbf{U}} = M_{D_7}U - M_{D_7}X(\hat{\beta} -$

β). By Lemma A.2(v), $M_{D_7} = I_{NMT} - P_{IJ}^* - P_{IT}^* - P_{JT}^*$, where

$$P_{IJ}^* = I_N \otimes (I_M - \frac{J_M}{M}) \otimes \frac{J_T}{T} = I_{NM} \otimes \frac{J_T}{T} - I_N \otimes \frac{J_{MT}}{MT},$$

$$P_{IT}^* = (I_N - \frac{J_N}{N}) \otimes \frac{J_M}{M} \otimes I_T = I_N \otimes \frac{J_M}{M} \otimes I_T - \frac{J_{NM}}{NM} \otimes I_T, \text{ and}$$

$$P_{JT}^* = \frac{J_N}{N} \otimes I_M \otimes (I_T - \frac{J_T}{T}) = \frac{J_N}{N} \otimes I_{MT} - \frac{J_N}{N} \otimes I_M \otimes \frac{J_T}{T}.$$

So a typical element of $M_{D_7} U$ is given by

$$u_{ijt} - (\bar{u}_{ij.} - \bar{u}_{i..}) - (\bar{u}_{i\cdot t} - \bar{u}_{..t}) - (\bar{u}_{\cdot jt} - \bar{u}_{\cdot j.}) \equiv \ddot{u}_{ijt}.$$

Therefore, $\hat{u}_{ijt} = \ddot{u}_{ijt} - (\hat{\beta} - \beta)' \ddot{x}_{ijt}$, where $\ddot{x}_{ijt}$ is defined analogously to $\ddot{u}_{ijt}$. Under the stated assumptions, we can readily show that $\hat{\beta} - \beta = O_p(\eta_{NMT})$. It follows that

$$\frac{1}{NMT_1}(\hat{\mathbf{Z}}'\hat{\mathbf{Z}} - \mathbf{Z}'\mathbf{Z}) = (\hat{\beta} - \beta)' \frac{1}{NMT_1} \sum_{i,j} \sum_{t=1}^{T-1} \ddot{x}_{ijt}\ddot{x}'_{ijt}(\hat{\beta} - \beta) - (\hat{\beta} - \beta)' \frac{2}{NMT_1} \sum_{i,j} \sum_{t=1}^{T-1} \ddot{x}_{ijt}\ddot{u}_{ijt}$$

$$= O_p(\eta_{NMT}) \cdot O_p(1) \cdot O_p(\eta_{NMT}) + O_p(\eta_{NMT}) \cdot O_p(1) = O_p(\eta_{NMT}),$$

where we use the fact that $\frac{1}{NMT_1} \sum_{i,j} \sum_{t=1}^{T-1} \ddot{x}_{ijt}\ddot{x}'_{ijt} = O_p(1)$ and that $\frac{1}{NMT_1} \sum_{i,j} \sum_{t=1}^{T-1} \ddot{x}_{ijt}\ddot{u}_{ijt} = O_p(1)$.

(ii) The analysis is analogous to that in (i) and thus omitted.

(iii) First we need to prove $(\mathbf{Z}'\mathbf{Z})^{-1} \mathbf{Z}'\mathbf{U} - \rho = O_p(\eta_{NMT}^2)$. When $p = 1$, we have $\rho = \rho_1$, $v_{ijt} = u_{ijt} - \rho u_{ij,t-1}$ and

$$\ddot{u}_{ij,t} - \rho \ddot{u}_{ij,t-1} = v_{ijt} - \bar{v}_{i\cdot t} - \bar{v}_{\cdot jt} + \bar{v}_{..t} + (1 - \rho)(\bar{u}_{i..} + \bar{u}_{\cdot j.} - \bar{u}_{ij.}).$$

Then

$$(\mathbf{Z}'\mathbf{Z})^{-1} \mathbf{Z}'\mathbf{U} - \rho = \left(\sum_{i,j,t} \ddot{u}_{ij,t-1}^2\right)^{-1} \left(\sum_{i,j,t} \ddot{u}_{ij,t-1}\ddot{u}_{ij,t}\right) - \rho$$

$$= \left(\sum_{i,j,t} \ddot{u}_{ij,t-1}^2\right)^{-1} \sum_{i,j,t} \ddot{u}_{ij,t-1} [v_{ijt} - \bar{v}_{i\cdot t} - \bar{v}_{\cdot jt} + \bar{v}_{..t} + (1 - \rho)(\bar{u}_{i..} + \bar{u}_{\cdot j.} - \bar{u}_{ij.})]$$

$$= \left(\sum_{i,j,t} \ddot{u}_{ij,t-1}^2\right)^{-1} \sum_{i,j,t} \ddot{u}_{ij,t-1}v_{ijt} - (1 - \rho) \left(\sum_{i,j,t} \ddot{u}_{ij,t-1}^2\right)^{-1} \sum_{i,j,t} \ddot{u}_{ij,t-1}\bar{u}_{ij.}$$

$$+ O_p((NMT)^{-1})$$

where the third equality follows from the fact that $\frac{1}{N} \sum_{i=1}^{N} \ddot{u}_{ijt} = \frac{1}{M} \sum_{j=1}^{M} \ddot{u}_{ijt} = \bar{u}$, $\frac{1}{NT_1} \sum_{i=1}^{N} \sum_{t=2}^{T} \bar{v}_{i\cdot t} = \frac{1}{MT_1} \sum_{j=1}^{M} \sum_{t=2}^{T} \bar{v}_{..t} = \bar{v}$, and both $\bar{u}$ and $\bar{v}$ are $O_p((NMT)^{-1/2})$. Similarly, noting that $\frac{1}{N} \sum_{i=1}^{N} \ddot{u}_{ijt} = \frac{1}{M} \sum_{j=1}^{M} \ddot{u}_{ijt} = \bar{u}$, we have

$$\frac{1}{NM} \sum_{i,j} \ddot{u}_{ijT}\bar{u}_{ij.} = \frac{1}{NM} \sum_{i,j} (u_{ijT} - \bar{u}_{ij.} - \bar{u}_{i\cdot T} - \bar{u}_{\cdot jT} + \bar{u}_{i..} + \bar{u}_{..T} + \bar{u}_{\cdot j.}) \bar{u}_{ij.}$$

$$= \frac{1}{NM} \sum_{i,j} (u_{ijT} - \bar{u}_{ij.}) \bar{u}_{ij.} + \bar{u}_{..T}\bar{u}$$

$$= O_p(T^{-1}) + O_p((NM)^{-1/2})O_p((NMT)^{-1/2}),$$

27

and

$$\frac{1}{NMT}\sum_{i,j,t}\ddot{u}_{ij,t-1}\overline{u}_{ij\cdot} = \frac{1}{NMT}\sum_{i,j}\left(\sum_{t=1}^{T}\ddot{u}_{ijt} - \ddot{u}_{ijT}\right)\overline{u}_{ij\cdot}$$

$$= \overline{u}\frac{1}{NM}\sum_{i,j}\overline{u}_{ij\cdot} - \frac{1}{NM}\sum_{i,j}\ddot{u}_{ijT}\overline{u}_{ij\cdot}$$

$$= O_p((NMT)^{-1}) + O_p(T^{-1} + (NM)^{-1}T^{-1/2}).$$

In addition

$$\frac{1}{NMT}\sum_{i,j,t}v_{ijt}\ddot{u}_{ij,t-1} = \frac{1}{NMT}\sum_{i,j,t}v_{ijt}\left(u_{ij,t-1} - \overline{u}_{i\cdot,t-1} - \overline{u}_{ij\cdot} - \overline{u}_{\cdot j,t-1} + \overline{u}_{i\cdot\cdot} + \overline{u}_{\cdot j\cdot} + \overline{u}_{\cdot\cdot t-1}\right)$$

$$= O_p((NMT)^{-1/2} + N^{-1} + M^{-1} + T^{-1} + (MT)^{-1} + (NT)^{-1} + (MN)^{-1})$$

$$= O_p(\eta_{NMT}).$$

It follows that $(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U} - \rho = O_p(\eta_{NMT})$. Then $(\hat{\mathbf{Z}}'\hat{\mathbf{Z}})^{-1}\hat{\mathbf{Z}}'\hat{\mathbf{U}} - \rho = O_p(\eta_{NMT})$ follows by noting the results in part (i)-(ii). ∎

Proof of Lemma B.2. (i) By the Assumptions A.5(iv) and A.1(iv)-(v) and noting that $\tilde{z}_{ijt,1} = x_{ijt} - \rho x_{ij,t-1}$, we have

$$Q_1 = \left\{\frac{1}{NMT_1}\sum_{i,j,t}v_{ijt}(x_{ijt} - \rho x_{ij,t-1})\right\}(X'X)^{-1}X'U$$

$$= O_p((NMT)^{-1/2})O_p((NMT)^{-1/2}) = O_p((NMT)^{-1}).$$

To prove (ii)-(vii), noting that $\tilde{z}_{ijt,m} = z_{ijt,m} - \rho z_{ij,t-1,m} = ((x_{ijt} - \rho x_{ij,t-1})', (d_{ijt,m} - \rho d_{ij,t-1,m})')'$ $\equiv (\tilde{x}'_{ijt}, \tilde{d}'_{ijt,m})'$ for $m = 2, ..., 7$, we first apply Lemma A.1 and make the following decomposition

$$Q_m = \frac{1}{NMT_1}\sum_{i,j,t}v_{ijt}(\tilde{x}'_{ijt}, \tilde{d}'_{ijt,m})(Z'_m Z_m)^{-1}\begin{pmatrix} X'U \\ D'_m U \end{pmatrix}$$

$$= \frac{1}{NMT_1}\sum_{i,j,t}v_{ijt}\left\{\tilde{x}'_{ijt}(X'M_{D_m}X)^{-1}X'M_{D_m}U - \tilde{d}'_{ijt}(D'_m D_m)^{-1}D'_m X(X'M_{D_m}X)^{-1}X'M_{D_m}U\right.$$

$$\left. + \tilde{d}'_{ijt,m}(D'_m D_m)^{-1}D'_m U\right\}$$

$$\equiv Q_{m,1} - Q_{m,2} + Q_{m,3}, \text{ say.} \tag{C.1}$$

Let $\bar{Q}_m = \frac{1}{NMT_1}\sum_{i,j,t}v_{ijt}\tilde{d}'_{ijt,m}(D'_m D_m)^{-1}D'_m X$ and $S_m = (X'M_{D_m}X)^{-1}X'M_{D_m}U$ for $m = 2, ..., 7$. Define

$$\xi^{a,0}_{ijt,m} \equiv A'D_m(D'_m D_m)^{-1}d_{ijt,m} \text{ and } \xi^{a}_{ijt,m} \equiv A'D_m(D'_m D_m)^{-1}\tilde{d}_{ijt,m} \text{ for } A = \{a_{ijt}\} \text{ and } m = 2, ..., 7. \tag{C.2}$$

Noting that $\xi^{a,0}_{ijt,2} = (\overline{a}_{i\cdot\cdot} - \overline{a}) + (\overline{a}_{\cdot j\cdot} - \overline{a}) + (\overline{a}_{\cdot\cdot t} - \overline{a})$, $\xi^{a,0}_{ijt,3} = \overline{a}_{ij\cdot} - \overline{a}$, $\xi^{a,0}_{ijt,4} = (\overline{a}_{ij\cdot} - \overline{a}) + (\overline{a}_{\cdot\cdot t} - \overline{a})$, $\xi^{a,0}_{ijt,5} = \overline{a}_{\cdot jt} - \overline{a}$, $\xi^{a,0}_{ijt,6} = (\overline{a}_{i\cdot t} - \overline{a}_{\cdot\cdot t}) + (\overline{a}_{\cdot jt} - \overline{a})$, and $\xi^{a,0}_{ijt,7} = (\overline{a}_{i\cdot t} - \overline{a}_{\cdot\cdot t}) + (\overline{a}_{ij\cdot} - \overline{a}_{i\cdot\cdot}) + (\overline{a}_{\cdot jt} - \overline{a}_{\cdot j\cdot})$, we have that

$$\begin{aligned}
\xi^{a}_{ijt,2} &= (1-\rho)(\overline{a}_{i\cdot\cdot} + \overline{a}_{\cdot j\cdot} - 2\overline{a}) + (1-\rho L)(\overline{a}_{\cdot\cdot t} - \overline{a}), \\
\xi^{a}_{ijt,3} &= (1-\rho)(\overline{a}_{ij\cdot} - \overline{a}) \\
\xi^{a}_{ijt,4} &= (1-\rho)(\overline{a}_{ij\cdot} - \overline{a}) + (1-\rho L)(\overline{a}_{\cdot\cdot t} - \overline{a}), \\
\xi^{a}_{ijt,5} &= (1-\rho L)(\overline{a}_{\cdot jt} - \overline{a}), \\
\xi^{a}_{ijt,6} &= (1-\rho L)[(\overline{a}_{i\cdot t} - \overline{u}_{\cdot\cdot t}) + (\overline{a}_{\cdot jt} - \overline{a})], \\
\xi^{a}_{ijt,7} &= (1-\rho)(\overline{a}_{ij\cdot} - \overline{a}_{i\cdot\cdot}) + (1-\rho L)[(\overline{a}_{i\cdot t} - \overline{a}_{\cdot\cdot t}) + (\overline{a}_{\cdot jt} - \overline{a}_{\cdot j\cdot})].
\end{aligned} \tag{C.3}$$

28

We are ready to prove (ii)-(vii) in order.

(ii) For $Q_{2,1}$ we have by Lemma A.4(i) and Assumptions A.1(v) and A.5(iv),

$$Q_{2,1} = \left\{ \frac{1}{NMT_1} \sum_{i,j,t} v_{ijt} \tilde{x}'_{ijt} \right\} (X'M_{D_2}X)^{-1} X'M_{D_2}U$$

$$= O_p((NMT)^{-1/2})O_p((NM)^{-1} + (MT)^{-1} + (NT)^{-1} + (NMT)^{-1/2}).$$

Note that $Q_{2,2} = \bar{Q}_2 S_2$. Using $\xi^a_{ijt,2}$ in (C.3) with $a = x$, we have by Assumptions A.3 and A.5 and the Cauchy-Schwarz inequality

$$\bar{Q}_2 = \frac{1}{NMT_1} \sum_{i,j,t} v_{ijt}[(1-\rho)(\overline{x}_{i\cdot\cdot} + \overline{x}_{\cdot j\cdot} - 2\overline{x}) + (1-\rho L)(\overline{x}_{\cdot\cdot t} - \overline{x})]$$

$$= \frac{1-\rho}{N} \sum_i \overline{v}_{i\cdot\cdot}(\overline{x}_{i\cdot\cdot} - \overline{x}) + \frac{1-\rho}{M} \sum_j \overline{v}_{\cdot j\cdot}(\overline{x}_{\cdot j\cdot} - \overline{x}) + \frac{1}{T_1} \sum_t \overline{v}_{\cdot\cdot t}(1-\rho L)(\overline{x}_{\cdot\cdot t} - \overline{x})$$

$$= O_p((MT)^{-1} + (NT)^{-1} + (MN)^{-1}).$$

Then by Lemma A.4(i) and Assumptions A.1(v), $Q_{2,2} = O_p((MT)^{-2} + (NT)^{-2} + (MN)^{-2} + (NMT)^{-1})$. Using $\xi^a_{ijt,2}$ in (C.3) with $a = u$

$$Q_{2,3} = \frac{1}{NMT_1} \sum_{i,j,t} v_{ijt} \tilde{d}'_{ijt,2}(D'_2 D_2)^{-1} D'_2 U$$

$$= \frac{1}{NMT_1} \sum_{i,j,t} v_{ijt} [(1-\rho)(\overline{u}_{i\cdot\cdot} + \overline{u}_{\cdot j\cdot} - 2\overline{u}) + (1-\rho L)(\overline{u}_{\cdot\cdot t} - \overline{u})]$$

$$= \frac{1}{NMT_1} \sum_{i,j,t} v_{ijt} [(1-\rho)(\overline{u}_{i\cdot\cdot} + \overline{u}_{\cdot j\cdot}) + (1-\rho L)\overline{u}_{\cdot\cdot t}] + O_p((NMT)^{-1}).$$

Summarizing the above results yields $Q_2 = \frac{1}{NMT_1} \sum_{i,j,t} v_{ijt} [(1-\rho)(\overline{u}_{i\cdot\cdot} + \overline{u}_{\cdot j\cdot}) + (1-\rho L)\overline{u}_{\cdot\cdot t}] + O_p ((MT)^{-2} + (NT)^{-2} + (MN)^{-2} + (NMT)^{-1})$.

(iii) For $Q_{3,1}$ we have by Lemma A.4(ii) and Assumptions A.1(v) and A.5(iv), $Q_{3,1} = O_p((NMT)^{-1/2}) \times O_p(T^{-1} + (NMT)^{-1/2})$. Note that $Q_{3,2} = \bar{Q}_3 S_3$. Using $\xi^a_{ijt,3}$ in (C.3) with $a = x$, we have by Assumptions A.3 and A.6

$$\bar{Q}_3 = \frac{1-\rho}{NMT_1} \sum_{i,j,t} v_{ijt}(\overline{x}_{ij\cdot} - \overline{x}) = \frac{1-\rho}{NM} \sum_{i,j} \overline{v}_{ij\cdot}(\overline{x}_{ij\cdot} - \overline{x}) = O_p(T^{-1}).$$

Then by Lemma A.4(ii) and Assumptions A.1(v), $Q_{3,2} = O_p(T^{-2} + (NMT)^{-1})$. Using $\xi^a_{ijt,3}$ in (C.3) with $a = u$,

$$Q_{3,3} = \frac{1}{NMT_1} \sum_{i,j,t} v_{ijt} \tilde{d}'_{ijt,3}(D'_3 D_3)^{-1} D'_3 U$$

$$= \frac{1-\rho}{NMT_1} \sum_{i,j,t} v_{ijt}(\overline{u}_{ij\cdot} - \overline{u}) = \frac{1-\rho}{NMT_1} \sum_{i,j,t} v_{ijt}\overline{u}_{ij\cdot} + O_p((NMT)^{-1}).$$

Summarizing the above results yields $Q_3 = \frac{1-\rho}{NMT_1} \sum_{i,j,t} v_{ijt}\overline{u}_{ij\cdot} + O_p(T^{-2} + (NMT)^{-1})$.

(iv) For $Q_{4,1}$ we have by Lemma A.4(iii) and Assumptions A.1(v) and A.5(iv), $Q_{4,1} = O_p((NMT)^{-1/2}) \times O_p(T^{-1} + (NM)^{-1} + (NMT)^{-1/2})$. Note that $Q_{4,2} = \bar{Q}_4 S_4$. Using $\xi^a_{ijt,4}$ in (C.3) with $a = x$, we have by Assumptions A.3 and A.6

$$\bar{Q}_4 = \frac{1}{NMT_1} \sum_{i,j,t} v_{ijt} [(1-\rho)(\overline{x}_{ij\cdot} - \overline{x}) + (1-\rho L)(\overline{x}_{\cdot\cdot t} - \overline{x})]$$

$$= \frac{1-\rho}{NM} \sum_{i,j} \overline{v}_{ij\cdot}(\overline{x}_{ij\cdot} - \overline{x}) + \frac{1}{T_1} \sum_t \overline{v}_{\cdot\cdot t}(1-\rho L)(\overline{x}_{\cdot\cdot t} - \overline{x}) = O_p(T^{-1} + (NM)^{-1}).$$

Then by Lemma A.4(i) and Assumption A.1(v), $Q_{4,2} = O_p(T^{-2} + (NM)^{-2} + (NMT)^{-1})$. Using $\xi^a_{ijt,4}$ in (C.3) with $a = u$,

$$
\begin{aligned}
Q_{4,3} &= \frac{1}{NMT_1} \sum_{i,j,t} v_{ijt} \left[(1-\rho)(\overline{u}_{ij\cdot} - \overline{u}) + (1-\rho L)(\overline{u}_{\cdot\cdot t} - \overline{u}) \right] \\
&= \frac{1}{NMT_1} \sum_{i,j,t} v_{ijt} \left[(1-\rho)\overline{u}_{ij\cdot} + (1-\rho L)\overline{u}_{\cdot\cdot t} \right] + O_p((NMT)^{-1}).
\end{aligned}
$$

Summarizing the above results yields $Q_4 = \frac{1}{NMT_1} \sum_{i,j,t} v_{ijt} \left[(1-\rho)\overline{u}_{ij\cdot} + (1-\rho L)\overline{u}_{\cdot\cdot t} \right] + O_p(T^{-2} + (NM)^{-2} + (NMT)^{-1})$.

(v) For $Q_{5,1}$ we have by Lemma A.4(iv) and Assumptions A.1(v) and A.5(iv), $Q_{5,1} = O_p((NMT)^{-1/2}) \times O_p(N^{-1} + (NMT)^{-1/2})$. Note that $Q_{5,2} = \bar{Q}_5 S_5$. Using $\xi^a_{ijt,5}$ in (C.3) with $a = x$, we have by Assumptions A.3 and A.6

$$
\bar{Q}_5 = \frac{1}{NMT_1} \sum_{i,j,t} v_{ijt}(1-\rho L)(\overline{x}_{\cdot jt} - \overline{x}) = \frac{1}{MT_1} \sum_{j,t} \overline{v}_{\cdot jt}(1-\rho L)(\overline{x}_{\cdot jt} - \overline{x}) = O_p(N^{-1}).
$$

Then by Lemma A.4(iv) and Assumptions A.1(v), $Q_{5,2} = O_p(N^{-2} + (NMT)^{-1})$. Using $\xi^a_{ijt,5}$ in (C.3) with $a = u$,

$$
Q_{5,3} = \frac{1}{NMT_1} \sum_{i,j,t} v_{ijt}(1-\rho L)(\overline{u}_{\cdot jt} - \overline{u}) = \frac{1}{NMT_1} \sum_{i,j,t} v_{ijt}(1-\rho L)\overline{u}_{\cdot jt} + O_p((NMT)^{-1}).
$$

Summarizing the above results yields $Q_5 = \frac{1}{NMT_1} \sum_{i,j,t} v_{ijt} \left[(1-\rho L)\overline{u}_{\cdot jt} \right] + O_p(N^{-2} + (NMT)^{-1})$.

(vi) For $Q_{6,1}$ we have by Lemma A.4(v) and Assumptions A.1(v) and A.5(iv), $Q_{6,1} = O_p((NMT)^{-1/2}) \times O_p(N^{-1} + M^{-1} + (NMT)^{-1/2})$. Note that $Q_{6,2} = \bar{Q}_6 S_6$. Using $\xi^a_{ijt,6}$ in (C.3) with $a = x$, we have

$$
\begin{aligned}
\bar{Q}_6 &= \frac{1}{NMT_1} \sum_{i,j,t} v_{ijt}(1-\rho L) \left[(\overline{x}_{i\cdot t} - \overline{x}_{\cdot\cdot t}) + (\overline{x}_{\cdot jt} - \overline{x}) \right] \\
&= \frac{1}{NT_1} \sum_{i,t} \overline{v}_{i\cdot t}(1-\rho L)(\overline{x}_{i\cdot t} - \overline{x}_{\cdot\cdot t}) + (\overline{x}_{\cdot jt} - \overline{x}) + \frac{1}{MT_1} \sum_{j,t} \overline{v}_{\cdot jt}(1-\rho L)(\overline{x}_{\cdot jt} - \overline{x}) \\
&= O_p(N^{-1} + M^{-1}).
\end{aligned}
$$

Then by Lemma A.4(v) and Assumptions A.1(v), $Q_{6,2} = O_p(N^{-2} + M^{-2} + (NMT)^{-1})$. Using $\xi^a_{ijt,6}$ in (C.3) with $a = u$,

$$
\begin{aligned}
Q_{6,3} &= \frac{1}{NMT_1} \sum_{i,j,t} v_{ijt} \tilde{d}'_{ijt,6} (D'_6 D_6)^{-1} D'_6 U \\
&= \frac{1}{NMT_1} \sum_{i,j,t} v_{ijt}(1-\rho L) \left[(\overline{u}_{i\cdot t} - \overline{u}_{\cdot\cdot t}) + (\overline{u}_{\cdot jt} - \overline{u}) \right] \\
&= \frac{1}{NMT_1} \sum_{i,j,t} v_{ijt}(1-\rho L)(\overline{u}_{i\cdot t} + \overline{u}_{\cdot jt}) + O_p((NM)^{-1} + (NMT)^{-1}).
\end{aligned}
$$

Summarizing the above results yields $Q_6 = \frac{1}{NMT_1} \sum_{i,j,t} v_{ijt}(1-\rho L)(\overline{u}_{i\cdot t} + \overline{u}_{\cdot jt}) + O_p(N^{-2} + M^{-2} + (NMT)^{-1})$.

(vii) For $Q_{7,1}$ we have by Lemma A.4(vi) and Assumptions A.1(v) and A.5(iv), $Q_{7,1} = O_p((NMT)^{-1/2}) \times O_p(N^{-1} + M^{-1} + T^{-1} + (NMT)^{-1/2})$. Note that $Q_{7,2} = \bar{Q}_7 S_7$. Using $\xi^a_{ijt,7}$ in (C.3) with $a = x$, we have

by Assumptions A.3 and A.6

$$
\begin{aligned}
\bar{Q}_7 &= \frac{1}{NMT_1}\sum_{i,j,t} v_{ijt}\left\{(1-\rho)(\overline{x}_{ij\cdot}-\overline{x}_{i\cdot\cdot})+(1-\rho L)\left[(\overline{x}_{i\cdot t}-\overline{x}_{\cdot\cdot t})+(\overline{x}_{\cdot jt}-\overline{x}_{\cdot j\cdot})\right]\right\} \\
&= \frac{1-\rho}{NM}\sum_{i,j}\overline{v}_{ij\cdot}\left[(\overline{x}_{ij\cdot}-\overline{x})-(\overline{x}_{i\cdot\cdot}-\overline{x})\right]+\frac{1}{NT_1}\sum_{i,t}\overline{v}_{i\cdot t}\left\{(1-\rho L)\left[(\overline{x}_{i\cdot t}-\overline{x})-(\overline{x}_{\cdot\cdot t}-\overline{x})\right]\right\} \\
&\quad +\frac{1}{MT_1}\sum_{j,t}\overline{v}_{\cdot j t}\left\{(1-\rho L)\left[(\overline{x}_{\cdot jt}-\overline{x})-(\overline{x}_{\cdot j\cdot}-\overline{x})\right]\right\} \\
&= O_p(T^{-1}+M^{-1}+N^{-1}).
\end{aligned}
$$

Then by Lemma A.4(vi) and Assumptions A.1(v), $Q_{7,2}=O_p(N^{-2}+M^{-2}+T^{-2}+(NMT)^{-1})$. Using $\xi^a_{ijt,7}$ in (C.3) with $a=u$,

$$
\begin{aligned}
Q_{7,3} &= \frac{1}{NMT_1}\sum_{i,j,t} v_{ijt}\left\{(1-\rho)(\overline{u}_{ij\cdot}-\overline{u}_{i\cdot\cdot})+(1-\rho L)[(\overline{u}_{i\cdot t}-\overline{u}_{\cdot\cdot t})+(\overline{u}_{\cdot jt}-\overline{u}_{\cdot j\cdot})]\right\} \\
&= \frac{1}{NMT_1}\sum_{i,j,t} v_{ijt}\left[(1-\rho)\overline{u}_{ij\cdot}+(1-\rho L)(\overline{u}_{i\cdot t}+\overline{u}_{\cdot jt})\right]+O_p((NM)^{-1}+(NT)^{-1}+(MT)^{-1}).
\end{aligned}
$$

Summarizing the above results, we have $Q_7=\frac{1}{NMT_1}\sum_{i,j,t} v_{ijt}\left[(1-\rho)\overline{u}_{ij\cdot}+(1-\rho L)(\overline{u}_{i\cdot t}+\overline{u}_{\cdot jt})\right]+O_p(N^{-2}+M^{-2}+T^{-2}+(NMT)^{-1})$. ∎

Proof of Lemma B.3. (i) This basically follows from the proof of Lemma B.2(ii) and (iv). The main difference is that some terms in the expansion of Q_2 are cancelling with the corresponding terms in Q_4. To see this, we continue to use the expansion in (C.1). Then $Q_4-Q_2=\sum_{l=1}^{3}(Q_{4,l}-Q_{2,l})$. It suffices to show that $Q_{4,3}-Q_{2,3}=\frac{1-\rho}{NM}\sum_{i,j}\overline{v}_{ij}\cdot\overline{u}_{ij\cdot}+O_p(T^{-2}+(NMT)^{-1})$ and $Q_{4,l}-Q_{2,l}=O_p(T^{-2}+(NMT)^{-1})$ for $l=1,2$. Recall that $\bar{Q}_m=\frac{1}{NMT_1}\sum_{i,j,t} v_{ijt}\tilde{d}'_{ijt,m}(D'_m D_m)^{-1}D'_m X$ and $S_m=(X'M_{D_m}X)^{-1}X'M_{D_m}U$ for $m=2,...,7$.

For $Q_{4,1}-Q_{2,1}$, we have

$$
\begin{aligned}
Q_{4,1}-Q_{2,1} &= \hat{Q}(S_4-S_2)=\hat{Q}[(X'M_{D_4}X)^{-1}X'(M_{D_2}-M_{D_4})X(X'M_{D_2}X)^{-1}]X'M_{D_4}U \\
&\quad +(X'M_{D_2}X)^{-1}X'(M_{D_4}-M_{D_2})U],
\end{aligned}
$$

where $\hat{Q}=\frac{1}{NMT_1}\sum_{i,j,t} v_{ijt}\tilde{x}'_{ijt}=O_p((NMT)^{-1/2})$ by Assumption A.5(iv). Noting that

$$
\frac{1}{NMT}X'(M_{D_2}-M_{D_4})X(X'M_{D_2}X)^{-1}X'M_{D_4}U=O_p\left(T^{-1}\right)O_p(T^{-1}+(NM)^{-1}+(NMT)^{-1/2})
$$

and

$$
\begin{aligned}
\frac{1}{NMT}X'(M_{D_4}-M_{D_2})U &= \frac{1}{NMT}\sum_{i,j}(\overline{x}_{ij\cdot}-\overline{x})(\overline{u}_{ij\cdot}-\overline{u})+O_p((NT)^{-1}+(MT)^{-1}) \\
&= \frac{1}{NMT}\sum_{i,j}\overline{x}_{ij\cdot}\overline{u}_{ij\cdot}+O_p((NT)^{-1}+(MT)^{-1})=O_p(T^{-1})
\end{aligned}
$$

by Lemmas A.12(i) and A.4(iv), we have $S_4-S_2=O_p\left(T^{-1}\right)$ and $Q_{4,1}-Q_{2,1}=O_p((NMT)^{-1/2})O_p(T^{-1})$ $=O_p(T^{-2}+(NMT)^{-1})$.

For $Q_{4,2}-Q_{2,2}$, we make the following decomposition

$$
Q_{4,2}-Q_{2,2}=\bar{Q}_4 S_4-\bar{Q}_2 S_2=(\bar{Q}_4-\bar{Q}_2)S_4+\bar{Q}_2(S_4-S_2).
$$

Noting that $\left(d_{ijt,4}(D'_4 D_4)^{-1}D'_4 X\right)'=(\overline{x}_{ij\cdot}-\overline{x})+(\overline{x}_{\cdot\cdot t}-\overline{x})$ and $\left(d_{ijt,2}(D'_2 D_2)^{-1}D'_2 X\right)'=(\overline{x}_{i\cdot\cdot}-\overline{x})+(\overline{x}_{\cdot j\cdot}-\overline{x})+(\overline{x}_{\cdot\cdot t}-\overline{x})$, we have

$$
\left(d_{ijt,4}(D'_4 D_4)^{-1}D'_4 X\right)'-\left(d_{ijt,2}(D'_2 D_2)^{-1}D'_2 X\right)'=\overline{x}_{ij\cdot}-\overline{x}_{i\cdot\cdot}-\overline{x}_{\cdot j\cdot}+\overline{x}
$$

31

and by Assumptions A.3 and A.5 and the Cauchy-Schwarz inequality

$$
\begin{aligned}
\bar{Q}_4 - \bar{Q}_2 &= \frac{1-\rho}{NMT_1} \sum_{i,j,t} v_{ijt} \left(\bar{x}_{ij\cdot} - \bar{x}_{i\cdot\cdot} - \bar{x}_{\cdot j\cdot} + \bar{x} \right) \\
&= \frac{1-\rho}{NM} \sum_{i,j} \bar{v}_{ij\cdot} \left(\bar{x}_{ij\cdot} - \bar{x} \right) - \frac{1-\rho}{N} \sum_i \bar{v}_{i\cdot\cdot} \left(\bar{x}_{i\cdot\cdot} - \bar{x} \right) - \frac{1-\rho}{M} \sum_j \bar{v}_{\cdot j\cdot} \left(\bar{x}_{\cdot j\cdot} - \bar{x} \right) \\
&= O_p(T^{-1} + (MT)^{-1} + (NT)^{-1}) = O_p(T^{-1}).
\end{aligned}
$$

Then

$$
\begin{aligned}
Q_{4,2} - Q_{2,2} &= (\bar{Q}_4 - \bar{Q}_2) S_4 + \bar{Q}_2 (S_4 - S_2) \\
&= O_p \left(T^{-1} \right) O_p(T^{-1} + (NM)^{-1} + (NMT)^{-1/2}) \\
&\quad + O_p((MT)^{-1} + (NT)^{-1} + (MN)^{-1} + (NMT)^{-1/2}) O_p \left(T^{-1} \right) \\
&= o_p \left(T^{-1} \right).
\end{aligned}
$$

For $Q_{4,3} - Q_{2,3}$, we use the fact that $d_{ijt,4}(D_4'D_4)^{-1}D_4'U - d_{ijt,2}(D_2'D_2)^{-1}D_2'U = \bar{u}_{ij\cdot} - \bar{u}_{i\cdot\cdot} - \bar{u}_{\cdot j\cdot} + \bar{u}$, Assumptions A.1-A.2 and A.5-A.6, and the Cauchy-Schwarz inequality to obtain

$$
\begin{aligned}
Q_{4,3} - Q_{2,3} &= \frac{1-\rho}{NMT_1} \sum_{i,j,t} v_{ijt} (\bar{u}_{ij\cdot} - \bar{u}_{i\cdot\cdot} - \bar{u}_{\cdot j\cdot} + \bar{u}) \\
&= (1-\rho) \left\{ \frac{1}{NM} \sum_{i,j} \bar{v}_{ij\cdot} \bar{u}_{ij\cdot} - \frac{1}{N} \sum_i \bar{v}_{i\cdot\cdot} \bar{u}_{i\cdot\cdot} - \frac{1}{M} \sum_j \bar{v}_{\cdot j\cdot} \bar{u}_{\cdot j\cdot} + \overline{vu}) \right\} \\
&= \frac{1-\rho}{NM} \sum_{i,j} \bar{v}_{ij\cdot} \bar{u}_{ij\cdot} + O_p((MT)^{-1} + (NT)^{-1}).
\end{aligned}
$$

In sum, we have shown that $Q_4 - Q_2 = \frac{1-\rho}{NM} \sum_{i,j} \bar{v}_{ij\cdot} \bar{u}_{ij\cdot} + o_p(T^{-1})$.

(ii) This basically follows from the proof of Lemma B.2(iii) and (iv).
(iii) This basically follows from the proof of Lemma B.2(iii) and (vii).
(iv) This basically follows from the proof of Lemma B.2(iv) and (vii).
(v) This basically follows from the proof of Lemma B.2(v) and (vi).
(vi) This basically follows from the proof of Lemma B.2(v) and (vii).
(vii) This basically follows from the proof of Lemma B.2(vi) and (vii). ∎

Proof of Lemma B.4. (i) Note that $\tilde{z}_{ijt,1} = x_{ijt} - \rho x_{ij,t-1}$ and by Assumption A.1 (iv)-(v), we have $L_1 = \frac{1}{NMT_1} U'X \left(X'X \right)^{-1} \sum_{i,j,t} (x_{ijt} - \rho x_{ij,t-1}) \cdot (x_{ijt} - \rho x_{ij,t-1})' \left(X'X \right)^{-1} X'U = O_p((NMT)^{-1})$.
To prove (ii)-(vii), we first apply Lemma A.1 to obtain the following decomposition

$$
\begin{aligned}
L_m &= (\zeta_{m,1}', \zeta_{m,2}') \frac{1}{NMT_1} \sum_{i,j,t} \tilde{z}_{ijt,m} \tilde{z}_{ijt,m}' \begin{pmatrix} \zeta_{m,1} \\ \zeta_{m,2} \end{pmatrix} \\
&= \zeta_{m,1}' \frac{1}{NMT_1} \sum_{i,j,t} \tilde{x}_{ijt,m} \tilde{x}_{ijt,m}' \zeta_{m,1} + \zeta_{m,2}' \frac{1}{NMT_1} \sum_{i,j,t} \tilde{d}_{ijt,m} \tilde{d}_{ijt,m}' \zeta_{m,2} \\
&\quad + 2\zeta_{m,1}' \frac{1}{NMT_1} \sum_{i,j,t} \tilde{x}_{ijt,m} \tilde{d}_{ijt,m}' \zeta_{m,2} \\
&\equiv L_{m,1} + L_{m,2} + 2L_{m,3}, \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad (\text{C.4})
\end{aligned}
$$

where $\zeta_{m,1} = (X'M_{D_m}X)^{-1}X'M_{D_m}U$ and $\zeta_{m,2} = (D_m'D_m)^{-1}D_m'U - (D_m'D_m)^{-1}D_m'X\zeta_{m,1}$ for $m = 2, ..., 7$. Note that $L_{m,1} = O_p(\|\zeta_{m,1}\|^2)$ whose exact order can be obtained from Lemma A.4 under Assumption

A.1(iv)-(v). Let $\xi^a_{ijt,m}$ be defined as in (C.2) whose expressions are given in (C.3) for $m = 2, ..., 7$. Note that

$$
\begin{aligned}
L_{m,2} &= \frac{1}{NMT_1} \sum_{i,j,t} (\xi^u_{ijt,m})^2 + \zeta'_{m,1} \frac{1}{NMT_1} \sum_{i,j,t} \xi^x_{ijt,m} \xi^{x\prime}_{ijt,m} \zeta_{m,1} - \frac{2}{NMT_1} \sum_{i,j,t} \xi^u_{ijt,m} \xi^{x\prime}_{ijt,m} \zeta_{m,1} \\
&\equiv L_{m,2}(1) + L_{m,2}(2) - 2L_{2,2}(3), \text{ say,}
\end{aligned}
\tag{C.5}
$$

and

$$
\begin{aligned}
L_{m,3} &= \zeta'_{m,1} \frac{1}{NMT_1} \sum_{i,j,t} \tilde{x}_{ijt,m} \xi^u_{ijt,m} - \zeta'_{m,1} \frac{1}{NMT_1} \sum_{i,j,t} \tilde{x}_{ijt,m} \xi^{x\prime}_{ijt,m} \zeta_{m,1} \\
&\equiv L_{m,3}(1) - L_{m,3}(2), \text{ say.}
\end{aligned}
\tag{C.6}
$$

We have several key observations under Assumptions A.1-A.3 and A.5-A.6: (1) $\frac{1}{NMT_1} \sum_{i,j,t} \xi^x_{ijt,2} \xi^{x\prime}_{ijt,2} = O_p(1)$, (2) $\frac{1}{NMT_1} \sum_{i,j,t} \tilde{x}_{ijt,m} \xi^{x\prime}_{ijt,m} = O_p(1)$, (3) both $\frac{1}{NMT_1} \sum_{i,j,t} \xi^x_{ijt,m} \xi^u_{ijt,m}$ and $\frac{1}{NMT_1} \sum_{i,j,t} \tilde{x}_{ijt,m} \xi^u_{ijt,m}$ have the same probability order as $\left\| \zeta_{m,1} \right\|$. As a result,

$$
L_m = L_{m,2}(1) + O(\left\| \zeta_{m,1} \right\|^2).
$$

By (C.3) and Assumptions A.1(iv), A.2, and A.5(iv) we have

$$
\begin{aligned}
L_{2,2}(1) &= \frac{1}{NMT_1} \sum_{i,j,t} [(1-\rho)(\overline{u}_{i\cdot\cdot} + \overline{u}_{\cdot j\cdot} - 2\overline{u}) + (1-\rho L)(\overline{u}_{\cdot\cdot t} - \overline{u})]^2 \\
&= \frac{(1-\rho)^2}{NM} \sum_{i,j,t} (\overline{u}^2_{i\cdot\cdot} + \overline{u}^2_{\cdot j\cdot}) + \frac{1}{T_1} \sum_t [(1-\rho L)\overline{u}_{\cdot\cdot t}]^2 + O_p((NMT)^{-1}),
\end{aligned}
$$

$$
\begin{aligned}
L_{3,2}(1) &= \frac{1}{NMT_1} \sum_{i,j,t} [(1-\rho)(\overline{u}_{ij\cdot} - \overline{u}) + (1-\rho L)(\overline{u}_{\cdot\cdot t} - \overline{u})]^2 \\
&= \frac{(1-\rho)^2}{NM} \sum_{i,j} \overline{u}^2_{ij\cdot} + \frac{1}{T_1} \sum_t [(1-\rho L)\overline{u}_{\cdot\cdot t}]^2 + O_p((NMT)^{-1}),
\end{aligned}
$$

$$
\begin{aligned}
L_{4,2}(1) &= \frac{1}{NMT_1} \sum_{i,j,t} [(1-\rho)(\overline{u}_{ij\cdot} - \overline{u}) + (1-\rho L)(\overline{u}_{\cdot\cdot t} - \overline{u})]^2 \\
&= \frac{(1-\rho)^2}{NM} \sum_{i,j} \overline{u}^2_{ij\cdot} + \frac{1}{T_1} \sum_t [(1-\rho L)\overline{u}_{\cdot\cdot t}]^2 + O_p((NMT)^{-1}),
\end{aligned}
$$

$$
\begin{aligned}
L_{5,2}(1) &= \frac{1}{NMT_1} \sum_{i,j,t} [(1-\rho L)(\overline{u}_{\cdot jt} - \overline{u})]^2 \\
&= \frac{1}{MT_1} \sum_{j,t} [(1-\rho L)\overline{u}_{\cdot jt}]^2 + O_p((NMT)^{-1}),
\end{aligned}
$$

$$
\begin{aligned}
L_{6,2}(1) &= \frac{1}{NMT_1} \sum_{i,j,t} [(1-\rho L)[(\overline{u}_{i\cdot t} - \overline{u}_{\cdot\cdot t}) + (\overline{u}_{\cdot jt} - \overline{u})]]^2 \\
&= \frac{1}{NMT_1} \sum_{i,j,t} \{[(1-\rho L)\overline{u}_{i\cdot t}]^2 + [(1-\rho L)\overline{u}_{\cdot jt}]^2\} + O_p(N^{-2} + M^{-2} + (NMT)^{-1}),
\end{aligned}
$$

and

$$
\begin{aligned}
L_{7,2}(1) &= \frac{1}{NMT_1} \sum_{i,j,t} [(1-\rho)(\overline{u}_{ij\cdot} - \overline{u}_{i\cdot\cdot}) + (1-\rho L)[(\overline{u}_{i\cdot t} - \overline{u}_{\cdot\cdot t}) + (\overline{u}_{\cdot jt} - \overline{u}_{\cdot j\cdot})]]^2 \\
&= \frac{(1-\rho)^2}{NM} \sum_{i,j} \overline{u}^2_{ij\cdot} + \frac{1}{NMT_1} \sum_{i,j,t} [(1-\rho L)(\overline{u}_{i\cdot t} + \overline{u}_{\cdot jt})]^2 \\
&\quad + O_p(N^{-2} + M^{-2} + T^{-2} + (NMT)^{-1}).
\end{aligned}
$$

33

Summarizing the above results yields the conclusion in Lemma B.4. ∎

Proof of Lemma B.5. The proof follows from that of Lemma B.4 by keeping track some mutually cancelling terms as in the proof of Lemma B.3. ∎

D Nickell biases for the seven estimators of the slope coefficient in the panel AR(1) model

In this section, we study the Nickell biases for the seven estimators of the slope coefficients in the panel AR(1) model. We continue to use the notation defined in the main text and write Model m as follows

$$
\begin{aligned}
Y &= \iota_{NMT}\beta_0 + Y_{-1}\beta_1 + D_m\pi_m + U \\
&= X\beta + D_m\pi_m + U,
\end{aligned}
\tag{D.1}
$$

where D_m, π_m, and U are as defined in Section 2. For the panel AR(1) model, $\beta = (\beta_0, \beta_1)'$, and $X = (\iota_{NMT}, Y_{-1})$, where ι_{NMT} is an $NMT \times 1$ vector of ones, and Y_{-1} is an $NMT \times 1$ vector that stacks $y_{ij,t-1}$ for $t = 1, ..., T$. Note that we assume that y_{ij0}'s are observed here.

Let $\hat{\beta}_1^{(m)}$ denote the least squares dummy variable (LSDV) estimator $\hat{\beta}_1^{(m)}$ of β_1 in (D.1) based on Model m. Balazsi, Matyas and Wansbeek (2018, BMW hereafter) study the Nickell biases of these estimators when $N = M$ pass to infinity and T is fixed. In this special case, they show that the asymptote bias of $\hat{\beta}_1^{(2)}$ is $O\left(\frac{1}{NT}\right)$, that of $\hat{\beta}_1^{(5)}$ and $\hat{\beta}_1^{(6)}$ is zero, whereas the asymptotic biases of $\hat{\beta}_1^{(m)}$, $m = 3, 4, 7$, are $O\left(\frac{1}{T}\right)$ and share the the same dominant term. In our setup, we allow N to be different from M and both diverge to infinity jointly with T. Based on Model m in (D.1), the estimator of β_1 is given by

$$
\hat{\beta}_1^{(m)} - \beta_1 = (Y_{-1}' M_{D_m}^* Y_{-1})^{-1}(Y_{-1}' M_{D_m}^* U), \ m = 1, ..., 7,
$$

where $M_{D_1}^* \equiv M_0 = I_{NMT} - \frac{\iota_{NMT}\iota_{NMT}'}{NMT}$, $M_{D_m}^* = M_0 - P_{D_m}$ for $m = 2, ..., 7$, and $P_{D_m} = D_m \left(D_m' D_m\right)^{-1} D_m$. Note that $M_{D_m}^*$ is a projection matrix due to the orthogonality between D_m and ι_{NMT}. Following the calculations in BMW (Section 5.1), we can show that the leading term of $E(\hat{\beta}_1^{(m)}) - \beta_1$ is given by

$$
\frac{E(\text{tr}(Y_{-1}' M_{D_m}^* U))}{E(\text{tr}(Y_{-1}' M_{D_m}^* Y_{-1}))} = \frac{1 - \beta_1^2}{2\beta_1}\left(\frac{\text{tr}(M_{D_m}^* \Psi) - \text{tr}(M_{D_m}^*)}{\text{tr}(M_{D_m}^* \Psi)}\right),
$$

where $\Psi = I_{NM} \otimes \Psi_0$ and

$$
\Psi_0 = \begin{pmatrix}
1 & \beta_1 & \cdots & \beta_1^{T-1} \\
\beta_1 & 1 & \ddots & \vdots \\
\vdots & \ddots & \ddots & \beta_1 \\
\beta_1^{T-1} & \cdots & \beta_1 & 1
\end{pmatrix}.
$$

Please note that due to the differences in the construction of the dummy variables in this paper and in BMW, our $\text{tr}(M_{D_m}^*)$ and $\text{tr}(M_{D_m}^* \Psi)$ are different from that of BMW (Table 2).

By Lemma A.2(iii)-(v) and the fact that $\text{tr}(A_1 \otimes A_2) = \text{tr}(A_1)\text{tr}(A_2)$, we have that

$$
\begin{aligned}
\text{tr}(P_I) &= N - 1, \ \text{tr}(P_J) = M - 1, \ \text{tr}(P_T) = T - 1, \\
\text{tr}(P_{IJ}) &= NM - 1, \ \text{tr}(P_{JT}) = MT - 1, \\
\text{tr}(P_{IT}^*) &= T(N - 1), \ \text{tr}(P_{IJ}^*) = N(M - 1), \ \text{and} \ \text{tr}(P_{JT}^*) = M(T - 1).
\end{aligned}
$$

With these results, we can show that

$$
\begin{aligned}
\text{tr}(P_I\Psi) &= (N - 1)\theta_T, \ \text{tr}(P_J\Psi) = (M - 1)\theta_T, \ \text{tr}(P_T\Psi) = T - \theta_T, \\
\text{tr}(P_{IJ}\Psi) &= (NM - 1)\theta_T, \ \text{tr}(P_{JT}\Psi) = MT - \theta_T, \\
\text{tr}(P_{IT}^*\Psi) &= T(N - 1), \ \text{tr}(P_{IJ}^*\Psi) = N(M - 1)\theta_T, \ \text{and} \ \text{tr}(P_{JT}^*\Psi) = M(T - \theta_T),
\end{aligned}
$$

where $\theta_T = \mathrm{tr}(\frac{1}{T}J_T\Psi_0) = 1 + \frac{2\beta_1}{1-\beta_1}\left(1 - \frac{1}{T}\frac{1-\beta_1^T}{1-\beta_1}\right)$. The results of $\mathrm{tr}(M_{D_m}^*)$ and $\mathrm{tr}(M_{D_m}^*\Psi)$ can be summarized in Table C1.

Table D1: $\mathrm{tr}(M_{D_m}^*)$ and $\mathrm{tr}(M_{D_m}^*\Psi)$ for model $m = 1, ..., 7$

Model	$\mathrm{tr}(M_{D_m}^*)$	$\mathrm{tr}(M_{D_m}^*\Psi)$
1	$NMT - 1$	$NMT - \theta_T$
2	$NMT + 2 - (N + M + T)$	$NMT - \theta_T(N + M - 2) - T$
3	$NMT - NM$	$NMT - \theta_T NM$
4	$NMT - NM - T + 1$	$NMT - \theta_T NM - T + \theta_T$
5	$NMT - MT$	$NMT - MT$
6	$NMT - MT - T(N - 1)$	$NMT - MT - T(N - 1)$
7	$NMT - (NM + MT + NT) + N + M + T - 1$	$NMT - \theta_T(1 + NM - N - M) - NT - MT + T$

Consequently, we can obtain the Nickell biases for $\hat{\beta}_1^{(m)}$ as summarized in Table D2 below.

Table D2: Nickell biases for Model m

Model	Biases: $\frac{1-\beta_1^2}{2\beta_1}\frac{\mathrm{tr}(M_{D_m}^*\Psi) - \mathrm{tr}(M_{D_m}^*)}{\mathrm{tr}(M_{D_m}^*\Psi)}$
1	$-(1 + \beta_1)\frac{1}{NMT} + o(\frac{1}{NMT})$
2	$-(1 + \beta_1)(\frac{1}{NT} + \frac{1}{MT}) + o(\frac{1}{NT} + \frac{1}{MT})$
3	$-(1 + \beta_1)\frac{1}{T} + o(\frac{1}{T})$
4	$-(1 + \beta_1)\frac{1}{T} + o(\frac{1}{T})$
5	0
6	0
7	$-(1 + \beta_1)\frac{1}{T} + o(\frac{1}{T})$

Note: Apparently, the bias $\frac{-(1+\beta_1)}{NMT}$ in Model 1 is always of smaller order than the variance term and thus can be always neglected asymptotically.

When the true model is given by Model 1 but one obtains the LSDV estimator $\hat{\beta}_1^{(m)}$ of β_1 based on Model m, the Nickell bias of $\hat{\beta}_1^{(m)}$ is identical to that given in Table C2 simply because Model 1 is nested in Models $2-7$.

REFERENCES

Balazsi, L., Mátyás, L. and Wansbeek, T. (2018). The estimation of multidimensional fixed effects panel data models. *Econometric Reviews* 37, 212-227.

Bernstein, D. S. (2005). *Matrix Mathematics: Theory, Facts, and Formulas with Application to Linear Systems Theory.* Princeton University Press, Princeton.

Greene, W. H. (2008). *Econometric Analysis.* 6th edition. Pearson/Prentice Hall, New Jersey.